UC San Diego
UC San Diego Electronic Theses and Dissertations

Title
HNF4 alpha bioactive ligands discovered by a high-throughput screen for modulators of the human insulin promoter

Permalink
https://escholarship.org/uc/item/64k4q3qr

Author
Kiselyuk, Alice Susan

Publication Date
2011-01-01

Peer reviewed|Thesis/dissertation
HNF4 ALPHA BIOACTIVE LIGANDS DISCOVERED BY A HIGH-THROUGHPUT SCREEN FOR MODULATORS OF THE HUMAN INSULIN PROMOTER

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

In

Bioengineering

by

Alice Susan Kiselyuk

Committee in charge:

Professor Shankar Subramaniam, Chair
Professor Fred Levine, Co-Chair
Professor Pamela Itkin-Ansari
Professor Palmer Taylor
Professor Shyni Varghese

2011
Copyright

Alice Susan Kiselyuk, 2011

All rights reserved.
The Dissertation of Alice Susan Kiselyuk is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

--------------------------------------------------

--------------------------------------------------

--------------------------------------------------

--------------------------------------------------

Co-Chair

Chair

University of California, San Diego

2011
# TABLE OF CONTENTS

Signature Page ....................................................................................................................................... iii
Table of Contents .................................................................................................................................. iv
List of Abbreviations .......................................................................................................................... viii
List of Figures ......................................................................................................................................... xi
List of Tables .......................................................................................................................................... xii
Acknowledgements ............................................................................................................................... xiii
Vita. ......................................................................................................................................................... xv
Abstract of the Dissertation .............................................................................................................. xvii

1 INTRODUCTION .......................................................................................................................... 1
  1.1 Scope of the Problem ................................................................................................................... 1
  1.2 Pancreas Function ....................................................................................................................... 3
  1.3 Diabetes Pathogenesis ................................................................................................................. 4
  1.4 Methods to Identify Regulatory Networks Important in Diabetes Pathogenesis: Selecting an Appropriate Platform ................................................................. 7
  1.5 Regulation of the Human Insulin Promoter ............................................................................. 10
  1.6 Liver-Enriched Transcription Factors that Play a Role in Diabetes Pathogenesis: HNF1 and HNF4α ............................................................................................................ 11
   1.6.1 HNF1α and HNF1β ............................................................................................................. 12
   1.6.2 HNF4α ............................................................................................................................... 14
  1.7 From Pancreas to Liver ................................................................................................................. 18
  1.8 Hepatocellular Carcinoma ......................................................................................................... 21
  1.9 Scope of the Dissertation ........................................................................................................... 22
  1.10 Figures ......................................................................................................................................... 25

2 PHENOTHIAZINE NEUROLEPTICS SIGNAL TO THE HUMAN INSULIN PROMOTER AS REVEALED BY A NOVEL HIGH-THROUGHPUT SCREEN ................................................................................................................................................................................................................................................. 27
  2.1 Abstract ....................................................................................................................................... 27
  2.2 Introduction ................................................................................................................................. 28
  2.3 Materials and Methods ............................................................................................................... 30
3.3 Methods ........................................................................................................ 61
  3.3.1 Cell Culture .......................................................................................... 61
  3.3.2 Compound library screening in T6PNE ............................................... 61
  3.3.3 Primary confirmatory assay for dose-responsiveness ...................... 62
  3.3.4 Counterscreen for estrogenic activity ................................................. 62
  3.3.5 Quantitative real-time PCR (Q-PCR) analysis .................................. 63
  3.3.6 Compounds ........................................................................................... 64
  3.3.7 In vitro analysis of ADME properties .................................................. 65
    3.3.7.1 Solubility .......................................................................................... 65
    3.3.7.2 Permeability ..................................................................................... 66
    3.3.7.3 Hepatic Microsome Stability ........................................................... 66
    3.3.7.4 Plasma Stability ............................................................................... 67
    3.3.7.5 Plasma Protein Binding ................................................................... 68
  3.3.8 PPRE-Luc reporter assays ................................................................. 69
  3.3.9 In silico docking studies ........................................................................ 70
  3.3.10 Quenching of intrinsic fluorescence of HNF4α aromatic amino acids
    Tyr/Trp ........................................................................................................ 71
  3.3.11 siRNA studies ..................................................................................... 71
  3.3.12 Chromatin immunoprecipitation (ChIP) assays ............................... 72
  3.3.13 Gene expression profiling ................................................................... 73
    3.3.13.1 Isolation of RNA and expression profiling ................................. 73
    3.3.13.2 Expression data analysis ................................................................. 73
  3.3.14 NextBio ................................................................................................. 74
  3.3.15 OTC transient transfection assay ...................................................... 74
  3.3.16 Liver panel profiling ............................................................................ 75
  3.3.17 Immunohistochemistry ....................................................................... 75
  3.3.18 Steatosis assay ............................................................................... 76
  3.3.19 Orthotopic xenograft mouse model ................................................. 76
  3.3.20 Mouse primary hepatocytes isolation .............................................. 77
  3.3.21 Statistical Analysis ............................................................................. 77
3.4 Results .......................................................................................................................... 78

3.4.1 Identification of a small molecule inhibitor of the insulin promoter:
BIM5078 .......................................................................................................................... 78

3.4.2 BIM5078 is structurally similar to FK614, a PPARγ agonist ....................... 78

3.4.3 BIM5078 binds directly to HNF4α and interacts with the ligand binding pocket in situ ......................................................................................... 79

3.4.4 BIM5078 mediates insulin promoter repression through HNF4α........ 80

3.4.5 BIM5078 disrupts binding of E47 and PDX-1 to the human insulin promoter .................................................................................................................. 81

3.4.6 BIM5078 affects the expression of known HNF4α target genes .......... 82

3.4.7 Fatty acids bound to HNF4α preferentially inhibit the human insulin promoter .................................................................................................................. 84

3.4.8 Development and characterization of lead compound BI6015 ............ 86

3.4.9 HNF4α antagonism induces loss of HNF4α expression and hepatic steatosis in vivo .................................................................................................................. 87

3.4.10 HNF4α antagonism induces fatty acid oxidation defects ................. 88

3.4.11 HNF4α antagonists are cytotoxic to human hepatocellular carcinoma (HCC) .................................................................................................................. 89

3.5 Discussion .................................................................................................................. 91

3.6 Acknowledgements ................................................................................................. 95

3.7 Figures ...................................................................................................................... 97

3.8 Tables ....................................................................................................................... 107

4 CONCLUSIONS .............................................................................................................. 109

4.1 Summary of work ................................................................................................. 109

4.2 Therapeutic implications ...................................................................................... 111

4.3 Future directions ................................................................................................. 112

References ...................................................................................................................... 114
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>%GFP+</td>
<td>Percent of green fluorescent protein positive cells</td>
</tr>
<tr>
<td>AF</td>
<td>Activation function</td>
</tr>
<tr>
<td>AMP</td>
<td>Adenosine monophosphate</td>
</tr>
<tr>
<td>AOX</td>
<td>Acyl-CoA oxidase</td>
</tr>
<tr>
<td>bHLH</td>
<td>Basic-helix-loop-helix</td>
</tr>
<tr>
<td>BMP</td>
<td>Bone morphogenic protein (BMP)</td>
</tr>
<tr>
<td>BrdU</td>
<td>Bromodeoxyuridine</td>
</tr>
<tr>
<td>CaPO₄</td>
<td>Calcium phosphate</td>
</tr>
<tr>
<td>CAT</td>
<td>Chloramphenicol acetyltransferase</td>
</tr>
<tr>
<td>CDKI</td>
<td>Cyclin-dependent kinase inhibitor</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>ChIP</td>
<td>Chromatin immunoprecipitation</td>
</tr>
<tr>
<td>CMV</td>
<td>Cytomegalovirus</td>
</tr>
<tr>
<td>DAPI</td>
<td>4',6-diamidino-2-phenyindole</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulfoxide</td>
</tr>
<tr>
<td>ECL</td>
<td>Electrochemiluminescence</td>
</tr>
<tr>
<td>eGFP</td>
<td>Enhanced green fluorescent protein</td>
</tr>
<tr>
<td>FACS</td>
<td>Fluorescence-activated cell sorting</td>
</tr>
<tr>
<td>FGF4</td>
<td>Fibroblast growth factor 4</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde 3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>GEO</td>
<td>Gene Expression Omnibus</td>
</tr>
<tr>
<td>HNF4α</td>
<td>Hepatocyte Nuclear Factor 4 alpha</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>HNF1</td>
<td>Hepatocyte Nuclear Factor 1</td>
</tr>
<tr>
<td>hTERT</td>
<td>Human telomerase reverse transcriptase</td>
</tr>
<tr>
<td>HTS</td>
<td>High-throughput screen</td>
</tr>
<tr>
<td>JDRF</td>
<td>Juvenile Diabetes Research Foundation</td>
</tr>
<tr>
<td>Kb</td>
<td>Kilobase pairs</td>
</tr>
<tr>
<td>K_d</td>
<td>Dissociation constant</td>
</tr>
<tr>
<td>LCFA</td>
<td>Long chain fatty acid</td>
</tr>
<tr>
<td>Luc</td>
<td>Luciferase</td>
</tr>
<tr>
<td>MCFA</td>
<td>Medium chain fatty acid</td>
</tr>
<tr>
<td>MER</td>
<td>Modified estrogen receptor</td>
</tr>
<tr>
<td>MAPK</td>
<td>Mitogen-activated protein kinase (MAPK)</td>
</tr>
<tr>
<td>MODY</td>
<td>Maturity onset diabetes of the young</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>MsigDB</td>
<td>Molecular signature database</td>
</tr>
<tr>
<td>Ngn3</td>
<td>Neurogenin 3</td>
</tr>
<tr>
<td>NIH</td>
<td>National Institute of Health</td>
</tr>
<tr>
<td>NJCC</td>
<td>NIH/JDRF Custom Collection</td>
</tr>
<tr>
<td>OTC</td>
<td>Ornithine carbamoyltransferase</td>
</tr>
<tr>
<td>Ptf1a</td>
<td>Pancreas transcription factor 1</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein data bank</td>
</tr>
<tr>
<td>PDX-1</td>
<td>Pancreatic duodenal homebox-1</td>
</tr>
<tr>
<td>PHHI</td>
<td>Persistent hyperinsulinemia and hypoglycemia of infancy</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>PI 3-kinase</td>
<td>Phosphoinositide 3-kinase</td>
</tr>
<tr>
<td>PP</td>
<td>Pancreatic polypeptide</td>
</tr>
<tr>
<td>PPARγ</td>
<td>Peroxisome proliferator-activated receptor gamma</td>
</tr>
<tr>
<td>PPRE</td>
<td>PPAR response element</td>
</tr>
<tr>
<td>Q-PCR</td>
<td>Quantitative real-time PCR</td>
</tr>
<tr>
<td>RFP</td>
<td>Red fluorescence protein</td>
</tr>
<tr>
<td>RIPA</td>
<td>Radioimmunoprecipitation assay</td>
</tr>
<tr>
<td>RL</td>
<td>Renilla luciferase</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Real Time - Polymerase chain reaction</td>
</tr>
<tr>
<td>SAR</td>
<td>Structure-activity relationship</td>
</tr>
<tr>
<td>SCFA</td>
<td>Short chain fatty acid</td>
</tr>
<tr>
<td>S.D.</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>S.E.M</td>
<td>Standard error of the mean</td>
</tr>
<tr>
<td>siRNA</td>
<td>Small interfering ribonucleic acid</td>
</tr>
<tr>
<td>shh</td>
<td>sonic hedgehog</td>
</tr>
<tr>
<td>TK</td>
<td>Thymidine kinase</td>
</tr>
<tr>
<td>VSV-G</td>
<td>Vesicular stomatitis virus-glycoprotein</td>
</tr>
<tr>
<td>Wt</td>
<td>Wild-type</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1.1: Insulin and glucagon expression in Islets of Langerhans ..................... 25
Figure 1.2: Schematic of the human insulin promoter and positions of the main cis-
acting regulatory components .............................................................................. 26
Figure 2.1: E47 expression ................................................................................ 50
Figure 2.2: Activation of E47 induces growth arrest in T6PNE cells ................. 51
Figure 2.3: Kip2 controls proliferation in T6PNE cells .................................... 52
Figure 2.4: Kip2 promoter responsiveness to E47 ............................................ 53
Figure 2.5: High-throughput screening for insulin promoter modulators .......... 54
Figure 2.6: Effect of phenothiazines on insulin promoter activity ................. 55
Figure 2.7: Phenothiazines do not exhibit estrogenic behavior .................... 56
Figure 2.8: SAR of phenothiazine compounds from the NIH/JDRF library .... 57
Figure 3.1: Identification of BIM5078, a regulator of insulin gene transcription .... 97
Figure 3.2: Only BIM5078 passed all confirmatory and counter screens .......... 98
Figure 3.3: BIM5078 represses insulin promoter activity through HNF4α antagonism ........................................................................................................ 99
Figure 3.4: BIM5078 disrupts binding of E47 and PDX-1 to the human insulin
promoter ............................................................................................................. 100
Figure 3.5: BIM5078 and BI6015 affect the expression of known and novel HNF4α
target genes ...................................................................................................... 101
Figure 3.6: Reported HNF4α ligands have variable effect on insulin promoter activity ........................................................................................................... 102
Figure 3.7: Blood chemistry of mice treated with BI6015 ............................... 103
Figure 3.8: HNF4α antagonism induces hepatic steatosis independent of fatty acid
oxidation defects ............................................................................................... 104
Figure 3.9: HNF4α antagonists are cytotoxic to hepatocellular carcinoma .... 105
Figure 3.10: NCI Panel .................................................................................... 106
LIST OF TABLES

Table 3.1: Summary of small molecule screening data.........................................................107
Table 3.2: Activity spectrum of compounds selected in the primary screen....................108
ACKNOWLEDGEMENTS

During my seven years as a graduate student, I have had the privilege of working with countless colleagues and collaborators across many disciplines. This dissertation is a cumulative representation of my work, which could not have been completed without the help of the people who advised and trained me.

First and foremost, I thank my advisor, Fred Levine, for his guidance, mentorship, advice, encouragement, and support. His door has always been open and his technical expertise has always been valuable and inspiring. Also deserving of special recognition for playing a direct role in my training and research are past and present members of the Levine Lab, including Seung-Hee Lee, Suzette Farber-Katz, Sonalee Athavankar, Ifat Geron, Tom Cohen, Li Huang, Ben Spangler and Pamela Itkin-Ansari. Their support has helped me in immeasurable ways.

I thank Shankar Subramaniam, my chair, who gave me invaluable input and guidance throughout my tenure in the Bioengineering Department. I thank the rest of my committee members, Pamela Itkin-Ansari, Palmer Taylor, and Shyni Varghese for their time, advice and support.

I thank Frank J. Gonzalez, Christopher Glass, Kees Murre, Michael German and Sam Okret for providing materials. I thank Roy Williams (Informatics Core) and Kang Liu (Genomics Core) for their support with the microarray studies. I thank the SBMRI Conrad Prebys Center for Chemical Genomics for support with high-throughput screening. I thank Anthony Pinkerton for providing insights into structural aspects of the HNF4α ligands.

Lastly, I want to give a heartfelt thank you to my family and friends. I am
especially grateful to my parents who have always encouraged and supported me. Also, thank you to my brother, Greg, and my fiancé, Adam, for being there for me every step of the way.

I would like to acknowledge the Sanford Children’s Health Research Center, the Department of Bioengineering at UCSD, the Skaggs School of Pharmacy and Pharmaceutical Sciences at UCSD, the UCSD Genetics Training Program (T32 GM008666) from the National Institute for General Medical Sciences, the UC Systemwide Biotechnology Research and Education Program Graduate Research and Education in Adaptive Bio-Technology Training Program, and the American Diabetes Association Clinical Scientist Training Award for their financial support.

Chapter 2, in part, is a reprint of the material as it appears in the Journal of Biomolecular Screening, 2010, with co-authors Alice Kiselyuk, Suzette Farber-Katz, Tom Cohen, Seung-Hee Lee, Ifat Geron, Berat Azimi, Suzanne, Heynen-Genel, Oded Singer, Jefferey Price, Mark Mercola, Pamela Itkin-Ansari, and Fred Levine. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in part, is currently being prepared for submitted for publication of the material as it may appear in Nature Chemical Biology, 2011, with co-authors Alice Kiselyuk, Seung-Hee Lee, Suzette Farber-Katz, Mingjun Zhang, Sonalee Athavankar, Tom Cohen, Anthony B. Pinkerton, Mao Ye, Paul Bushway, Adam Richardson, Heather A. Hostetler, Li Huang, Benjamin Spangler, Jennifer Higginbotham, John Cashman, Pamela Itkin-Ansari, Marcia I. Dawson, Friedhelm Schroeder, Yong Cang, Mark Mercola, and Fred Levine. The dissertation author was the primary investigator and author of this paper.
VITA

EDUCATION

2004  Bachelor of Science, Massachusetts Institute of Technology
2011  Doctor of Pharmacy, University of California, San Diego
2011  Doctor of Philosophy, University of California, San Diego

AWARDS

2009 – 2011  ADA Clinical Scientist Training Award
2007 – 2011  AAAS/Science Program for Excellence in Science
2006 – 2008  UC Systemwide Biotechnology Research & Education
Program Training Grant
2005 – 2006  UC NIH/NIGMS funded UCSD Genetics Training Program

PUBLICATIONS


INVITED ORAL PRESENTATIONS


FIELDS OF STUDY

Major Field: Bioengineering
   Studies in Diabetes and Molecular Biology
   Professor Fred Levine

Major Field: Chemical Engineering
   Studies in Biotechnology and Materials Science
   Professor Robert Langer

   Studies in Metabolic Engineering
   Professor Gregory Stephanopoulos.
ABSTRACT OF THE DISSERTATION

HNF4 ALPHA BIOACTIVE LIGANDS DISCOVERED BY A HIGH-THROUGHPUT SCREEN FOR MODULATORS OF THE HUMAN INSULIN PROMOTER

by

Alice Susan Kiselyuk

Doctor of Philosophy in Bioengineering

University of California, San Diego, 2011

Professor Shankar Subramaniam, Chair
Professor Fred Levine, Co-Chair

A number of diabetogenic stimuli interact to influence insulin promoter activity, making it an attractive target for both mechanistic studies and therapeutic interventions. Our lab developed a unique cell-based high-throughput screening assay for molecules that control insulin gene expression. The assay platform is based on an engineered cell line, T6PNE, derived from human fetal islets. T6PNE cells express insulin in response to inducible E47, an important insulin gene transactivator and were adapted for high-throughput assays using a lentiviral vector expressing the insulin-eGFP transgene. Through a screen of a subset of the ChemBridge Diverset small molecule library, we identified BIM5078, which represses insulin promoter activity. A cheminformatic analysis revealed structural similarity to a PPARγ ligand, FK614.
Consistent with that result, BIM5078 was found to be a PPARγ agonist when tested in a PPAR response element reporter assay. However, FK614 was inactive when tested for its ability to modulate the insulin promoter. In a search for other potential targets of BIM5078, we found that it binds strongly to HNF4α (Kd: 11.9 ± 2.9nM) and modulates known HNF4α target genes. Like BIM5078, siRNA to HNF4α in T6PNE reduces insulin gene expression, suggesting that BIM5078 is an HNF4α antagonist. In order to advance mechanistic studies and in vivo delivery of this novel class of potent synthetic HNF4α ligands, we developed a structurally related small molecule (BI6015) that selectively inhibits HNF4α by eliminating off-target effects mediated through PPARγ. As predicted by the phenotype observed in the liver-specific HNF4α knockout, both pharmacologic inhibition with BI6015 and genetic silencing of HNF4α results in accelerated lipid accumulation. Interestingly, BIM5078 and BI6015 were noted to be selectively cytoxic to cancer cell lines in vitro. In vivo, BI6015 induces apoptosis of transplanted human hepatocellular carcinoma cells but not the normal liver in an orthotopic xenograft model, although potency was limited by suboptimal pharmacokinetic properties. The discovery of bioactive ligands for HNF4α raises the possibility that diseases involving HNF4α, such as diabetes and cancer, might be amenable to pharmacologic intervention and provide a powerful tool to study the physiological role of HNF4α.
1 INTRODUCTION

1.1 Scope of the Problem

Diabetes is a major health concern, affecting greater than 8% of the U.S. population, which has climbed from 5% in 1990 [1]. This number is expected to continue to grow in coming years, resulting in not only a reduced quality of life directly associated with the chronic complications of diabetes but also a significant economic burden to individual patients and the U.S. healthcare system.

A major goal of our lab is to gain a better understanding of the pathophysiology underlying diabetes as well as to identify novel therapeutic targets. Diabetes is a multifaceted disease and can be difficult to treat despite an extensive number of treatment options. Currently, there are 9 classes of anti-diabetic agents available on the market, including 6 oral medications and 3 injectables. These pharmacologic agents have a wide variety of targets and mechanisms, some of which are capable of stimulating insulin release from pancreatic \( \beta \)-cells and others which act as insulin sensitizers by reducing hepatic glucose output and increasing uptake of
glucose in the periphery. In spite of this array of drugs available commercially, glycemic control remains difficult to achieve in diabetic patients and warrants the development of new therapies. One approach taken by our lab is to study the mechanisms and mediators that facilitate insulin gene regulation, as insulin sits at the very heart of diabetes, whether the result of its insufficient production or inadequate function.

Interestingly, many of the same signaling pathways that mediate insulin gene expression and have been found to play a role in diabetes pathogenesis have also been linked to maintaining the differentiated function of the liver. A group of phylogenetically unrelated liver-enriched transcription factors named hepatocyte nuclear factors (HNF) are particularly important in coordinating liver-specific expression of genes central to differentiated function, but also have roles in glucose homeostasis and fatty acid metabolism. In fact, defects in HNF4α, HNF1α and HNF1β are not only linked to hepatic dysregulation but also result in a form of early-onset diabetes known as Maturity Onset Diabetes of the Young (MODY) [2]. While there is no evidence to show that HNF1α and HNF1β are capable of being regulated by a ligand, HNF4α is an especially attractive target for pharmacologic stimulation. In addition to having a central role in maintaining the expression of genes important in β-cell and hepatocyte function, HNF4α is a member of the nuclear receptor superfamily, a class of proteins successfully targeted by pharmaceuticals at a high frequency. Thus, in searching for small molecule regulators of the human insulin gene, it was gratifying to find one that targeted HNF4α. The applications of such ligands directed at HNF4α are numerous, ranging from diabetes to atherosclerosis and cancer.
This dissertation focuses on the discovery of small molecules directed at HNF4α through a powerful high-throughput screening assay for modulators of insulin gene expression. We find that perturbations to HNF4α-dependent pathways using these ligands reveal transcriptional regulation of the human insulin promoter as well as selective toxicity to tumor cells. The discovery of novel transcriptionally active ligands directed at HNF4α suggests that diseases linked to HNF4α, including diabetes and cancer, might be amenable to pharmacologic intervention.

1.2 Pancreas Function

The pancreas is a dual function organ, having both exocrine and endocrine components. It produces digestive enzymes secreted by exocrine acinar cells as well as a number of hormones (insulin, glucagon, somatostatin, and pancreatic polypeptide, PP) secreted into the bloodstream by the endocrine cells. The endocrine pancreas forms many small clusters, collectively known as the Islets of Langerhans (Figure 1.1). Islets comprise 1-2% of the entire pancreas. Humans have approximately one million islets composed of 15-20% α-cells (produce glucagon), 60-80% β-cells (insulin), 5-10% δ-cells (somatostatin) and less than 2% PP cells (PP) [3]. The endocrine pancreas regulates blood glucose levels through the secretion of insulin. Upon release, insulin binds receptors on peripheral tissues that allow glucose transport into cells [3]. Without insulin, glucose remains in the blood, resulting in diabetes.
1.3 Diabetes Pathogenesis

Diabetes mellitus is the seventh leading cause of death in the U.S., affecting an estimated 25.8 million individuals, or 8.3% of the U.S. population in 2011. It is a major cause of heart disease and stroke and is the leading cause of kidney failure, limb amputations and blindness [1]. In 2007, an estimated $174 billion was spent on direct and indirect costs associated with diabetes [1].

Diabetes is a metabolic disorder characterized by elevated blood glucose levels. Classically, diabetes has been grouped into 3 categories: type 1 (failure to produce insulin, most commonly immune-mediated), type 2 (failure to use insulin appropriately, oftentimes leading to insulin deficiency in the chronic state) and gestational diabetes (features of type 2 diabetes present in pregnancy). However, there are a number of other causes of diabetes, including genetic defects in β-cell function (as seen with MODY diabetes), defects in insulin processing and even drug-induced diabetes (as occurs with glucocorticoids). At the very heart of these processes is a deficiency in the ability of insulin to function properly.

Diabetes results from the hyposecretion and/or hypoactivity of the hormone insulin. Central to the regulation of carbohydrate and fat metabolism, insulin facilitates glycogen storage in the liver and muscle, thereby playing a critical role in increasing “energy” availability to cells. In response to stimuli that includes protein and carbohydrates, β-cells begin to synthesis insulin in the form of a proinsulin precursor. Proteolytic cleavage by prohormone convertases and carboxypeptidases results in the release of C-peptide, which is used clinically as a marker to distinguish between type 1
and type 2 diabetes, and the active form of insulin. Mature insulin then gets released from β-cells in two phases, the first of which occurs rapidly in response to elevated blood glucose levels and a second, much slower phase involving the release from newly formed secretory vesicles.

The regulation of insulin expression and secretion can occur at a number of different steps, including the initial transcription of the protein, stabilization and/or translation of the insulin messenger RNA (mRNA), and post-translational modifications. Dysfunction at any of these stages of insulin synthesis or destruction of the β-cells themselves can result in impaired insulin secretion, and thereby diabetes. Alternatively, diabetes can be the result of impaired insulin sensitivity (insulin resistance). Models for insulin resistance suggest that defects can occur at the level of the insulin receptor or in related downstream signaling cascades, the result of increased lipid levels in the blood and increased intracellular fatty acid metabolites. In the first model, phosphorylation of serine/threonine sites on the insulin receptor result in decreased association with phosphoinositide 3-kinase (PI 3-kinase) and decreased activation of glucose transport [4]. Alternatively, a high level of fat in the blood, as occurs with obesity, stimulates accumulation of triglycerides in muscle cells. This can activate protein kinases which results in decreased glucose uptake independent of insulin levels [5].

The β-cell is both the effector cell in glucose-responsive insulin secretion and a major end organ in diabetes pathogenesis. A number of the pathways that play a role in β-cell dysfunction ultimately contributing to the progression of type 1 and type 2 diabetes affect the transcription of the insulin gene [6]. While type 1 diabetes results
from an absolute insulin deficiency, overt diabetes in type 2 diabetes results from progressive failure of β-cells to produce adequate levels of insulin to compensate for insulin resistance [7]. Initially, the β-cell mounts a compensatory response to impaired sensitivity, consisting of an increase in β-cell mass as well as enhanced insulin production [8]. As the disease progresses, the islets become desensitized to the persistently high levels of glucose and fail to aptly compensate [6]. This results in reduced insulin production, mediated in part through effects of insulin promoter transcription factors [9]. It has been reported that elevated serum glucose and lipid levels inhibit PDX-1, MafA and NeuroD activity, critically important transcription factors that activate the insulin promoter, postulated to occur through reactive oxygen species [10-12]. A similar phenomenon occurs early in the pathogenesis of type 1 diabetes, in which proinflammatory cytokines injure the β-cell through changes in gene expression needed for β-cell function, including reduced MafA and PDX-1 expression, and proapoptotic signals [13].

An understanding of the pathways that regulate diabetes progression is critical in developing novel therapeutic agents. Because a number of diabetogenic stimuli, including glucotoxicity and lipotoxicity, interact to affect insulin promoter activity, the factors that regulate the insulin promoter are attractive targets for both mechanistic studies as well as the development of novel therapeutic interventions.
1.4 Methods to Identify Regulatory Networks Important in Diabetes Pathogenesis: Selecting an Appropriate Platform

Insulin promoter transcription factors implicated in β-cell dysfunction and diabetes pathogenesis also contribute to β-cell neogenesis and have a role in the global maintenance of β-cell differentiated function. The most obvious example comes from pancreatic duodenal homeobox-1 (PDX-1), a central regulator of pancreatic organogenesis and β-cell maturation. The global PDX-1 knockout results in pancreatic agenesis, while the β-cell specific conditional knockout developed diabetes secondary to immature β-cell formation [14]. This is consistent with what is observed clinically, in which homozygous PDX-1 mutations prevent proper formation of the pancreas, while heterozygous mutations in PDX-1 in humans results in MODY4 [2]. Because PDX-1 forms a crucial link in the cascade of transcription factors that control appropriate expression of β-cell genes, including insulin, it was used to reprogram adult pancreatic exocrine cells into β-cells [15]. Together with MafA and Neurogenin 3 (Ngn3), PDX-1 was found to be capable of reprogramming pancreatic exocrine cells into functional β-cells, a process described as trans-differentiation, by delivering these regulators directly into the pancreas in vivo [15]. Based on this evidence, one would expect that therapies aimed at stimulating β-cell transcription factors have the potential to reverse disease progression by promoting β-cell neogenesis and preservation of differentiated function.
Adenoviral vector delivery of intact β-cell transcription factors, as was done above, is important for mechanistic and proof-of-principle studies but impractical with current technology as a therapeutic option. An alternative is to target important insulin promoter transcription factors that are capable of being regulated by ligands. This has prompted our pursuit of cell-based high-throughput screening assays to screen small molecule libraries for regulators of the insulin promoter.

The only cell in the body that expresses the insulin gene is the pancreatic β-cell. However, primary β-cells are in limited supply and have a tendency to undergo apoptosis when manipulated in culture. Therefore, a robust *in vitro* model of endocrine cell growth and differentiation would be extremely useful for studying insulin gene regulation. Although rodent insulinoma cell lines have been studied for many years, a comparative analysis of insulin gene promoters across species has shown surprising divergence between the rodent and human. The rodent insulin promoters have roughly 65% homology with their human counterpart in the 300bp region upstream of the transcription start site. Beyond those 300bp, homology declines to 30% [16, 17]. While the sequences are divergent, it has been argued that the factors bound are nearly identical, and that expression of the human insulin promoter is appropriate when expressed in transgenic rodents. This suggests that a sufficient number of cis-acting tissue-specific insulin gene enhancer elements are conserved between rat and human insulin genes [18].

It is possible, however, that while human and rodent insulin promoters do not exhibit significant physiological differences under normal conditions, phenotypic species specific differences may become apparent when the transcriptional network is
challenged by gene mutation or knockout. In fact, this may explain divergent phenotypes observed with regulatory networks implicated in β-cell biology. Of particular interest to us is that of the liver-enriched transcription factors, HNF1α, HNF1β and HNF4α. Heterozygous mutations in each of these is implicated in early onset diabetes in the human: MODY3 (HNF1α), MODY5 (HNF1β) and MODY1 (HNF4α). While these genes are part of an autoregulatory network present in both rodents and humans, heterozygous inactivation studies in the rodent have resulted in divergent phenotypes from what is clinically observed in the human. Mice with heterozygous mutations in HNF1α or HNF4α do not display signs of diabetes, in contrast to their MODY counterparts in humans. HNF1α+/− heterozygous mice do not exhibit any defects in insulin secretion or glucose tolerance [19], while patients affected with MODY3 have characteristic defects in insulin secretion [20]. Similarly, HNF4α+/− heterozygous mice are also normal [21], while MODY1 affected patients exhibit macrosomia, hyperinsulinemia and hypoglycemia at birth [22, 23] and ultimately impairment of glucose-stimulated insulin secretion by β-cells later in life. The reasons for these discrepancies remain unclear, though one could speculate that β-cells in mice have a greater capacity to compensate for the presence of just a single functional copy of HNF1α or HNF4α, unlike the human which exhibits haploinsufficiency in the adult [24]. Taken together, these phenotypic variances between the rodent and human suggest that studies done in rodents may not detect subtle variations of the normal and extrapolating conclusions from studies done in the rat or mouse should be done cautiously.
1.5 Regulation of the Human Insulin Promoter

The expression of the insulin gene exclusively in pancreatic β-cells is the result of a series of factors that stimulate transcription through the 300-400 base pair region upstream of the transcription start site, defined as the insulin promoter (Figure 1.2). Multiple sequences contained within the promoter region regulate insulin gene expression by serving as binding sites for proteins (transcription factors) that only bind to specific DNA sequences. Furthermore, a number of other signaling molecules, co-activators and repressors and the RNA polymerase complex contribute to complex regulatory network controlling the insulin gene. Many of the regulatory proteins involved in insulin gene expression have been characterized. These include both islet-specific factors (NeuroD, PDX-1 and MafA) as well as ubiquitous proteins such as E47 [25]. Several conserved motifs in the insulin promoter are widely believed to be responsible for restricting the expression of the insulin gene to β-cells. These include the E, A, and C1 elements [14, 16, 17].

The E boxes share the consensus sequence CANNTG and serve as binding sites for heterodimers of the basic helix-loop-helix family (bHLH) [26, 27]. In particular, heterodimerization occurs between the β-cell specific factor, NeuroD, and ubiquitous products of the E2A gene (E47 and HEB). The mouse NeuroD knockout did not develop mature islets and as a result became diabetic. Conversely, E2A disruption displayed no β-cell phenotype, suggesting redundancy in ubiquitous bHLH proteins. A boxes are regions of the promoter that are rich in AT sequences, where homeodomain family proteins bind. In particular, PDX-1 (pancreatic duodenal
homeobox-1), which is important in the development of the pancreas and in β-cell maturation, binds at this location [14]. PDX-1 knockout mice are unable to properly form a pancreas. The β-cell specific conditional knockout using the Cre-loxP system resulted in diabetic mice, while heterozygous mutations in PDX-1 in humans results in MODY4. Furthermore, interaction of these transcription factors that bind the E and A elements can act synergistically to activate insulin transcription. Additional regulatory sites have been evaluated as well, including the C element which binds transcription factor MafA, cyclic AMP response elements which bind a broad array of transcription factors, negative regulatory elements, enhancer sites and HNF1α which binds an A box adjacent to PDX-1 and is regulated by HNF4α. The aforementioned factors represent strongly conserved regulatory elements across species. In addition to these, there are a number of sites that appear to have diverged between species [14, 16, 17]. As previously discussed, the biological importance of divergent sites in the insulin promoter, including HNF4α, remains unclear. Subtle species-specific characteristics of the promoters may preclude detection of previously unrecognized regulators of insulin gene expression in the human when studied in the rodent, such as ligands for the liver-enriched transcription factor HNF4α.

1.6 Liver-Enriched Transcription Factors that Play a Role in Diabetes Pathogenesis: HNF1 and HNF4α

The liver-enriched transcription factors named hepatocytes nuclear factors (HNF) are a diverse, yet critically important group of proteins involved in directing
the expression of genes essential to normal liver function. Defects in these proteins have also been linked to diabetes, cancer, atherosclerosis, hemophilia and others. Mutations in three members of this set (HNF1α, HNF1β and HNF4α) are directly responsible for a form of early-onset diabetes known as Maturity Onset Diabetes of the Young (MODY) [2].

1.6.1 HNF1α and HNF1β

HNF1α and HNF1β are members of the variant homeodomain (homebox) containing family of proteins, each of which have three functional domains: 1) amino-terminal dimerization domain, 2) DNA-binding homeodomain-like domain and 3) carboxy-terminal transactivation domain. Both are expressed in the liver, kidney, pancreas and the gut [28] are central to the transcription of a myriad of liver-specific genes, including albumin, α-1 antitrypsin, apolipoproteins, fibrinogen and clotting factors [29-31]. Despite binding to palindromic sequences also recognized by HNF1α, the transactivation potential of HNF1β is about two-fold lower [32].

Complete genetic ablation of HNF1α in mice is associated with enlarged liver, decreased expression of liver-specific genes and defective insulin secretion from β-cells. Most striking however are defects in renal tube absorption of glucose (Fanconi syndrome). While a fraction of mice lacking HNF1α die soon after weaning from progressive wasting syndrome, the majority of these mice die from renal failure [19]. Surprisingly, the blood glucose levels of heterogeneous HNF1α+/- mice are normal and do not exhibit insulin secretion defects, unlike the human counterparts, which exhibit MODY3, as previously described.
HNF1α has been shown to be a weak transactivator of the rat insulin 1 gene [33] and the human insulin gene [24]. In fact, it was found by positional cloning that mutations in HNF1α are responsible for MODY3, suggesting a role for this gene in glucose regulation and β-cell function in humans [20]. Studies carried out by Okita et. al. demonstrate about 30-fold activation of a fragment of the human insulin promoter by HNF1α and HNF1β. They go on to show significantly reduced insulin expression in the presence of HNF1α carrying mutations identified in MODY patients [20]. This suggests that decreased insulin content observed in MODY may be, at least in part, attributed to reduced insulin gene transcription.

Contrary to what occurs with the HNF1α global knockout mouse, mice with a deficiency in HNF1β are embryonic lethal [34]. The β-cell conditional knockout of HNF1β exhibited impaired insulin secretion and glucose tolerance as well as increased HNF1α and PDX-1 expression with decreased HNF4α mRNA [35]. These results indicate that HNF1β is important in the transcriptional network in β-cells.

Despite the findings that HNF1α and HNF1β are important in the pathogenesis of diabetes and the transcriptional network governing β-cell biology, there is no evidence to suggest they are capable of being regulated by a ligand, making them unattractive targets for pharmacologic manipulation. HNF4α, however, does bind ligands, and also plays an equally important regulatory role in liver function and diabetes pathogenesis.
1.6.2 HNF4α

HNF4α was originally identified as a member of the nuclear receptor superfamily of transcription factors capable of binding DNA response elements central to hepatic gene expression [36]. In the adult, HNF4α is expressed in liver hepatocytes, kidney, small intestine, colon and pancreatic β-cells [37]. HNF4α is an important regulator of gene expression in the visceral endoderm and is necessary during gastrulation. During development, HNF4α is required to maintain a viable embryo; the HNF4α homozygous knockout mouse has an embryonic lethal phenotype during gastrulation [38, 39], consistent with there being no reports of homozygous mutations in human HNF4α [40].

As previously described, mice with heterozygous mutations in HNF4α do not display signs of diabetes, in contrast to humans with heterozygous inactivating mutations in HNF4α who develop MODY1 [22, 23]. The reason for this discrepancy has not been elucidated. Targeted homozygous deletion of HNF4α has been studied in the adult mouse liver and β-cells using a conditional cre-lox strategy. The conditional liver HNF4α knockout shows lipid accumulation in the liver, reduced serum cholesterol and triglyceride levels and increased serum bile acid concentrations [41]. Two β-cell specific knockouts have been generated and confirm a role for HNF4α in regulating insulin secretion. Klaus Kaestner’s group at the University of Pennsylvania showed that mice with β-cell-restricted knockout of HNF4α exhibited lowered blood glucose levels in both the fed and fasting states, elevated plasma insulin levels and decreased Kir6.2 expression in the islets of their HNF-4α mutant mice [42]. Using the same RIP-Cre transgene, Shimomura’s group in Japan generated HNF4α β-cell
specific knockouts, although different background strains of mice were used. They too found impaired glucose-stimulated insulin secretion but pancreatic islet morphology, β-cell mass, and insulin content were normal, as was Kir6.2 expression [43]. No definitive explanation is available for these discrepancies between the knockouts, but differences in strains and differential mRNA processing have been suggested.

The role that HNF4α plays in regulating the insulin promoter and how this relationship contributes to the MODY1 phenotype is not entirely clear. The HNF4α-binding site in the rodent insulin I promoter is located between nucleotides -57 and -69 and is important for promoter activity [44]. It is adjacent to a PDX1-binding site and in close proximity to the E1 site, both of which are crucial for promoter activation. In 2002, Michael Walker’s group at the Weizmann Institute of Science in Israel showed that HNF4α is able to directly binds and activates the rat insulin promoter [44], suggesting regulation of the insulin promoter by HNF4α occurs directly through a previously unrecognized HNF4α consensus sequence, in addition to the already known indirect action of HNF4α on the insulin promoter through HNF1α. However, this direct relationship has not proven to be conserved in humans. ChIP-ChIP data, which found that over 40% of genes expressed in hepatocytes and in the pancreas had promoters that bound HNF4α, many of which are involved in glucose and lipid homeostasis, [45-49], failed to show occupation of the human insulin promoter by HNF4α [45]. In spite of the discordance in the direct effect of HNF4α on the human insulin promoter, there is clearly a regulatory role for HNF4α in the pancreatic β-cell and on the human insulin promoter, as demonstrated both by heterogenous mutations resulting in MODY and our own data, in which we use a human cell based-assay to
find small molecules HNF4α ligands capable of regulating insulin gene expression. The sensitivity of our assay to HNF4α modulation, a feature that may have been missed if conducted in a rodent model, is more consistent with the human MODY1 phenotype than the heterozygous HNF4α mutation in mice.

Nuclear receptors, including HNF4α, are regulated by binding to small lipophilic ligands [50]. Ligand binding to the ligand binding domain (LBD) triggers changes in the conformation of the nuclear receptor to stabilize the activation function (AF) domain in the active position. These changes promote association with co-activators, important in transcriptional activation. The ligand binding pocket is formed by a combination of α-helices, β-strands and connecting loops that form a deep cavity in which lipophilic ligands can bind. The specificity and efficiency of ligand binding is largely the result of the size of the ligand binding pocket and the hydrophobic amino acids protruding from the α-helices forming the cavity [50]. Although HNF4α was clearly characterized to have all the functional domains consistent with being a nuclear receptor, the ligand remained unidentified – until recently.

HNF4α was long considered to be an orphan receptor. In 2002, Shoelson’s group at the Joslin Diabetes Center (Boston, Massachusetts) solved the x-ray crystal structure of the rat HNF4α ligand binding domain [51]. In doing so, they noted that in the absence of exogenously added ligand, the ligand binding pocket was occupied by fatty acids. Transfection studies with HNF4α have also found it to be transcriptionally active in the absence of ligand addition. Furthermore, although fatty acids were present in all HNF4α molecules, it was found that one molecule of each homodimer was in the active (open) conformation in which α helix 12 was extended while the
second molecule was in the active (closed) conformation with helix 12 in contact with the ligand binding domain. A model has since been proposed in which the constitutive occupation of the ligand binding pocket of HNF4α play a role in stabilizing the conformation of HNF4α but are unable to lock and regulate that protein’s transcriptional state [52]. Moreover, this suggests that there is a regulatory component in addition to the ligand that is involved in HNF4α activation. To address this question, Shoelson et. al. went on to solve the structure of human HNF4α bound to a free fatty acid ligand as well the co-activator SRC-1. In this ternary structure, both monomers of the homodimer adopted the active conformation, indicating that co-activator recruitment locks the complex in the active conformation [52].

The finding that HNF4α was transcriptionally active in the absence of ligand addition drew controversy in the field over whether HNF4α could be regulated by ligand interactions [51, 53, 54]. However, in 2009 studies in Drosophila demonstrated that a GAL4-dHNF4 ligand sensor could be activated by starvation or exogenous long chain fatty acids [46]. Almost simultaneously, Yuan et. al. used affinity isolation followed by mass spectrometry (AIMS) to reveal that linoleic acid (LA, C18:2ω6) was bound to HNF4α in the livers of fed but not fasted mice, suggesting a ligand could reversibly bind to the ligand binding pocket [55].

There have also been a number of reports suggesting that exogenous ligands targeting HNF4α have been discovered, including an acyl-CoA thioester (“Medica-16”) and a nitro-naphthofuran derivative [56, 57]. Although fatty acyl-coenzyme A thioesters have been proposed to modulate HNF4α activity [56], it has since been suggested that these molecules are too large for the HNF4α ligand binding pocket
More recently, a group in France has developed nitro-naphthofuran based ligands, which they report to bind the HNF4α ligand binding domain (LBD) and alter its ability to activate transcription. However, their prototype ligand, “compound 5” reduces both basal and HNF4α-induced transcriptional levels [57]. Its nonspecific effects on transcriptional regulation raise questions about its specificity and render it unfavorable as a targeted therapeutic. Thus, no truly potent and credible exogenous ligand for HNF4α has yet been proposed.

Because of its role in diabetes, hemophilia, atherosclerosis, hepatitis and cancer, synthetic small molecule ligands to HNF4α can be tremendously useful. They can not only facilitate defining the physiological role of HNF4α, but also serve as pharmacologic interventions, particularly in those organs in which HNF4α is highly expressed, such as the pancreas and liver.

1.7 From Pancreas to Liver

The molecular and cellular mechanisms that govern the development of the liver and pancreas draw striking parallels. The progenitors to these organs originate from the same embryonic endoderm, and, like all cell types, rely heavily on signal transduction cascades from nearby mesoderm cells to specify fate and initiate morphogenic budding from distinct domains of the embryonic foregut at 8.5 days of mouse gestation (3 weeks of human gestation) [59]. Early inductive signals are the same for the liver and pancreas, beginning with suppression of mesodermal Wnt and fibroblast growth factor 4 (FGF4) signaling in the foregut. Anterior-posterior
determination is established in the foregut by retinoic acid signals. At this point, patterning of the liver and pancreas occurs through a series of distinct signals that include bone morphogenic protein (BMP), mitogen-activated protein kinase (MAPK), FGF, sonic hedgehog (shh), PDX-1 and pancreas transcription factor 1 (Ptf1a) – the result of which is specification of hepatoblasts and pancreatic progenitors. The events that follow in promoting differentiation and organ morphogenesis are, for the most part, different for the liver and pancreas. Interestingly, the “liver-enriched” HNF transcription factors play an important role in promoting organogenesis and differentiation of both the liver and pancreas, as well as in regulating growth and function in mature cells within these organs. For instance, HNF6 plays a central role in fate specification of both liver and pancreatic progenitors and acts as part of a loop involving HNF1α, HNF1β and HNF4α. This was confirmed by Odom et. al., who used chromatin immunoprecipitation with promoter microarrays to detect genes occupied by HNF transcription factors including HNF6, finding that HNF6 serves as an important regulator of transcriptional circuits in hepatocytes and pancreatic islets [45]. Several of the other HNF transcription factors (HNF1α, HNF1β and HNF4α) have already been discussed, including the finding that greater than 40% of genes expressed in hepatocytes and in the pancreas had promoters that bound HNF4α [45], many of which are involved in glucose and lipid homeostasis [45-49]. All of these findings are consistent with the HNF transcription factors sitting at the heart of a complex network of feedback loops that maintain the differentiated state in hepatocytes and β-cells [45].
Despite acquisition of seemingly different functions upon receiving inductive signals, the liver and pancreas play a coordinated role in mediating cell metabolism and blood glucose homeostasis [59]. Hepatocytes are important in modifying nutrients and maintaining glucose levels through gluconeogenesis and glycolysis in response to glucagon released by α-cells in the pancreas. This hepatic glucose released during the fasted state prevents hypoglycemia. Conversely, β-cells in the pancreas regulate blood glucose by releasing insulin when triggered by stimulatory signals, including elevated glucose levels. This release, used as mechanism to store and provide energy, also helps to prevent hyperglycemia by suppressing hepatic gluconeogenesis and glycolysis and facilitating glycogen synthesis [60]. The intricacy with which the liver and pancreas coordinate metabolic function is consistent with convergence of liver-enriched HNF signaling transduction pathways.

Defects in any of these governing signaling cascades, however, can give rise to profound hepatic and pancreatic dysregulation, resulting in liver failure or diabetes. Deficiencies in the HNF1α and HNF4α circuitry are not only linked to Maturity Onset Diabetes of the Young (MODY), as previously described, but also can result in hepatic dysfunction and malignancy of the liver [61]. The importance of HNF1α and HNF4α as master regulators in multiple organ systems and contributors to disease pathogenesis makes it more compelling to therapeutic strategies that target this network of transcriptional regulators. Despite the findings that HNF1α is important in the pathogenesis of diabetes and the transcriptional network governing β-cell biology, there is no evidence to suggest that it is capable of being regulated by a ligand. HNF4α, however, has been shown to bind endogenous fatty acids. Differences in the
fed and fasted states have suggested that HNF4α is capable of being regulated by ligands, and our work advances this notion, with the development of therapeutically active HNF4α ligands.

1.8 Hepatocellular Carcinoma

HNF4α plays a crucial role in many disease processes. In MODY1, HNF4α is directly mutated and in type II diabetes it is strongly downregulated [62]. In other conditions, such as liver disease and carcinogenesis, HNF4α has an apparent, but less understood role. Specifically, an inconsistency exists in the literature with regards to the role of HNF4α in hepatocellular carcinoma (HCC). Some groups report down-regulation of HNF4α in the setting of HCC [63, 64], while others report its upregulation in HCC [61]. Clarifying the role of HNF4α in HCC can help identify new therapeutics aimed at the molecular basis for disease progression.

Hepatocellular carcinoma is the most prevalent liver tumor originating from hepatocytes, the main functional cell in the liver. HCC is the third leading cause cancer mortality worldwide, and the incidence has tripled between 1975 and 2005 [65]. Development of malignant HCC is associated with increased proliferation and dedifferentiation through a loss of epithelial morphology, as is common to all cancer development. A variety of signaling pathways important in maintaining liver function and promoting proliferation have been described, but the molecular basis of HCC progression has yet to be fully elucidated [61]. Hepatocyte nuclear factors maintain hepatocyte differentiation and control the expression of liver-specific genes, but its
role in liver tumorigenesis and HCC progression is poorly understood and has yielded conflicting reports in different biological systems.

Several groups have implicated HNF4α in hepatocellular carcinoma and other cancers, but the regulatory role it plays has yet to be understood. Some have shown that downregulation of HNF4α promotes tumorigenesis in hepatocellular carcinoma and other cancers [63, 64]. However, others demonstrate that HNF4α is upregulated in human hepatocellular carcinoma, in which researchers suggest that regulatory factors suppress the expression of C-EBPa but stimulate HNFs in HCC [61]. These results would be consistent with an acquired dependence of cancer cells on metabolic processes controlled by HNF4α. As a whole, the involvement of HNF4α in HCC is apparent, but how it does so remains ambiguous. Small molecules targeting HNF4α can provide a powerful tool for elucidating this discrepancy and for developing novel therapeutic agents for HCC.

1.9 Scope of the Dissertation

This dissertation describes the novelty and utility of HNF4α bioactive ligands discovered through a robust and highly sensitive high-throughput screening assay designed initially to identify regulators of the insulin promoter. The sensitivity of our human cell based assay to HNF4α modulation, a feature that may have been missed if conducted in a rodent model, is more reflective of the human MODY1 phenotype than the observed heterozygous HNF4α mutation in mice. HNF4α is a particularly attractive target for pharmacologic manipulation because it serves a central role in
maintaining β-cell and hepatocyte differentiated function and is implicated in a number of disease states, including diabetes and cancer. In searching for small molecule regulators of the human insulin gene, it was gratifying to find one that targeted HNF4α. We found that perturbations to HNF4α-dependent pathways with these ligands revealed transcriptional regulation of the human insulin promoter as well as selective toxicity to tumor cells, forming the basis for its therapeutic utility in hepatocellular carcinoma.

Chapter 2 describes the development a high-throughput screen to identify previously unrecognized regulators of the insulin promoter as a means to understand the underlying pathophysiology to diabetes and advance pharmacologic therapeutics in applicable fields. Because the insulin promoter is a target of multiple diabetogenic stimuli [66], we hypothesized that high-throughput screening for compounds that modulate the human insulin promoter activity may yield valuable insight into the pathogenesis of diabetes and also provide hits leading to novel therapeutic agents. To achieve this goal, we engineered a cell line derived from the human endocrine pancreas to express an insulin promoter eGFP-reporter transgene. Using this assay, we screened a library of known drugs and identified phenothiazine-based antipsychotics as modulators of insulin promoter activity. We show that a structurally diverse array of antipsychotics, spanning the typical and atypical classes, potently inhibit insulin gene expression following chronic exposure. This is consistent with what is observed clinically, as both generations of antipsychotics are known to have deleterious metabolic side effects, including diabetes.
The work detailed in Chapter 3 uses this powerful high-throughput screening tool as a means to identify a novel class of HNF4α ligands, capable of potently modulating insulin gene expression and exhibiting selective toxicity to a series of transformed cells, including hepatocellular carcinoma. Using the T6PNE insulin promoter assay to screen a diverse synthetic chemical library, we found a repressor of insulin gene expression to be an HNF4α antagonist. This small molecule, BIM5078, binds to HNF4α with high affinity and modulates HNF4α target genes. Moreover, BIM5078 and a structurally related analog, BI6015, are selectively cytotoxic to transformed cells in vitro. In vivo, BI6015 induces apoptosis of a human hepatocellular carcinoma cell line. The development of bioactive ligands for HNF4α, a goal attempted unsuccessfully by many in the pharmaceutical industry, suggests that diseases involving HNF4α, including diabetes and cancer, may be susceptible to pharmacologic intervention.

Chapter 4 summarizes these studies, their contributions to our understanding of the pathogenesis of diabetes, and the utility of HNF4α bioactive ligands as therapeutics for hepatocellular carcinoma. Further development of structurally related compounds could yield pharmacologic interventions for diseases associated with HNF4α, including diabetes and cancer, as well as providing a powerful tool to study the physiological role of HNF4α.
1.10 Figures

Figure 1.1: Insulin and glucagon expression in Islets of Langerhans
Islets are composed of 15-20% α-cells (producing glucagon, shown in red) and 60-80% β-cells (insulin, shown in green). They are also composed of 5-10% δ-cells (somatostatin, not shown) and less than 2% PP cells (PP, not shown) [used with permission from Seung-Hee Lee]
Figure 1.2: Schematic of the human insulin promoter and positions of the main cis-acting regulatory components

The human insulin gene is controlled by its promoter, located in the 300-400bp region upstream of transcription start site. The insulin promoter is highly regulated by a number of cis-acting elements, charted above. Positions are relative to the transcription start site at +1.
2 PHENOTHIAZINE NEUROLEPTICS SIGNAL TO
THE HUMAN INSULIN PROMOTER AS REVEALED
BY A NOVEL HIGH-THROUGHPUT SCREEN

2.1 Abstract

A number of diabetogenic stimuli interact to influence insulin promoter activity, making it an attractive target for both mechanistic studies and therapeutic interventions. High-throughput screening (HTS) for insulin promoter modulators has the potential to reveal novel inputs into the control of that central element of the pancreatic β-cell. A cell line from human islets in which the expression of insulin and other β-cell-restricted genes are modulated by an inducible form of the bHLH transcription factor E47 was developed. This cell line, T6PNE, was adapted for HTS by transduction with a vector expressing green fluorescent protein under the control of the human insulin promoter. The resulting cell line was screened against a library of
known drugs for those that increase insulin promoter activity. Members of the phenothiazine class of neuroleptics increased insulin gene expression upon short-term exposure. Chronic treatment, however, resulted in suppression of insulin promoter activity, consistent with the effect of phenothiazines observed clinically to induce diabetes in chronically treated patients. In addition to providing insights into previously unrecognized targets and mechanisms of action of phenothiazines, the novel cell line described here provides a broadly applicable platform for mining new molecular drug targets and central regulators of β-cell differentiated function.

### 2.2 Introduction

Although both type I and type II diabetes ultimately result in β-cell destruction, β-cell dysfunction precedes outright loss of β-cell mass, contributing to disease progression [6]. A number of the pathways that play a role in β-cell dysfunction affect insulin gene transcription. This includes fatty acids and glucose in type II diabetes and inflammatory mediators that are present early in the course of type I diabetes. In addition, transcription factors, including PDX-1, MafA, NeuroD1, and HNF1 that are important in insulin promoter transactivation also act globally to maintain many aspects of β-cell function and are targets for β-cell glucolipotoxicity [67]. Because of the central role of insulin promoter activity in β-cell function, stimulating or maintaining promoter activity in β-cell stem/progenitors or in mature β-cells has the potential to preserve and enhance β-cell function in diabetes.
To discover compounds that act on the insulin promoter, we developed a novel human cell line that mimics several important aspects of β-cell biology. TRM6, a cell line derived from human fetal islets, was previously engineered to express the homeodomain transcription factor PDX-1 and the bHLH factor NeuroD1, resulting in low levels of insulin promoter activity [68]. To further augment the degree of insulin promoter activity, we expressed the NeuroD1 dimerization partner, E47. E47 activation resulted in strong upregulation of insulin gene expression and other β-cell genes. E47 induction also resulted in cell cycle arrest mediated by Kip2, a β-cell specific CDKI (cyclin-dependent kinase inhibitor).

This cell line, termed T6PNE, was adapted for high-throughput screening by introducing a human insulin promoter-eGFP cassette. As an initial screen, we chose to study a library of known drugs for those that activated the insulin promoter. This had the advantage that many of the compounds have known mechanisms of action, unlike commonly used large and chemically diverse compound libraries. Moreover, many commonly used drugs act on pathways that are active in β-cells, but little is known about how those drugs affect β-cell function. Surprisingly, phenothiazine neuroleptics were found to modulate the insulin promoter, inducing an acute stimulatory response followed by repression upon chronic administration. This finding demonstrates a novel mechanism by which phenothiazines can affect glucose homeostasis and may explain the propensity of patients on these drugs to develop diabetes. T6PNE provides a simple in vitro model to study how pharmacological agents may influence β-cell function. The rapid responsiveness of the insulin promoter in this cell line to pharmacologic manipulation sets it apart from other β-cell models and makes it useful
for characterization of many aspects of β-cell biology that impinge on insulin promoter activity.

### 2.3 Materials and Methods

#### 2.3.1 Cell Culture

T6PN cell cultures [68] were infected with retrovirus expressing hTERT [69]. After 3 passages, the cell culture was further infected with VSV-G pseudotyped retrovirus expressing E47 as a fusion protein with the modified estrogen receptor (E47mer) (kind gift from Kees Murre [70]. The vector also contained the CD25 cell surface marker, permitting the use of FACS to sort infected cells. The cell line expressing telomerase and E47 was termed T6PNE. It was maintained in DMEM, and 5.5 mM glucose supplemented with 10% fetal bovine serum. To induce E47 activity in cells under experimental conditions, 4-hydroxy4-hydroxytamoxifen (0.5-10 µM) was added to the culture media. BrdU incorporation studies were performed with BrdU (GE) added to cells for up to 2 hours. For SAR studies compounds were added to cells for 48 hours. For chronic ethopropazine treatment, T6PNE cells were seeded in 0.5 µM 4-hydroxytamoxifen. At 24 hours either 15 µM ethopropazine or a vehicle control was added. At 96 hours the cells were passed into either a 384 well clear bottom black well plate for imaging or to a culture dish with fresh ethopropazine or vehicle. The same passing procedure was performed every 3 days for the remainder of the 12 day experiment (n=3, experiment was performed 3 times, error bars are standard error, *p-
value < 0.05). HeLa cells were maintained in DMEM (Gibco) and 10% FBS. Human islets were obtained from the NIH/JDRF Islet Cell Resource Center Program.

2.3.2 Vector Design and Transfection Studies

HeLa cells (ATCC) were transfected by CaPO₄ method with full length Kip2 promoter or Kip2 promoter deletion constructs provided Dr. S. Okret (Huddinge University Hospital, Sweden) [71]. To analyze E-box activity, HeLa cells were transfected with (4RTK-luc) [72] reporter and a plasmid encoding either wtE47 or E47mer [70]. Sixteen hours after transfection, the culture medium was changed and maintained for 48 hr with or without 4-hydroxytamoxifen.

Human Kip2 cDNA was cloned from pBlueskriptR (Open Biosystems) into the MSCV-IRES-GFP vector [73]. Lentivirus expressing the insulin promoter driving GFP was produced as previously described [74]. For lentivirus expressing shKip2, four different siRNA target sequences were chosen from the open reading frame of human Kip2 (accession no. BC067842), using the siRNA selection algorithm from sfold (http://sfold.wadsworth.org/).

An H1 promoter driven shRNA expression cassette was constructed by amplification of an oligonucleotide containing the entire shRNA, as described [75]. Each shRNA cassette was digested with XbaI and ligated into pBluescript. Validation of the construct was done by cotransfection with a plasmid expressing Kip2. The most effective shRNA cassette, as determined by Western blot, was transferred into a lentiviral vector for virus production [75, 76]. Kip2 oligonucleotides – shown is the start position of the siRNA target sequence and the sequence of the oligonucleotide
used for amplification: (shKip2-1448) 5'-CTGTCTAGACAAAAACCAAGGATGTAA
AGCTTTAATCTCTTGATTAAAGCCTTACACCTTGGGGGGATCTGTGGTCT
CATACA-3’

2.3.3 RT-PCR

Quantitative RT-PCR was performed on cDNA corresponding to 100 ng of RNA using the Opticon Real-Time System (MJ Research), with SYBR Green. RT-PCR for GAPDH was used for normalization. To measure the number of copies of mRNA, standard curves were constructed by PCR with a known number of copies of the target sequence.

Primers:

| Glucokinase     | for: gaagggaatgtgcccgtcgttgg  
|                 | rev: cacactggtcttcattggtcctc  |
| SUR-1           | for: cctcgtgctcaacacagaa       
|                 | rev: cagcttctctccttacgaa       |

Insulin, MafA, and GAPDH primers were described previously [77]

2.3.4 Western Blot

Whole-cell extracts were prepared by incubation in RIPA buffer containing protease inhibitors (Calbiochem). Protein (20µg) was separated on 4–20% Longlife gels (Lifegels) and transferred to Immobilon-P membrane (Millipore). After overnight blocking in PBS-Tween (PBST) with 3% milk, membrane was incubated with Kip2 antibody (Santa Cruz), followed by secondary antibody conjugated to horseradish peroxidase (Amersham/GE, Buckinghamshire, UK), and signal revealed by ECL (Amersham/GE, Buckinghamshire, UK).
2.3.5 ChIP Assay

T6PNE cells treated with or without 10µm 4-hydroxytamoxifen were treated with formaldehyde for in vivo cross-linking of Kip2 promoter and E47. Cross-linking was quenched by addition of glycine. The cells were then harvested by centrifugation resuspended in lysis buffer, according to EZ-ChIP protocol (Upstate).

Cellular DNA was sheared by sonication and immunoprecipitated with E47 antibody (Santa Cruz). Following purification, PCR was used to analyze immunoprecipitated DNA. PCR of 10ng DNA was performed with E-box oligonucleotide primers from the region between 4Kb and 3Kb in Kip2 promoter deletion constructs. PCR products were analyzed by electrophoresis on 1.2% agarose gels.

Primers:

<table>
<thead>
<tr>
<th>Primer</th>
<th>5’ Sequence</th>
<th>3’ Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1ST F</td>
<td>GGGCGTTACATA</td>
<td>CCACCAACCCAGCTAATTT</td>
</tr>
<tr>
<td>1ST R</td>
<td>5’</td>
<td>3’</td>
</tr>
<tr>
<td>2ND F</td>
<td>AAGCCTGGGGCAACATA</td>
<td>TAGAGGGCGTCTTGCTCTG</td>
</tr>
<tr>
<td>2ND R</td>
<td>5’</td>
<td>3’</td>
</tr>
<tr>
<td>3RD F</td>
<td>CAATCCTGGTTGAT</td>
<td>CAGTGTGTTTGCTCTGGAA</td>
</tr>
<tr>
<td>3RD R</td>
<td>5’</td>
<td>3’</td>
</tr>
<tr>
<td>4-6TH F</td>
<td>TTCCAGAGGCAACACACT</td>
<td>CATTCCTAGCCCTGTG</td>
</tr>
<tr>
<td>4-6TH R</td>
<td>5’</td>
<td>3’</td>
</tr>
</tbody>
</table>

2.3.6 Mutagenesis study

The 4Kb Kip2 deletion construct was used to make point mutations in each E-box between 4Kb and 3Kb by QuikChange II XL Site-Directed Mutagenesis kit (Stratagene). The PCR-based protocols were allowed to proceed for 18 cycles with each E-box point mutated primers. The product was treated with Dpn I restriction enzyme at 37°C for 1hr to digest the parental DNA template and to select for mutation
containing synthesized DNA. The mutated DNA was transformed in XL10-Gold Ultracompetent cells and purified with Endo-free Plasmid Maxi kit (Qiagen).

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kip2 promoter (Sac4.0) 1st E box mutation primer</td>
<td>5’-GGAGACCAAGGCCCCGGCTCGCTTG-3’ 5’-CAAGCGAGCCCGGCTTTGGTCTCC-3’</td>
</tr>
<tr>
<td>Kip2 promoter (Sac4.0) 2nd E box mutation primer</td>
<td>5’-GGTGTGGTGGAATTCCTGTGGTCC-3’ 5’-CCACACCACCTAAGGACACCAGG-3’</td>
</tr>
<tr>
<td>Kip2 promoter (Sac4.0) 3rd E box mutation primer</td>
<td>5’-TCGTTAAGCCCCCTTTCCAGGGG-3’ 5’-CCCCTGGGAAACCACGGCTTAACGA-3’</td>
</tr>
</tbody>
</table>

2.3.7 Luciferase and CAT Assays

Cells transfected with Kip2 promoter constructs driving a luciferase reporter gene as well as CMV-CAT were harvested with lysis buffer. Kip2 promoter activity was measured by Luciferase Reporter gene assay kit (Roche) and data was normalized to CAT activity (Roche) [68].

2.3.8 Immunohistochemistry

Immunostaining was performed on monolayers fixed with 4% paraformaldehyde and blocked with 5% donkey serum (Jackson ImmunoResearch Laboratories) and 1% bovine serum albumin (Sigma) in PBS. Primary antibodies were mouse monoclonal anti- E47 (Santa Cruz), mouse anti-Kip2 (Diagnostic Biosystems, mouse anti-BrdU (GE). Secondary antibodies were Rhodamine-red donkey anti-mouse (Jackson ImmunoResearch Laboratories), and fluorophore-labeled donkey anti-mouse (Alexa Flou 488, Molecular Probes, Inc.). Samples were mounted in Vectashield (Vector Labs, California, USA). Digital images of stained sections were captured using a fluorescence microscope with a digital camera (Nikon, Tokyo, Japan)
2.3.9 **Arrays and analysis**

Microarray analysis was used to determine genes expressed in T6PNE cells. RNA isolated from T6PNE cells treated with and without 4-hydroxytamoxifen was assayed for the expression of 24,000 human genes using Illumina BeadArray microarrays and Bead Studio software. Microarray data have been submitted in MIAME compliant format to the Gene Expression Omnibus (GEO) database (GSE18821).

2.3.10 **High Throughput Screening**

2.3.10.1 **Compound Library Screening**

T6PNE cells expressing the insulin-eGFP transgene were plated at 2,000 cells per well in a 384 well plate in the presence of 0.5μM 4-hydroxytamoxifen to induce a submaximal level of insulin expression. That dose of 4-hydroxytamoxifen was chosen to allow the screen to be particularly sensitive to compounds that upregulated insulin promoter activity. A higher dose of 4-hydroxytamoxifen would be optimal to detect compounds that inhibited the insulin promoter. Compound addition was performed 24 hours after 4-hydroxytamoxifen administration with the BiomekFX (Beckman Coulter). This was done by adding a single compound from the NIH/JDRF Custom Collection compound library (Micro Source Discovery Systems). All compounds were to a final concentration in the well of 5μM. Forty-eight hours after compound addition, the cells were fixed in 4% paraformaldehyde and stained with DAPI (0.167 μg/ml) to allow DNA content on a per cell basis (a surrogate for cell number) to be measured. The plates were stored in 50% glycerol.
2.3.10.2 Image Acquisition and Analysis

Images were obtained on a Beckman Coulter IC 100 high throughput microscopy system equipped with a Hamamatsu ORCA-ER scientific camera using a Nikon S Flour 10x NA 0.5 objective. Each 384-well plate was scanned and DAPI (40 ms integration time), GFP (200 ms integration time) and RFP channels were collected at one 3-color, 1280 x 1024 image per well.

Images of each well were captured in the blue channel to evaluate DAPI fluorescence as a measure of cell number, the green channel to assess expression of the insulin promoter-GFP transgene, and the red channel to assess autoflorescence. DAPI fluorescence produced a nuclear mask that allowed the microscope to focus from well to well on the plate.

Image processing (Cytoshop, Beckman Coulter, USA) was used to perform shade correction, nuclear segmentation and cytometry. Due to lamp variations throughout the imaging process, shade correction and plate-to-plate normalization techniques were utilized by equilibrating plate and global medians. Cell-by-cell analysis was performed by tessellation and pixel intensity measurements. The segmentation protocol involved the following steps: 1. The nucleus of each cell was identified using Cytoshop’s nuclear segmentation algorithm that included an “open” morphological operation. 2. Equidistance tessellation lines were drawn between the centroids of the identified nuclei, effectively breaking up the images into cellular regions. 3. An Object Extraction Correlation radius of 30 pixels was set, inside which pixels were assumed to belong to the cytoplasm of the cell.
Hits were then determined by applying threshold intensity gates on the green channel and surveying the number of cells above or below that gate in a given well by analyzing the average green pixel intensity under the cytoplasmic mask. These counts were then normalized to the total number of cells per well to give a percent GFP positive cells quotient. To construct the final assay-wide database from the individual plate databases, a Matlab algorithm was utilized to collate the individual plate data.

2.3.11 Chronic Ethopropazine Treatment

T6PNE cells were seeded in a 10cm dish with 0.5uM 4-hydroxytamoxifen. At 24 hours either 15uM Ethopropazine or a Vehicle Control was added. At 96 hours the cells were passed into either a 384 well clear bottom black well plate for imaging or a 10cm dish for continued culture along with fresh Ethopropazine or vehicle. The same passing procedure was performed every 3 days for the remainder of the 12 day experiment. The 384 well plates seeded for imaging were fixed in 4% Formaldehyde 24 hours after they were seeded. DAPI was then added at a final concentration of 167µg/ml. The plates were imaged on the IC 100 in the blue and green channels and Cytoshop analysis was then performed to determine the percent of cells in each well that contained GFP above a threshold determined by a MATLAB algorithm (% GFP+). Fold change in % GFP+ cells per well treated with Ethopropazine (15uM) versus vehicle control was measured over 12 days. (n=3, error bars are standard error, *p-value<.05).
2.3.12 Phenothiazine Structure Activity Relationship

T6PNE cells were seeded at 1750 cells per well in clear bottom black 384 well plates with 0.5µM 4-hydroxytamoxifen. At 24 hours, the phenothiazine or vehicle was added at a concentration range of 1.25-20µM. Cells were fixed with 4% formaldehyde at 72 hours. DAPI was then added at a final concentration of 0.167 µg/ml. The plates were imaged on the IC 100 in the blue and green channels and Cytoshop analysis was then performed to determine the percent of cells in each well that contained GFP above a threshold determined by a MATLAB algorithm (% GFP+).

2.4 Results

2.4.1 E47 activity is limiting for insulin gene expression in T6PN

TRM-6 is a cell line derived from human fetal islets. Engineering of TRM-6 to express the homeodomain transcription factor PDX-1 and the class II bHLH factor NeuroD1 (T6PN cells) resulted in the induction of a very low level of insulin expression [68]. High levels of the suppressor of E-box transactivators, Hes1, in cell lines that we have derived from human islets [77] suggested that the heterodimer partner of NeuroD1 on the insulin promoter, E47 [78], might be limiting insulin gene expression. As an initial test of that idea, we examined E47 expression in T6PN. Interestingly, the level of E47 mRNA was equal between T6PN and primary human islets (Figure 2.1A). E47 protein was undetectable by immunohistochemistry in either T6PN (Figure 2.1C, left panel) or primary islets (not shown).
To increase the level of E47, we overexpressed its mRNA in T6PN using a retroviral vector encoding E47\textsuperscript{MER}, in which E47 is fused to a modified estrogen receptor ligand binding domain (MER), rendering E47 inducible by 4-hydroxytamoxifen [79] (Figure 2.1B). In the absence of 4-hydroxytamoxifen, E47 was detectable at a very low level in these cells, termed T6PNE (Figure 2.1C, middle panel). In the presence of 4-hydroxytamoxifen, the level of nuclear E47\textsuperscript{MER} increased dramatically (Figure 2.1C, right panel). Induction of E47 by 4-hydroxytamoxifen resulted in a potent and dose-dependent upregulation of insulin gene expression (Figure 2.1D,E), but still approximately 1,000-fold lower than in primary β-cells, in which insulin mRNA constitutes close to 10% of the mRNA [80].

2.4.2 E47 induces multiple genes expressed in β-cells

To extend the finding that E47 plays a limiting role in insulin gene expression in T6PNE, a subset of genes that are important for β-cell differentiation and function were evaluated for E47 responsiveness. RT-PCR analysis of selected β-cell genes found that E47 induced glucokinase, SUR-1, and MafA (Figure 2.1E). The myosin heavy chain gene, which is cardiac specific and highly controlled by multiple E-boxes within its promoter [81], was not expressed in T6PNE and was not induced by 4-hydroxytamoxifen.

To probe the global pattern of gene expression in T6PNE in an unbiased fashion, oligonucleotide microarray analysis was performed on T6PNE in the presence and absence of E47 induction by 4-hydroxytamoxifen. Transcripts detected were then compared to a list of 747 genes enriched in pancreatic islets [82] and similar lists of
genes expressed in other tissues in the body derived from the Neurocrine Body Atlas (GSE 3526 of the GEO database) [83]. Of the 17,980 transcripts detected in T6PNE (p<0.05) and 18,482 transcripts detected in T6PNE treated with 4-hydroxytamoxifien (p<0.05), a total of 17,422 genes were expressed in both cell lines. 681 of the 747 β-cell specific genes were detected in T6PNE induced with 4-hydroxytamoxifien, suggesting inherently similar genetic profiling to β-cells.

To assess further the extent to which T6PNE retains a pattern of gene expression that predisposes it to endocrine differentiation in response to E47 induction, we compared data from the Illumina oligonucleotide microarray of T6PNE treated with 4-hydroxytamoxifien to that of a number of tissues including tongue, heart and adipocytes. These lists were derived from the Neurocrine Body Atlas by selecting differentially expressed genes when comparing the tissue of interest to all other tissues present in the atlas. Data obtained from 50 trials of randomly selected sets of 747 genes compared with genes expressed in T6PNE treated with 4-hydroxytamoxifien revealed a 74% average overlap by random chance. Similarly, comparisons of adipose-enriched, cardiac-enriched and tongue-enriched gene lists to 4-hydroxytamoxifien-treated T6PNE yielded 76%, 77% and 76% overlap, respectively. A Chi-square analysis demonstrated that T6PNE induced with 4-hydroxytamoxifien was more consistent with the transcriptional profile of β-cells, having a 91% intersection with β-cell enriched genes (p<0.0005), as compared with the other tissues studied, which were not statistically similar to islets (Tongue: p=0.28, Heart: p=0.14, Adipose: p=0.41).
2.4.3 E47 induces growth arrest mediated by upregulation of Kip2

As stated above, E47 was not constitutively expressed in T6PNE, but was introduced as a fusion protein with a modified estrogen receptor, rendering it inducible with 4-hydroxytamoxifen. This was done because E47 caused growth arrest in T6PN cells, as evidenced by decreased growth rate (Figure 2.2A) and BrdU incorporation (Figure 2.2B).

To determine the mechanism responsible for the growth arrest, the microarray data were examined for genes involved in cell cycle control that were induced by 4-hydroxytamoxifen. Most prominently, the cyclin dependent kinase inhibitor p57Kip2 was strongly induced in T6PNE by 4-hydroxytamoxifen (Figure 2.2C). This was confirmed by RT-PCR, demonstrating that 4-hydroxytamoxifen induction of E47 resulted in an increase in the level of Kip2 mRNA from almost undetectable to a level similar to that in islets. Interestingly, p57Kip2 is expressed in a β-cell restricted manner in the adult human (but not rodent) pancreas and is deleted in the β-cell hyperproliferative disorder focal PHHI (persistent hyperinsulinemia and hypoglycemia of infancy) [84].

To determine whether Kip2 was sufficient to induce growth arrest in T6PNE cells, it was overexpressed using a bicistronic retroviral vector also expressing GFP. In the absence of 4-hydroxytamoxifen (no E47 induction) 59% of the cells infected with a control virus expressing GFP alone incorporated BrdU. In contrast, no cells infected with the Kip2 and GFP expressing virus were positive for BrdU (Figure 2.3A-C).

To determine whether Kip2 was sufficient to account the growth arrest observed in T6PNE cells when E47 activity was induced, a bicistronic lentiviral vector
expressing both GFP and a Kip2 shRNA was used. The specific shRNA but not control, scrambled, shRNA suppressed Kip2 expression in 4-hydroxytamoxifen treated cells (Figure 2.3E). Even though knockdown of Kip2 by shRNA was incomplete and it is unknown how much Kip2 expression is required to inhibit proliferation, the degree of inhibition that was achieved resulted in significant reversal of the inhibitory effects of 4µM 4-hydroxytamoxifen on growth, increasing the percent of cells incorporating BrdU from 22 to 31% (p=.04) (Figure 2.3C).

To investigate the mechanism by which E47 induced Kip2 gene expression, we used co-transfection, ChIP assays, and site directed mutagenesis, finding that E47 directly activated the p57Kip2 promoter through a specific E-box between -3kb and -4kb. Seven E47 binding motifs within this region were identified but deletion analysis revealed that the 2nd E-box (E2) site was critical for promoter activity (Figure 2.4).

2.4.4 Adaptation of T6PNE for high-throughput screening (HTS)

The ability to precisely modulate insulin gene expression with 4-hydroxytamoxifen suggested that T6PNE cells could be used as a platform for HTS to detect compounds that affect insulin promoter activity. To adapt the cell line for HTS assays, T6PNE cells were infected with a lentiviral vector expressing the eGFP gene from a 1.4kb fragment containing the human insulin promoter [74, 85], providing for direct visualization of changes in insulin promoter activity. The specificity of the insulin promoter-eGFP transgene has been previously demonstrated [74].

4-hydroxytamoxifen dose-dependent eGFP fluorescence, precisely in concert with endogenous insulin mRNA level, was observed in T6PNE cells infected with the
transgene construct (Figure 2.5A). Therefore, we were able to prime T6PNE with an intermediate level of insulin promoter activity by administering low levels of 4-hydroxytamoxifen, allowing for detection of insulin promoter activators and inhibitors. Assay optimization of T6PNE cells in 384 well plates was performed using increasing doses of 4-hydroxytamoxifen from 0.5 to 4µM, yielding Z’ scores that ranged between 0.2 and 0.6. The variation in the Z’ occurred because the extent to which GFP expression was induced by 4-hydroxytamoxifen varied in different pilot studies for reasons that are unclear, but appears to be related to growth characteristics of the T6PNE cells.

2.4.5 High-Throughput Screening

Despite the variability in the Z’ values, we proceeded with a screen of a library of 1,040 known drugs from the NIH/JDRF Custom Collection (NJCC), from MicroSource Discovery Systems, Inc. Because of the variability, the primary screen was done in duplicate in separate plates. A number of compounds, including merbromin, acriflavinium hydrochloride, calcein, aklavine hydrochloride, isoreserpine, and pyrvinium pamoate, were found to exhibit autofluorescence and thus were not pursued further. Others were cytotoxic and increased GFP by creating fluorescent cell aggregates. As expected, known estrogenic compounds such as estrone, clomiphene citrate, estradiol diacetate, estrone acetate, diethylstilbestrol, estradiol cypionate, estiol, and estradiol acetate, were positive in the assay through their ability to activate the E47MER transgene.
Of the remaining compounds, three compounds were identified as increasing the number of GFP-positive cells in both wells from the primary screen. These compounds, chlorpromazine, ethopropazine, and berbamine were subjected to a primary confirmatory assay, consisting of a repeat of the initial assay but with a dose response and eight replicates for each dose. Image acquisition and analysis was performed as for the primary screen. All three of these compounds demonstrated dose-responsive increases in the percentage of GFP-positive cells.

2.4.6 **A subset of phenothiazines activate the insulin promoter**

The three compounds that passed the primary counterscreens, ethopropazine, chlorpromazine, and berbamine (Figure 2.5B), were subjected to a secondary assay consisting of quantitative RT-PCR for GFP and endogenous insulin mRNA. Berbamine increased GFP but not endogenous insulin mRNA and so was not studied further (Figure 2.6A). However, ethopropazine stimulated an increase in both GFP and insulin mRNA above 0.1% DMSO by \(3.86 \pm 0.33\) and \(2.89 \pm 0.13\) fold, respectively (Figure 2.6A). Chlorpromazine acted similarly, stimulating an increase in GFP and insulin mRNA of \(2.33 \pm 0.14\) and \(2.45 \pm 0.19\) fold, respectively (Figure 2.6A).

To ensure that the phenothiazines were not acting in a manner similar to the estrogenic compounds that were true but biologically uninteresting positives in the initial screen, we tested their ability to activate the insulin promoter in the absence of added 4-hydroxytamoxifen, finding that there was no effect under that condition (Figure 2.7).
2.4.7 Structure-Activity Relationship

Given that the NJCC library contained many phenothiazines, it was interesting that the only members of the class that were detected as hits in the primary screen were chlorpromazine and ethopropazine. To examine the structural specificity of phenothiazines on insulin gene expression further, we retested all of the phenothiazines using the methodology employed for the primary confirmatory screen described above. Results of this inquiry indicated that many phenothiazines in the NJCC library were weakly positive, but fell below the threshold value used in the high-throughput screen to detect hits. This analysis also revealed a false-negative in the assay, promethazine, which indeed turned out to activate insulin promoter activity (Figure 2.6B). Examining the structures of the positive and negative phenothiazines revealed that the most active compound was chlorpromazine, which differs from promazine only by the addition of a chlorine onto the carbon at the 3 position. This suggests that adding an electron-withdrawing group at that position increases activity. The increasing activity from promazine to ethopropazine suggests that increasing the hydrophobicity at the terminal amine by adding alkyl substituents also increases activity. The presence of a piperazine rather than an alkyl group at the terminal amine had a strong negative effect on activity. This suggests the possibility that this side group is limiting interaction with a specific receptor by steric hindrance (Figure 2.8).
2.4.8 Effect of Chronic Exposure to Ethopropazine

A puzzling aspect of the finding that some phenothiazines increase insulin expression in our assay is that this class of drugs has been reported to be diabetogenic in chronically treated patients [86]. To address that discrepancy, we investigated the effect of chronic exposure of T6PNE cells to ethopropazine. While the positive effect on insulin expression was confirmed early in the course of treatment, chronic exposure to the drug led not only to the loss of stimulatory activity, but to repression of insulin promoter activity (Figure 2.6C).

2.5 Discussion

This study describes the generation of a new model of human β-cell growth and differentiation and its application to high-throughput screening for compounds that affect the activation state of the insulin promoter, a key feature of the differentiated β-cell. The principal finding from this small screen of known drugs was that a subset of phenothiazine neuroleptics was able to acutely activate the insulin promoter, while chronic treatment led to downregulation of the insulin promoter.

A key to the development of the T6PNE β-cell line model described here was the finding that E47 controls both proliferation, through direct binding to the Kip2 promoter, and β-cell differentiation. Thus, selection for growth, as inevitably occurs in cells cultured \textit{in vitro}, resulted in downregulation of E47 activity and loss of β-cell differentiation, including insulin expression. Of note, the effect of E47 on insulin and other β-cell genes was replicated in another cell line that we have studied, blox5,
which also loses differentiated gene expression over time [87]. The low level of E47 activity observed in T6PNE compared with T6PN produces a small effect on proliferation, but this is sufficient to result in selection against maintenance of E47MER expression over time. Thus, it was important to start periodically with earlier passage or cloned T6PNE cells.

Although glucose-responsive insulin secretion is not seen with induction of E47 activity in T6PNE, the upregulation of many β-cell genes in this system provides a model to study some aspects of β-cell differentiation and modulation of the insulin promoter. The microarray analysis revealed that T6PNE cells express a large number of the genes that are preferentially expressed in islets. However, it is clear that the level of expression is much lower in T6PNE than in primary β-cells.

While for most purposes, the low level of expression of β-cell genes, particularly insulin, in T6PNE would be regarded as a drawback. However, it was precisely this low level expression that made it possible to use high-throughput screening for insulin promoter modulators. Detecting differences in expression from the insulin promoter-GFP transgene is straightforward in T6PNE because the level of GFP can be titrated by adjusting the 4-hydroxytamoxifen concentration. However, in primary β-cells or even conventional insulinoma cell lines such as MIN6, where the insulin gene is expressed at an enormous level, detecting the rapid change in GFP expression that is required for a practical high-throughput screen in primary islets is problematic.

It was surprising that the only class of drugs within the 1,040-compound library that influenced insulin gene expression was the phenothiazines, a class of drugs
known to predispose to diabetes upon chronic treatment [86]. It should be noted that the screen was of short duration, with only two days of exposure to compound. Prolonged treatment with ethopropazine in T6PNE led to repression of insulin promoter activity. This could result from desensitization of a pathway that plays an important role in β-cell function and could contribute to the deleterious effects on β-cells that are found in most published studies, in which phenothiazines have been shown to inhibit glucose-responsive insulin secretion [88]. Thus, as with other diabetogenic molecules such as fatty acids, acute exposure can stimulate aspects of β-cell function, while chronic exposure is harmful [67]. Unfortunately, it was not possible to perform such chronic exposure studies with primary islets due to toxicity of the compounds. Using lower doses that are not toxic would require much longer times of exposure, exceeding the time in which islets can be maintained in vitro. However, longer term in vivo studies of the effects of phenothiazines on islet function would be of value.

The targets of phenothiazine neuroleptics are diverse, having been shown to affect a variety of proteins, including dopamine receptors, protein kinase C, calmodulin, and others [89]. We have not as yet identified the target in T6PNE that signals to the insulin promoter. However, it is interesting to note that only a subset of phenothiazines were able to increase insulin promoter activity and that a specific structural feature of the terminal amine determined the level of activity. This suggests not only that a specific molecule is being targeted, but that it is likely to be a molecule that is different than those relevant to the psychiatric actions of the phenothiazines. There is poor concordance between the phenothiazines that we have found to act on
the insulin promoter with those that have been reported to inhibit insulin secretion [88]. With additional understanding of the target that is involved in insulin promoter modulation, it may be possible to design compounds that are more specific for the psychiatrically relevant target, with fewer effects on glucose homeostasis.

2.6 Acknowledgements

We would like to thank Kees Murre, Michael German and Sam Okret for providing materials. Steve Vasile and the Burnham Institute High Content Screening Core provided support with high-throughput screening. Roy Williams (Informatics Core) and Kang Liu (Genomics Core) provided support for the microarray studies. This work was supported by the Sanford Children’s Health Research Center (FL) and grants from the NIH (1R21NS057001 and R01 DK055283-08S1 to FL), Juvenile Diabetes Research Foundation (FL, PI-A), the UCSD Genetics Training Grant (AK, TC), a California Institute for Regenerative Medicine (CIRM) fellowship (S-HL), and the UC Systemwide Biotechnology Research and Education Program Graduate Research and Education in Adaptive Bio-Technology (GREAT) Training Program (AK).

Chapter 2, in part, is a reprint of the material as it appears in the Journal of Biomolecular Screening, 2010, with co-authors Alice Kiselyuk, Suzette Farber-Katz, Tom Cohen, Seung-Hee Lee, Ifat Geron, Berat Azimi, Suzanne, Heynen-Genel, Oded Singer, Jefferey Price, Mark Mercola, Pamela Itkin-Ansari, and Fred Levine. The dissertation author was the primary investigator and author of this paper.
2.7 Figures

Figure 2.1: E47 expression

(A) E47 mRNA was measured by RT-PCR in 3 independent preparations of T6PNE and 2 human islet donors. (B) E47mer is localized in the cytoplasm bound to hsp90. With 4-hydroxytamoxifen, E47 dissociates from hsp90 and translocates to the nucleus where it heterodimerizes with NeuroD1, driving insulin gene expression. (C) T6PNE and uninfected cells (T6PN) were assayed by immunohistochemistry for E47 (red) in the presence and absence of 4µM 4-hydroxytamoxifen. Without 4-hydroxytamoxifen, most cells did not have detectable E47 either in the cytoplasm or the nucleus. Cells with weak nuclear E47 was presumably due to leakiness of the modified estrogen receptor. With 4-hydroxytamoxifen, nuclear E47 was greatly increased. DAPI (blue). (D) Insulin expression was dose-responsive to 4-hydroxytamoxifen, saturating at 4µM. Error bars are SEM of 3 replicates. (E) E47 induction by 4-hydroxytamoxifen resulted in upregulation of glucokinase, SUR-1 and MafA. Assay was done twice and GAPDH control was equivalent between samples.
Figure 2.2: Activation of E47 induces growth arrest in T6PNE cells

(A) Cells cultured in the absence or presence of 4-hydroxytamoxifen (4 µM) were counted on the indicated days, demonstrating substantial growth arrest with the addition of 4-hydroxytamoxifen. (B) BrdU (green) and DAPI (blue) staining of T6PNE cells, demonstrating the absence of BrdU incorporation in the presence of 4-hydroxytamoxifen (4 µM). (C) Microarray analysis was used to determine genes in T6PNE cells that were affected by E47, with particular attention to CDKIs. T6PNE were treated with 4-hydroxytamoxifen (4 µM) and assayed for the expression of 24,000 human genes using Illumina BeadArray microarrays. Each dot on the scatter plot (power function) represents a single gene. The values on the X and Y axes represent the level of hybridization to the oligonucleotide on the array, which is an indirect measure of the level of mRNA in the sample. p57^{Kip2} was undetectable in the absence of 4-hydroxytamoxifen and was expressed at an approximately 2,000-fold higher level in cells in which E47 was induced by 4-hydroxytamoxifen. p21^{Cip1} was induced 2-fold.
Figure 2.3: Kip2 controls proliferation in T6PNE cells
A retroviral vector expressing GFP alone (A) or expressing GFP and Kip2 (B) was used to infect T6PNE cells in the absence of 4-hydroxytamoxifen. Cell cycle status of GFP-positive cells was assessed by BrdU staining (red). No GFP-BrdU double-positive cells were found in wells infected with the bicistronic GFP-Kip2 vector (B), while many cells infected with the GFP control vector did take up BrdU (arrows in A). DAPI is shown in blue. To inhibit Kip2 expression, T6PNE cells were transfected with a plasmid encoding a shKip2 (D) or scrambled (C) RNA along with GFP. The efficacy of the Kip2 shRNA construct was assessed by Western blot analysis (E). In T6PNE cells in which growth was inhibited by 4-hydroxytamoxifen (4 µM)-mediated induction of E47, transfection of the shKip2 but not control shRNA (visualized by co-co-expression of GFP) caused cell cycle re-entry as determined by colocalization of BrdU (red) and GFP (these are green cells with yellow nuclei. Six such cells are marked by white arrows in D).
Figure 2.4: Kip2 promoter responsiveness to E47

(A) A 6.3kb and several shorter fragments from the 5’ region of the Kip2 gene were cloned upstream of a luciferase reporter gene and luciferase activity measured in the absence and presence of 4µM 4-hydroxytamoxifen, with the finding that the region responsive to E47 was localized to a 900 bp fragment between 3 and 4 kb 5’ of the transcription start site. (B) This region was found to contain 7 E box elements in 4 distinct locations, with E4-E7 clustered together. ChIP analysis revealed weak binding to the E4-E7 cluster and strong binding to E1-3. (C) Site directed mutagenesis of E1 (green), E2 (red), E3 (blue) demonstrated that only E2 was required for E47-responsiveness.
Figure 2.5: High-throughput screening for insulin promoter modulators
(A) T6PNE cells infected with a lentiviral vector expressing an insulin promoter-eGFP cassette [74] exhibited a 4-hydroxytamoxifen dose-dependent increase in green fluorescence similar to that seen with endogenous insulin mRNA levels (Figure 1C). (B) Scatter plot of a high-throughput screen with T6PNE expressing the insulin promoter-GFP cassette. 1,040 known drugs from the NIH/JDRF Custom Collection were screened in duplicate. Each set of replicates is depicted as either red squares or blue triangles. The duplicates for each compound were assayed on separate 384 well plates, but the values for each compound are shown consecutively for ease of visualization. Duplicate values for the three compounds that passed the primary confirmatory and secondary assays (berbamine, chlorpromazine, and ethopropazine) are circled.
Figure 2.6: Effect of phenothiazines on insulin promoter activity

(A) Quantitative RT-PCR of GFP mRNA and endogenous insulin mRNA. Ethopropazine and chlorpromazine increased endogenous insulin mRNA and eGFP mRNA relative to untreated cells. Insulin mRNA levels were normalized to GAPDH. Error bars are SEM of three independent experiments. (B) Retesting of all phenothiazines in the NIH/JDRF library for effect on the insulin promoter. Values are expressed as fold change in GFP-positive cells relative to vehicle treated cells. Each compound is shown at its maximal tolerated dose, which was 20µM for all compounds except fluphenazine, perphenazine, and trifluoperazine, which were used at 10µM because of toxicity at higher doses (n=6, experiment was performed 3 times, error bars are standard error) (**p-value <0.01, *p-value<0.05) (C) Effect of chronic administration of ethopropazine on insulin promoter activity. Ethopropazine (15µM) was administered to T6PNE cells for 12 days. Media was changed every 3 days. A lower concentration of drug was used than in acute treatment to minimize toxic effects on the cells. Values are expressed as fold change in GFP-positive cells relative to vehicle treated cells. (n=3, error bars are S.E.M., *p-value<0.05).
Figure 2.7: Phenothiazines do not exhibit estrogenic behavior
T6PNE cells were exposed to 4-hydroxytamoxifen (Tam.) for three days and/or ethopropazine (Eth.) for the final 48 hours, followed by extraction of RNA and quantitative real-time RT-PCR for insulin and GAPDH. Insulin mRNA was normalized to GAPDH, which did not vary significantly between the samples. In the absence of 4-hydroxytamoxifen, insulin mRNA was not induced by ethopropazine, demonstrating that ethopropazine did not exhibit activation of the E47MER fusion protein in a manner similar to 4-hydroxytamoxifen. Error bars are S.D. of three biological replicates.
Figure 2.8: SAR of phenothiazine compounds from the NIH/JDRF library
Numbers in parenthesis are the fold change in GFP fluorescence in T6PNE cells induced by that compound (taken from Figure 2.6B).
3 HNF4α ANTAGONISTS DISCOVERED BY A HIGH-THROUGHPUT SCREEN FOR MODULATORS OF THE HUMAN INSULIN PROMOTER

3.1 Abstract

HNF4α is a central regulator of gene expression in cell types that play a critical role in metabolic homeostasis, including pancreatic β-cells, enterocytes, and hepatocytes. Although fatty acids occupy the HNF4α ligand-binding pocket, they have not been found to influence its transcriptional state, yielding controversy about whether HNF4α is regulated by ligand binding. Here, we report the discovery of a novel class of potent synthetic HNF4α ligands through a high-throughput screen for modulators of the human insulin promoter. These molecules bind to HNF4α with high affinity and modulate the expression of known HNF4α target genes. Notably, they
were found to be selectively cytotoxic to cancer cell lines \textit{in vitro} and \textit{in vivo}, although \textit{in vivo} potency was limited by suboptimal pharmacokinetic properties. The discovery of bioactive ligands for HNF4α raises the possibility that diseases involving HNF4α, such as diabetes and cancer, might be amenable to pharmacologic intervention.

### 3.2 Introduction

HNF4α is a member of the nuclear receptor superfamily of transcription factors and binds as a homodimer to a relatively degenerate consensus DNA sequence [90]. It is expressed at high levels in hepatocytes, enterocytes, pancreatic epithelial cells (including β-cells), and renal tubular epithelial cells [37, 91]. In these cells, it sits at the heart of a transcriptional regulatory network that controls the expression of many genes but particularly those involved in intermediary metabolism and maintenance of epithelial differentiation.

Because of the important role of HNF4α in regulating metabolic processes such as glucose and lipid homeostasis [21, 41, 45], and the high frequency with which nuclear receptor transcription factors have been targeted by pharmaceuticals, there has been great interest in developing synthetic ligands for HNF4α [56, 57]. Interestingly, medium and long chain fatty acids (MCFAs and LCFAs, respectively) are invariably found bound to the HNF4α ligand binding pocket in structural studies of HNF4α purified from bacteria [51, 52]. More recently, linoleic acid was found to be bound in the ligand-binding pocket of HNF4α purified from COS-7 cells. \textit{In vivo}, linoleic acid was present in the ligand-binding pocket of HNF4α in the fed state in mice and was
completely absent in the fasted state, suggesting that ligand binding is regulated [55]. However, there has been no evidence indicating that ligand binding influences the state of transcriptional activity mediated by HNF4α [55], limiting studies of HNF4α primarily to genetic deletion [38, 39, 41, 42, 92] or overexpression [93-95].

Previously, we described an assay for insulin promoter modulators based on a cell line derived from human fetal islets, T6PNE, which was engineered to express the β-cell transcription factors PDX-1, NeuroD1, and E47 (as a fusion protein with a modified estrogen receptor ligand binding domain to render it 4-hydroxytamoxifen-inducible; E47\textsuperscript{MER})[96]. Induction of E47 by 4-hydroxytamoxifen resulted in dose-responsive expression of the insulin gene, as well as a number of other genes expressed in β-cells. T6PNE cells were adapted for high-throughput screening by transduction with a lentiviral vector expressing green fluorescent protein under the control of the human insulin promoter [97].

Here, we used the T6PNE insulin promoter assay to screen a diverse synthetic chemical library. A screening hit, BIM5078, potently repressed insulin expression in that assay. BIM5078 bound to HNF4α with high affinity and modulated HNF4α target genes and metabolic processes controlled by HNF4α. Interestingly, BIM5078 and a related analog, BIM6015, were selectively cytotoxic to transformed cells in vitro. In vivo, BIM6015 induced apoptosis of a human hepatocellular carcinoma cell line. Further development of related compounds could lead to pharmacologic therapies for a variety of diseases, including diabetes and cancer, as well as providing a powerful tool to study the physiological role of HNF4α.
3.3 Methods

3.3.1 Cell Culture

T6PNE cells were maintained in RPMI (5.5 mM glucose) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin (pen-strep) and grown in 5% CO₂, 37°C. To induce E47 activity, 0.5 or 1 μM 4-hydroxy-4-hydroxytamoxifen (Sigma-Aldrich; MO, USA) was added to culture media. HepG2, Hep3B-luc, and HeLa cells were cultured in Dulbecco’s Modified Eagle Medium (DMEM) supplemented with 10% FBS and 1% pen-strep. HepG2 and Hep3B-luc cells were grown at 5% CO₂, 37°C. HeLa cells were grown at 10% CO₂, 37°C. 5-Bromo-2-deoxyuridine (BrdU) incorporation experiments were performed in Hep3B-luc cells using a 2 hours BrdU pulse (1:1000 dilution, GE Healthcare, UK), followed by fixation for immunohistochemistry.

3.3.2 Compound library screening in T6PNE

A collection of 8,064 diverse small molecules, a subset of the ChemBridge DiverSet library (ChemBridge Corporation; CA, USA) was screened in a high-throughput assay for activity against in T6PNE expressing the insulin-eGFP transgene, as previously described [96]. Briefly, T6PNE cells were seeded at 2,000 cells per well in 384-well tissue culture plates (Greiner Bio-One; NC, USA) in the presence of 0.5 μM 4-hydroxytamoxifen to induce a submaximal level of insulin expression. Compound addition (active compound in DMSO or vehicle alone) occurred 24 hours after 4-hydroxytamoxifen administration with the BiomekFX (Beckman Coulter; CA,
USA). Forty-eight hours after compound addition, cells were fixed in 4% paraformaldehyde and stained with DAPI (0.167 µg/ml; Invitrogen; CA, USA). Blue (DAPI) and green (human insulin promoter driving GFP) channels were imaged using the GE/Amersham InCell 1000 high-throughput microscopy system. Image processing (Cytoshop, Beckman Coulter; CA USA) was used to evaluate the percentage of cells containing GFP greater than a threshold, as determined by an algorithm in MATLAB. Fold change over vehicle (DMSO) was reported as %GFP+ cells [96].

3.3.3 Primary confirmatory assay for dose-responsiveness

Compounds selected as hits in the primary screen were tested for their reproducibility and dose-responsiveness in T6PNE. As described for the primary screen, compounds were dissolved in DMSO and added to T6PNE 24 hours after 4-hydroxytamoxifen addition. Compounds were tested over a range of concentrations (0.3-20 µM) and compared against vehicle (DMSO) after 48 hours using the imaging techniques described for the primary screen.

3.3.4 Counterscreen for estrogenic activity

As previously reported [96], estrogenic activity was monitored by co-transfection of a reporter plasmid containing a multimerized E-box 5’ of a minimal promoter fused to the Firefly luciferase gene (4RTK-luc) with wild-type E47 or E47MER. As described in the PPRE-Luc reporter assay, HeLa cells were transfected using PEI, 0.2 µg 4RTK-Luc plasmid and either 0.3 µg of human E47, E47MER or pMSCVhph vector in 50 μl of serum-free DMEM per well. Transfections included
Renilla luciferase (pRL-TK) plasmid as a control for transfection efficacy. Transfection conditions are as described in the PPRE-Luc reporter assay. Sixteen hours after transfection, culture media were changed and maintained for 48 hours with 4-hydroxytamoxifen and/or compound or vehicle (DMSO). Cells were then lysed and assayed for luciferase activity using the Promega Dual Luciferase kit (Promega Corp., Wisconsin, USA), and luminescence was measured using the Veritas Microplate Luminometer (Turner Biosystems, California, USA). Data were normalized to Renilla luciferase (pRL-TK) and are expressed as fold-change over vehicle alone.

3.3.5 Quantitative real-time PCR (Q-PCR) analysis

RNA was purified using the RNeasy Kit (Qiagen; CA, USA). 2 µg of RNA was used to synthesize cDNA using the qScript cDNA SuperMix (Quanta BioSciences; Gaithersburg, Maryland, USA). Q-PCR was performed on cDNA corresponding to 100 ng of RNA using the Opticon Real-Time System (MJ Research; MA, USA) and Q-PCR SuperMix (BioPioneer; CA, USA). All mRNA values were normalized to 18S rRNA or GAPDH values and are expressed as fold changes over vehicle-treated control.

**Primers:**

<table>
<thead>
<tr>
<th></th>
<th>FWD</th>
<th>REV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insulin</td>
<td>CTACCTAGTGTGGGGGAAC</td>
<td>GCTGGTAGAGGGAGCAGATG</td>
</tr>
<tr>
<td>HNF4α</td>
<td>CAGAATGAGCGGGACCAGTC</td>
<td>CAGCAGCTGCTCCTCTTCTAAT</td>
</tr>
<tr>
<td>Kip2</td>
<td>GGCGATCAAGAAGCTGTCC</td>
<td>GGGCTCTTTGGGCTCTTCAAT</td>
</tr>
<tr>
<td>18S</td>
<td>GATATGCTCATGTTGTTT</td>
<td>AATCTTCTTCTAGTCGGCTCCA</td>
</tr>
<tr>
<td>GAPDH</td>
<td>ACAGTCAGCGCCTCTTT</td>
<td>AATGAAGGGGTCATTGATGG</td>
</tr>
</tbody>
</table>
3.3.6 Compounds

BIM5078 (1-(2’-chloro-5’-nitrobenzenesulfonyl)-2-methylbenzimidazole) is commercially available from Princeton Bio (NJ, USA), CAS 337506-43-1, Cat# OSS1290849. BIM5078 was also synthesized in larger quantity by Mao Ye, Ph.D. and Marcia I. Dawson, Ph.D. Sample purity of >95% was confirmed by $^1$H NMR and HPLC. BI6015 (2-Methyl-1-(2-methyl-5-nitrophenylsulfonyl)-1H-benzo[d]imidazole) was synthesized by Mao Ye, Ph.D. and Marcia I. Dawson, Ph.D., and is now commercially available from Princeton Bio (NJ, USA), CAS 93987-29-2, Cat# OSS1290848. Sample purity of >95% was confirmed by $^1$H NMR and LC/MS.

Medica-16 (3,3,14,14-tetramethylhexadecanedioic acid) is commercially available from Cayman Chemical (MI, USA), CAS 87272-20-6, Cat# 90290. Bezafibrate (2-[4-[2-[(4-chlorobenzoyl) amino]ethyl]phenoxy]-2-methyl-propanoic acid) is commercially available from Cayman Chemicals (MI, USA), CAS 41859-67-0, Cat# 10009145. Compound 5 (naphtho(2,1-b)furan,1-methyl-2-nitro-) CAS 86539-67-5 was synthesized by Mao Ye, Ph.D. and Marcia I. Dawson, Ph.D, according to the procedure published by [57]. All fatty acids were purchased from Sigma-Aldrich (MO, USA) unless otherwise specified. Acetic acid, CAS 64-19-7, Cat# A6283. Butyric acid (butanoic acid), CAS 107-92-6, Cat# B103500. Caproic acid (hexanoic acid), CAS 142-62-1, Cat# H12137. Caprylic acid (octanoic acid), CAS 124-07-2, Cat# O3907. Capric acid (Decanoic acid), CAS 334-48-5, Cat# C1875. Myristic acid (Tetradecanoic acid), CAS 544-63-8, Cat# M1328. Palmitic acid (Hexadecanoic acid), CAS 57-10-3, Cat# P0500. Suberic acid (Octanedioic acid), CAS 505-48-6, Cat# S5200. Phytanic acid (3,7,11,15-Tetramethylhexadecanoic acid), CAS 14721-66-5,
Cat# P4060. Linolic acid (α-Lnn, cis,cis,cis-9,12,15-Octadecatrienoic acid), CAS 463-40-1, Cat# L2376. Oleic acid (cis-9-Octadecenoic acid, Elainic acid), CAS 112-80-1, Cat# O1008.

3.3.7 In vitro analysis of ADME properties

All pharmacokinetic studies were done by Layton Smith, Ph.D., at the Sanford-Burnham Institute for Medical Research at Lake Nona.

3.3.7.1 Solubility

Solubility analysis was performed using a direct UV kinetic solubility method in a 96-well format. All liquid dispensing and transfer steps were performed with the Freedom Evo automated liquid handler (Tecan USA). Solubility measurements were performed in an aqueous buffer solution (System Solution, pION Inc, P/N 110151) at pH 5.0, 6.2, and 7.4, in duplicate. Samples were incubated at room temperature for a minimum of 18 hours to achieve equilibrium, then filtered (filter plate, pION Inc, P/N 110322) to remove any precipitate formed. The concentration of the compounds was measured by UV absorbance (250-498 nm) using the Infinite M200 (Tecan US) and compared to the spectra of the precipitation-free reference solutions. Spectroscopically pure 1-propanol (Sigma-Aldrich P/N 256404; St. Louis, Missouri, USA) was used as a cosolvent to suppress precipitation in the reference solutions. The solubility of each compound was determined using µSOL Evolution Plus software v3.2 (pION Inc) and is expressed as the concentration (µg/mL) of a solute in a saturated solution.
3.3.7.2  **Permeability**

Permeability was assessed using the Parallel Artificial Membrane Permeability Assay (PAMPA) in a 96-well format. A “sandwich” plate (pION Inc, P/N 110212) consisting of a donor bottom plate and an acceptor filter plate was used. The donor wells contained the compounds in 180 µl system solution and magnetic stir bars. The filter on the bottom of each acceptor well was coated with GIT-0 lipid (pION Inc, P/N 110669) and filled with 200 µl of Acceptor Sink Buffer, pH 7.4 (pION Inc, P/N 110139), containing a surfactant to mimic the function of serum proteins. The permeation time was 30 minutes, and moderate stirring (equivalent to 40 µm Aqueous Boundary Layer thickness) was applied using the Gut-Box™ (pION, Inc, P/N 110205). After the permeation time, the sandwich was disassembled, and the amount of compound present in both the donor and acceptor wells was measured by UV absorbance (250-498 nm) using the Infinite M200 (Tecan USA) and compared to spectra obtained from reference standards. Mass balance was used to determine the amount of material embedded in the membrane filter. The effective permeability, Pe, was calculated using the software PAMPA Evolution Plus, version 3.2 (pION Inc).

3.3.7.3  **Hepatic Microsome Stability**

Metabolic stability was assessed in the presence of human liver microsomes (XenoTech, P/N H0630, lot # 0810063). NADPH, a required cofactor for CYP450 metabolism, was provided by the NADPH Regenerating System, Solutions A (BD Biosciences, P/N 451220) and B (BD Biosciences, P/N 451200). Compound stock solutions were initially prepared in 100% DMSO and subsequently diluted in
acetonitrile for the assay. The pH of the reactions was kept at ~ 7.4 with potassium phosphate buffer (BD Biosciences, P/N 451201). The reactions were started after adding NADPH to the reaction plate containing microsomes and compounds, and time 0 minutes aliquots were promptly collected and mixed with ice-cold acetonitrile (spiked with internal standards) to quench the reactions. The remainder of the reaction volume was incubated at 37 °C with shaking. Additional aliquots were collected 60 minutes after the start of the reaction and promptly quenched with ice-cold acetonitrile (spiked with internal standards). Samples were centrifuged at 3000 rpm for 10 minutes. The amount of compound in the supernatant was determined by LC/MS/MS (Applied Biosystems, Sciex API4000 Q-Trap), and the percent of parent compound remaining after 60 minutes was calculated by the following formula:

\[
\text{% parent compound remaining} = \left( \frac{\text{Concentration at 60 min}}{\text{Concentration at 0 min}} \right) \times 100
\]

All reactions were run in triplicate, except negative controls (no NADPH), which were performed as single reactions. Results are reported as the means of each reaction triplicate, normalized to the internal standard, and expressed as percent compound remaining after the incubation time.

### 3.3.7.4 Plasma Stability

Plasma was allowed to thaw at room temperature prior to mixing with the appropriate buffer in a 1:1 (v/v) ratio. Two buffers were prepared from a 10X PBS stock solution, 1X PBS (pH 7.4) and a 1X Protease Inhibitor cocktail in 1X PBS (pH
7.4) solution. These 4 solutions were placed in a warm water bath at 37°C for 30 minutes prior to use. All stock reference solutions were diluted to 40 µM using DMSO. From each 40 µM test solution, 5 µL was added to the appropriate wells of the 96-well plate, followed by the addition of 195 µL of the appropriate solution. After the addition of the solution, samples were mixed by gently aspirating and dispensing in the plate using a multichannel pipette. Immediately after mixing, 50 µL of each sample was aspirated from the reaction plate and dispensed into the quench plate containing ice cold acetonitrile (containing an internal standard). This plate was then vortexed for 5 minutes and centrifuged for 10 minutes at 3000 rpm. The supernatant was then transferred to a new 96-well plate for analysis by LC/MS/MS as the t = 0 plate. The reaction plate was covered with breathable sealing tape and placed in a shaking incubator for 180 minutes (3 hours) at 37°C. After 3 hours, the same procedure as described above was followed for the t = 180 plate. The supernatant was then transferred to a new 96-well plate for analysis by LC/MS/MS (Applied Biosystems, Sciex API4000 Q-Trap).

3.3.7.5 Plasma Protein Binding

Teflon® Base Plate wells were rinsed with 20% ethanol for 10 minutes. Ethanol was then removed, and wells were rinsed with ultrapure water and allowed to dry. RED Inserts from Thermo Scientific (Pierce) were placed (open end up) into the wells of the base plate. The sample chambers (red ring) contained 300 µl of a mixture of plasma and compound. Also, 500 µl of dialysis buffer (1X PBS, pH 7.4) was added to the buffer chambers of the inserts. Duplicate inserts were made for each
concentration tested. The base plate was covered with sealing tape and incubated at 37°C on an orbital shaker at 300 rpm for 4 hours. After the incubation, equal volumes from both chambers were removed and transferred to a 96-well plate containing either plasma or buffer. To precipitate proteins and release compounds, ice-cold acetonitrile (with an internal standard) was added. Samples were centrifuged for 10 minutes at 3000 rpm. The amount of compound in the supernatant was determined by LC/MS/MS (Applied Biosystems, Sciex API4000 Q-Trap). The percent of free and bound compounds were calculated with the following formulas:

\[
\text{% of free parent compound} = \frac{\text{amt of compound in receiver chamber}}{\text{amt of compound in donor chamber}} \times 100
\]

\[
\text{% of bound parent compound} = [100 - \% of free compound]
\]

Results are reported as the mean of each reaction duplicate, normalized to the internal standard, and expressed as the percent compound bound after the incubation time.

3.3.8 PPRE-Luc reporter assays

PPAR activity was monitored by transfection of an (AOx)₃-tk-Luc plasmid, which contains three copies of the acyl-CoA oxidase PPRE (Peroxisome Proliferator-Activated Receptors) upstream of the basal thymidine kinase (TK) promoter in the tk-Luc plasmid. The PPRE–luciferase promoter-reporter vector (AOx)₃-TK-Luc plasmid and human PPARγ expression vector were kind gifts from Professor Christopher K. Glass (University of California, San Diego) [98]. HeLa cells \((7 \times 10^4)\) were
transfected in 24-well tissue culture plates (1 mL cell suspension), using 1 μg polyethylenimine (PEI, Polyscience Inc., Warrington, Pennsylvania, USA), 0.2 μg (AOx)₃-tk-Luc plasmid, and either 0.3 μg of human PPARγ or pMSCVhph vector in 50 μl of serum-free DMEM per well. Transfections included Renilla luciferase (pRL-TK) plasmid as a control for transfection efficacy. The PEI mixture was incubated for 20 minutes at room temperature. The medium was replaced with fresh medium plus serum, and the PEI mixture was added. After overnight incubation at 37°C under 10% CO₂, the medium was replaced with DMEM containing 10% FBS and troglitizone as the positive control, compound or vehicle (DMSO) up to a final maximal concentration of 0.5% DMSO. Following a 48 hours incubation, cells were lysed and assayed for luciferase activity using the Promega Dual Luciferase kit (Promega Corp., Madison, Wisconsin, USA) according to manufacturer’s protocol. Luminescence was measured on a Veritas Microplate Luminometer (Turner Biosystems, Sunnyvale, California, USA). Data were normalized to Renilla luciferase (pRL-TK) and are expressed as fold-activation above (AOx)₃-tk-Luc activation induced by vehicle alone.

### 3.3.9 In silico docking studies

Molecules were docked into HNF4α receptor (PDB Code: 1m7w) with the program BioMedCache. The docking pose was superimposed with the original ligand fatty acid. Modeling studies were conducted by Dr. Mao Ye.
3.3.10 Quenching of intrinsic fluorescence of HNF4α aromatic amino acids Tyr/Trp

The effect of compound binding on intrinsic fluorescence of full-length HNF4α aromatic amino acids was examined as previously described for fatty acyl-CoAs [99]. Briefly, 100 nM of full-length HNF4α protein was titrated with increasing concentrations of each compound. Fluorescence emission spectra were obtained at 24°C with a PC1 photon counting spectrofluorometer (ISS Inc.; IL, USA), corrected for background (protein only and fluorescent ligand only), and maximal intensities were measured at 330 nm upon excitation at 280 nm. The number of binding sites (n) and binding affinity (K_d) were estimated by fitting the binding curve to a Hill plot according to the equation \( y = \frac{ax^b}{c^b+x^b} \), where \( a \) is the maximal fluorescence intensity, \( b \) is the number of binding sites, and \( c \) is \( K_d \) [99]. Binding studies were performed by Heather A. Hostetler, Ph.D. and Friedhelm Schroeder, Ph.D.

3.3.11 siRNA studies

Ambion silencer siRNA (HNF4α siRNA ID#4013) was delivered into T6PNE cells by reversible transfection by mixing 1 µl of individual siRNA (0.5 µM) and 10 µl of diluted (1:100 in Optimem) Lipofectamine RNAiMAX (Invitrogen; CA, USA) per well of a 384-well plate, followed by incubation for 15 minutes at room temperature. T6PNE cells (2,000 cells per well) diluted in 40 µL of RPMI supplemented with 10% FBS and 1% pen-strep were added to the transfection mix and incubated for 48 hours at 37°C, 5% CO₂, followed by the addition of 1 µM 4-hydroxytamoxifen. Transfected
cells were incubated at 37°C for an additional 48 hours with 4-hydroxytamoxifen. Images were obtained as described in compound library screening in T6PNE. A mixture of 48 nonspecific control siRNAs was tested on T6PNE cells, as well. A similar method was used to perform siRNA transfections in 6-well tissue culture plates to quantify mRNA content. RNA was purified, cDNA was synthesized, and Q-PCR were performed as described above.

### 3.3.12 Chromatin immunoprecipitation (ChIP) assays

T6PNE cells were treated with vehicle (DMSO) or 5 µM BIM5078 for 48 hours and fixed per GENpathway’s cell fixation protocol. Using Factor Path Query analysis (GENpathway; CA, USA), chromatin was immunoprecipitated with rabbit polyclonal anti-E47 (#sc-763, Santa Cruz Biotechnology; CA, USA) or goat polyclonal anti-PDX-1 antibody (#sc-14664, Santa Cruz Biotechnology; CA, USA). Precipitated DNA was quantified using Q-PCR and was performed with primers targeting E47 or PDX-1 response elements (E-box or A-box, respectively) in the human insulin promoter, as indicated below.

**Primers:**

<table>
<thead>
<tr>
<th></th>
<th>FWD:</th>
<th>REV:</th>
</tr>
</thead>
<tbody>
<tr>
<td>hINS +40</td>
<td>GTAGGGGAGATGGGCTCTG</td>
<td>GCTTGATGGCCTCTTCTGAT</td>
</tr>
<tr>
<td>hINS -173</td>
<td>AGCTGTGAGCAGGGACAG</td>
<td>CTTGGTCGTCAGCACCTCTT</td>
</tr>
<tr>
<td>hINS –300</td>
<td>GGACAGCAGCGCAAAGAG</td>
<td>GGGCCACTTTCCACATTAGA</td>
</tr>
</tbody>
</table>
3.3.13 Gene expression profiling

3.3.13.1 Isolation of RNA and expression profiling

Total RNA was extracted from cells grown in six-well tissue culture plates using RNeasy kits (Qiagen; CA, USA) according to the manufacturer’s recommended protocol. A Nanodrop spectrophotometer was used to measure RNA concentrations. RNA quality was determined by an Agilent 2100 Bioanalyzer. RNA (500 ng) was reverse-transcribed by M-MLV reverse transcriptase, amplified, and labeled using an Illumina RNA Amplification kit (Ambion; TX, USA). Labeled cRNA was hybridized to bead-linked 50-mer cDNA oligonucleotide probes arrayed on Illumina HumanRef-8 v.1 Expression Bead Chips for whole-genome expression analysis. The HumanRef-8 Expression Bead Chips contain a 24,000-gene array representing the 23,000 RefSeq collection and 1,000 controls. Microarrays were read using a BeadStation Array Reader and analyzed using BeadStudio (Illumina; CA, USA) at the Sanford-Burnham Institute for Medical Research microarray core. Downstream data analysis was performed using GeneSpring (Agilent; CA, USA).

3.3.13.2 Expression data analysis

We analyzed Illumina HumanRef-8 v.1 Expression BeadChips using the manufacturers BeadStation Array Reader and collected primary data using the supplied Scanner software. Data analysis was done in three stages. First, expression intensities were calculated for each gene probed on the array for all hybridizations using Illumina’s Beadstudio#2 software. Second, intensity values were quality controlled and normalized. To eliminate genes that were not expressed at a statistically significant level, a p-value detection threshold of 0.05 was used as a cutoff. All arrays
were then normalized using the `normalize.quantiles` routine in Bioconductor. This procedure accounted for variation in hybridization intensities between individual arrays. These normalized data were then imported into GeneSpring and analyzed for differentially expressed genes. Genes that were significantly differentially expressed between T6PNE treated with vehicle (DMSO) and T6PNE treated with BIM5078 were determined on the basis of t-tests and fold-difference changes in expression level.

### 3.3.14 NextBio

Genes altered by BIM6078 treatment 2-fold in T6PNE cells were loaded into the web-based systems biology software NextBio to correlate transcription changes observed with BIM5078 treatment to clusters of genes containing previously recognized regulatory motifs in human promoters [-2kb, 2kb] around their respective transcription start sites. Analysis was restricted to the Molecular Signatures Database (MSigDB) containing motif gene sets annotated by the Broad Institute [100].

### 3.3.15 OTC transient transfection assay

HepG2 and CV-1 cells were cultured in Dulbecco’s Modified Eagle Medium (DMEM) supplemented with 10% FBS and 1% pen-strep. Cells were seeded in 24-well tissue culture plates and transfected with 0.4 µg of wild-type OTC promoter[94] and 0.0625 µg/well of pRL-TK using Lipofectamine 2000 reagent (Invitrogen; CA, USA) per the manufacturer's recommendations. After 48 hours of BI6015 treatment, cells were then lysed and assayed for luciferase activity using the Promega Dual Luciferase kit (Promega Corp.; WI, USA), and luminescence was measured using the
Veritas Microplate Luminometer (Turner Biosystems; CA, USA). Data were normalized to *Renilla* luciferase (pRL-TK) and are expressed as fold-change over vehicle alone.

### 3.3.16 Liver panel profiling

One hundred microliters of mouse blood was collected by retro-orbital bleeding in heparin-coated tubes. The samples from BI6015- and vehicle (DMSO)-treated HCC mice were analyzed using a mammalian liver enzyme profile rotor on a VetScan VS2 analyzer (Abaxis; CA, USA).

### 3.3.17 Immunohistochemistry

Samples were harvested from BI6015 and DMSO treated mice, fixed in 4% paraformaldehyde (USB; OH, USA), and embedded in paraffin or OCT freezing media (Sakura Finetek; CA, USA). Slides of 5µm thickness were washed four times with PBS and treated with 0.3% Triton in PBS for 10 minutes. Antigen retrieval was done using with CitriSolvTM (Fisher Scientific, PA) for 10 minutes in sub-boiling temperature. After washing with PBS for 10 minutes, slides were incubated in blocking solution with 5% normal donkey serum (Jackson Immuno Research, PA) for 60 minutes at room temperature. Cells were fixed in 4% paraformaldehyde for 15 minutes on 4°C and washed with PBS, treated with 0.3% Triton in PBS for 10 minutes and blocked as previously described for slides samples.

Primary antibodies: Cleaved Caspase3 (#9664, Cell Signaling; MA, USA), HNF4α (#sc-6556, Santa Cruz Biotechnology; CA, USA; #3113, Cell Signaling; MA,
USA) and BrdU (1:500; #RPN20AB; GE Healthcare, UK). For fluorescent imaging, samples were incubated with Alexa 488 (Invitrogen; CA, USA) or Rhodamine (Jackson ImmunoResearch Laboratories, Inc.; PA, USA) fluor-labeled anti-mouse, rabbit or goat and nuclei were counterstained with DAPI (Invitrogen; CA, USA). Controls using secondary antibodies alone were utilized to ensure specificity of immunostaining. Fluorescently labeled sections were analyzed with a conventional inverted microscope (Olympus, PlanFl 40x/0.60; PA, USA) or with a confocal microscope (Bio-Rad Laboratories Inc.; CA, USA) equipped with krypton/argon laser.

### 3.3.18 Steatosis assay

Steatosis was assessed using the Oil Red O Method for Fats kits (#K043-8OZ, Poly Scientific; NY, USA), per manufacturers guidelines. Briefly, frozen tissue slides or fixed cells were incubated in propylene glycol absolute for 2 minutes and Oil Red O solution for 15 hours for slides or 1 hour for fixed cells, differentiated in 85% propylene glycol solution for 1 minute, washed twice with distilled water and stained in Hematoxylin of 10 seconds. Slides were mounted with glycerin jelly mounting medium.

### 3.3.19 Orthotopic xenograft mouse model

The orthotopic mouse model for liver tumor (HCC) was established in male athymic (nu/nu) nude mice (Harlen, 4–5 weeks old). Mice were anesthetized with avertin (15-17 mg/kg) given intraperitoneally (IP). Hep3B-Luc cells (2 x 10^6 in 30–50 µl of cold PBS) were injected into the upper left lobe of the liver. Mice receiving
Hep3B-Luc were expected to develop a tumor at the injected site beginning 2 weeks post-injection. Growth of injected cells was monitored by measuring the luciferase activity with the IVIS 200 Imaging System (Caliper LifeSciences; MA, USA) twice weekly by injecting 150 mg/kg of D-luciferin (#122796, Caliper LifeSciences; MA, USA) into the mice 10 minutes prior to imaging. Mice that demonstrated a doubling of luciferase counts on 3 consecutive reads by bioluminescent imaging were given 30 mg/kg BI6105 or vehicle (DMSO) IP daily or every other day, as tolerated. Orthotopic xenograft mouse model studies and primary hepatocyte isolations were done in collaboration with Seung-Hee Lee, Ph.D. in Dr. Fred Levine’s lab and Mingjun Zhang, Ph.D. in Dr. Yong Cang’s lab.

3.3.20 Mouse primary hepatocytes isolation

Mouse primary hepatocytes isolated following a standard protocol involving collagenase IV (Sigma, Cat# C5138) perfusion of the whole liver and low-speed centrifugation of detached cells. Isolated hepatocytes were culture with DMEM media in 5% CO2 incubator on collagen coated plate.

3.3.21 Statistical Analysis

Unless otherwise indicated, data are presented as mean ± s.e.m. of three or more independent cultures. Statistical significance was assessed using two-tailed unpaired Student's t-test or Pearson chi-square test, as indicated.
3.4 Results

3.4.1 Identification of a small molecule inhibitor of the insulin promoter: BIM5078

To identify molecules with novel mechanisms of action on the insulin promoter, a high-throughput screen of structurally diverse compounds, consisting of a subset of the ChemBridge DiverSet Library, was performed using the T6PNE insulin promoter assay [96] (Figure 3.1A). As previously described, primary hits were subjected to secondary counterscreens to assess activity on the endogenous insulin promoter and the modified estrogen receptor [96] (Figure 3.2). From the screen, we isolated one small molecule, 1-(2’-chloro-5’-nitrobenzenesulfonyl)-2-methylbenzimidazole, hereafter referred to as BIM5078, which passed the counterscreens and exhibited dose-responsive inhibition of endogenous insulin expression, with an IC50 = 930 nM (Figure 3.1B).

3.4.2 BIM5078 is structurally similar to FK614, a PPARγ agonist

A cheminformatic analysis of BIM5078 revealed structural similarity to an atypical PPARγ agonist, FK614, in clinical trials as a therapeutic for type II diabetes [101] (Figure 3.1C). Accordingly, BIM5078 was tested in a PPAR response element (PPRE)-luciferase reporter assay. BIM5078 activated the PPRE 2-fold, and activation was enhanced nearly 4-fold by co-transfection with a PPARγ expression vector (Figure 3.1D).
In contrast to BIM5078, the PPARγ agonist FK614 was inactive in the T6PNE insulin promoter assay (data not shown). We were unable to document effects on PPARα and PPARδ when co-transfected in the PPRE-luciferase reporter assay (data not shown). This suggests that the repressive effect of BIM5078 on the insulin promoter is due to a target other than the PPARs.

### 3.4.3 BIM5078 binds directly to HNF4α and interacts with the ligand binding pocket in situ

Amongst transcription factors other than the PPARs known to bind to PPREs is HNF4α [102-104], another member of the nuclear hormone receptor superfamily. Accordingly, we measured the direct binding of BIM5078 and FK614 to HNF4α. Binding was determined based on the intrinsic fluorescence of full-length HNF4α aromatic amino acids Tyr/Trp [99]. The resultant K_d’s ± standard deviations were 11.9 ± 2.9nM and 254 ± 55nM for BIM5078 and FK614, respectively (Figure 3.3A).

Shoelson and colleagues at the Joslin Diabetes Center (Boston, Massachusetts) solved the x-ray crystal structures of the human and rat HNF4α ligand binding domains (LBD), noting that fatty acids were always bound in the absence of exogenously added ligand [51, 52]. Using Shoelson’s crystal structures (PDB Code 1m7w and 1pzl), we docked BIM5078, FK614, and fatty acids (FFAs) into the LBD of HNF4α (Figure 3.3B). These in silico studies suggest that not only is it reasonable for BIM5078 to bind the same pocket as the putative ligand (FFA), but they also form the molecular basis for predicting substrate specificity. Superimposition of the docking pose with the original ligand fatty acid shows that the oxygen of the nitro group forms
a stronger salt bridge ion interaction and hydrogen bond with the residues Arg 226 and Gly 237 than FK614. Moreover, the chloro group of BIM5078 extrudes into a wide pocket inside the active site and forms a hydrophobic interaction with residue Val 178.

3.4.4 BIM5078 mediates insulin promoter repression through HNF4α

If HNF4α is the target of BIM5078, it must be expressed in T6PNE cells. To test this prediction, we performed quantitative RT-PCR for HNF4α in T6PNE cells, and we found that HNF4α is potently induced by E47 through 4-hydroxytamoxifen administration by greater than 40-fold in a saturable process (Figure 3.3C). Thus, expression of HNF4α correlates with conditions in which the insulin gene is active in T6PNE cells. This uncovers a previously unrecognized mode of regulation of the HNF4α promoter.

A further prediction of a model in which HNF4α acts on the insulin promoter is that direct inhibition of HNF4α expression by siRNA should inhibit insulin promoter activity. While there is a well-characterized binding site for HNF4α in the rodent insulin promoter [44], the cognate region of the human insulin promoter varies in sequence, and HNF4α has not been detected in ChIP-Chip studies as being directly bound to the human insulin promoter [45]. However, there is a well-characterized HNF1α site in the human insulin promoter [24], suggesting one possible route through which HNF4α can act.

As predicted, transfection of HNF4α siRNA into T6PNE cells led to potent inhibition of insulin promoter expression. We observed a 15-fold reduction in the
percent of GFP-positive cells and a nearly 120-fold reduction in the level of endogenous insulin mRNA (Figure 3.3D). Taken together, these data demonstrate that endogenous HNF4α expressed in T6PNE cells has a role in maintaining the activation state of the human insulin promoter in these cells.

3.4.5 BIM5078 disrupts binding of E47 and PDX-1 to the human insulin promoter

Because of the evidence that HNF4α does not bind directly to the human insulin promoter, we further investigated the mechanism by which BIM5078 inhibits insulin gene expression. One possibility was that HNF4α acts indirectly, through effects of other insulin promoter transactivators. To test this hypothesis, ChIP assays were used to probe the binding of transcriptional co-activators E47 and PDX-1 to their regulatory sequences on the insulin promoter (E-box [CANNTG] and A-box [TAAT] motifs, respectively) in the presence and absence of BIM5078. After 48 hours, treatment with 5µM BIM5078 significantly decreased association of E47 to the distal E-box (E2) but not the proximal E-box (E1) of the insulin promoter (Figure 3.4A). Furthermore, association of PDX-1 was significantly disrupted at all tested A-boxes (Figure 3.4B). These results are consistent with the hypothesis that BIM5078-mediated repression of insulin promoter activity occurs indirectly through disrupted binding of E47 and PDX-1.
3.4.6 BIM5078 affects the expression of known HNF4α target genes

To provide additional evidence that BIM5078 inhibits HNF4α, we examined its effect on known HNF4α target genes. HNF4α autoregulates its own transcription through a complex regulatory feedback loop between HNF4α and HNF1α, part of which involves direct binding of HNF4α to its own promoter. As predicted, BIM5078 potently repressed HNF4α expression in T6PNE cells (Figure 3.5A), as well as in the HepG2 hepatoma line, which was previously recognized to have exceptionally high levels of HNF4α expression (Figure 3.5B).

We expanded the analysis of the effects of BIM5078 on gene expression by performing gene expression profiling of T6PNE cells in the presence and absence of BIM5078. We have previously shown that T6PNE cells exhibit a pattern of gene expression similar to that of pancreatic islets (GSE 18821) [96]. Accordingly, we performed a bioinformatic analysis to compare genes affected by BIM5078 in T6PNE cells and genes affected by genetic deletion of HNF4α in mouse islets [47]. Of the 156 identifiable genes altered in HNF4α-deleted islets compared to normal islets, 20 were exact matches to genes affected by BIM5078. An additional 36 genes were closely related to genes altered by HNF4α knockout, e.g., cyclinD1 and Rab3a versus cyclinD2 and Rab3b, respectively. Thus, 36% of the genes affected by genetic deletion of HNF4α are either identical or closely related to genes affected by pharmacologic inhibition of HNF4α. A Chi-square analysis demonstrated that T6PNE cells treated with BIM5078 are statistically similar to the transcriptional profile of HNF4α genetic deletion (36% intersection, p<0.0001) when compared with 50,000 trials of randomly generated gene sets of the same size selected from genes expressed in T6PNE cells.
(11% intersection predicted). The concordance between the genes affected by genetic deletion and pharmacologic antagonism of HNF4α is particularly striking given that the analysis was done in different species (human versus mouse), different cell types (the cell line T6PNE cultured in vitro versus primary mouse islets), and genetic deletion in murine β-cells occurred during embryonic development, as soon as the insulin promoter became active.

To further examine the effects of BIM5078 on gene expression, we utilized the web-based systems biology software NextBio to compare lists of genes containing cis-regulatory motifs with the list of genes modulated by BIM5078. We restricted our analysis to the Molecular Signatures Database (MSigDB) containing motif gene sets annotated by the Broad Institute to enable us to associate changes in our microarray studies with conserved, putative cis-regulatory elements [100]. The most significantly correlated regulatory motifs were the E12, AP4 and MYOD binding sites, all of which contain the core CANNTG E-box sequence that we had found to be altered in the ChIP analysis of the insulin promoter. It is interesting to note that there was a statistically significant association with “HNF4α binding site geneset 1” (containing the motif VTGAACCTTTGMMB) but not 4 other binding genesets representing alternate consensus binding sites for HNF4α, raising the possibility that BIM5078 may affect the activity of HNF4α at some sites preferentially over others, similar to the effect of other nuclear receptor ligands on their target receptors [105].

Given the effect of BIM5078 on genes containing E-boxes in their promoter, as well as its effect on E47 binding to the insulin promoter in T6PNE cells, we compared the genes altered by BIM5078 to a set of previously published genes altered by E47
induction in T6PNE cells [96]. Of the 214 genes altered by BIM5078, 67 (31%) were also altered by E47 induction. This association was enhanced when only the genes containing E-boxes were compared. Specifically, of the 96 E-box containing genes altered by BIM5078, 42 were also altered by E47 induction. This global analysis supports the idea that BIM5078-mediated HNF4α modulation of the human insulin promoter may be the result of effects on transcriptional activity mediated by E-box sequences.

The effect of BIM5078 on E-box binding in the insulin promoter prompted us to examine the effects of BIM5078 on other genes regulated by E-boxes. We have previously shown that the cyclin-dependent kinase inhibitor p57Kip2 is directly regulated by E47 occupation of a particular E-box within the p57Kip2 promoter [96]. Although p57Kip2 does not contain an HNF4α binding site, BIM5078 potently reduced Kip2 expression in T6PNE cells (Figure 3.5C). The effects of BIM5078 on E-box activity, combined with the induction of HNF4α expression by E47, suggest HNF4α and E47 form a complex, multicomponent regulatory network that can function as a feedback loop, similar to what has been described by Odom et al. between HNF4α and HNF1α [45].

3.4.7 Fatty acids bound to HNF4α preferentially inhibit the human insulin promoter

Because of uncertainty in the literature about the role of natural or synthetic ligands in controlling HNF4α activity, we sought to determine whether ligands that have previously been reported to bind HNF4α had activity in our T6PNE assay. In the
absence of exogenously added ligand, MCFAs and LCFAs have always been found in the HNF4α ligand-binding pocket in structural studies of HNF4α purified from bacteria[51, 52]. Recently, linoleic acid was found to be bound in the ligand binding pocket of HNF4α in COS-7 cells, as well as in the livers of fed but not fasted mice, suggesting that it might be a natural ligand [55]. Furthermore, this demonstrated that binding within the ligand-binding pocket of HNF4α could be reversible. Based on the widely accepted mechanism of nuclear receptor regulation, one would expect binding of the endogenous ligand (linoleic acid) to its native receptor (HNF4α) to induce changes in HNF4α activity. To test this, a broad spectrum of fatty acids was assessed for effects on insulin promoter activity. We found that MCFAs and LCFAs, including the polyunsaturated fatty acid linoleic acid, inhibited the insulin promoter (Figure 3.6A), while short chain fatty acids (SCFA), such as acetic and butyric fatty acids, which have not been isolated in the ligand-binding pocket of HNF4α, did not have an effect on insulin gene expression. The correlation between fatty acids bound to HNF4α and those that have a repressive effect on the insulin promoter assay are consistent with the effect of fatty acids on the insulin promoter, and these effects are mediated, at least in part, by inhibition of HNF4α, which suggests a role for HNF4α in β-cell lipotoxicity.

In addition to fatty acids, a number of other compounds have been proposed to be HNF4α ligands, including bezafibrate, an acyl-CoA thioester (“Medica-16”) and a nitro-naphthofuran derivative (“Compound 5”) [56, 57]. Medica-16 was weakly repressive at low levels of E47 and HNF4α expression (0.5 μM 4-hydroxytamoxifen), and effects were lost at higher E47 and HNF4α expression levels (1 μM 4-
hydroxytamoxifen). It has been suggested that acyl-CoA thioesters are too large for the HNF4α ligand binding pocket [53, 58]. Bezafibrate was weakly repressive at all doses of drug and 4-hydroxytamoxifen tested. Finally, compound 5, which was reported to be an HNF4α agonist solely based on its ability to repress basal transcriptional levels more than HNF4α-induced transcriptional levels, had no effect on the insulin-GFP reporter assay at low 4-hydroxytamoxifen concentrations but was repressive at higher 4-hydroxytamoxifen concentrations (Figure 3.6B,C).

3.4.8 Development and characterization of lead compound BI6015

Due to its striking behavior in vitro, pharmacokinetic studies were performed to evaluate the translational potential of BIM5078 for in vivo studies. BIM5078 was found to have relatively low plasma stability and only moderate microsomal stability (8% remaining after 3 h and 32% after 1.25 h, respectively), high binding to plasma proteins (98% bound after 4 h), and low solubility (0.17 µg/ml after 18 h).

In studying structural analogs of BIM5078, we identified BI6015, which turned out to have better pharmacokinetic properties than the parent compound. Replacement of the chloro group of BIM5078 with a methyl group resulted in BI6015 (2-Methyl-1-(2-methyl-5-nitrophenylsulfonyl)-1H-benzo[d]imidazole). BI6015 reduced endogenous insulin gene expression 10-fold in T6PNE cells (Figure 3.5C). Furthermore, HNF4α gene expression in HepG2 cells was reduced by up to 6-fold (Figure 3.5B), as compared with the 3-fold reduction previously described with BIM5078. In addition, we tested BI6015 for activity on the ornithine transcarbamoylase (OTC) promoter, which has been well characterized as responsive
to HNF4α in transient transfection assays[94]. BI6015 potently inhibited luciferase expression driven by that promoter in both HepG2 cells and CV-1 cells (Figure 3.5D). Interestingly, BI6015, unlike BIM5078, which differs only by the substitution of a methyl for a chloro group, was not found be a PPARγ agonist (Figure 3.1D). Thus, despite structural similarity between synthetic and natural ligands that bind PPARγ[106] and those that bind HNF4α, we have found that minor structural modifications can dissociate their effects, thereby reducing off-target effects.

In vivo PK studies of BI6015 in mice revealed a half-life of approximately 90 minutes on delivery by oral gavage or intraperitoneal (IP) injection, with moderate plasma levels (AUC = 1.6 µg·min/mL after a 30 mg/kg IP injection). In addition, high levels of BI6015 were found in the liver (3.1 µM at 24 hr after a 30 mg/kg IP injection in mice), supporting its use in hepatocellular carcinoma models. Thus, BI6015 is a potent HNF4α antagonist that has a more favorable stability profile over its predecessor, BIM5078, making it suitable for in vivo studies. In addition, a Ricerca hit profiling panel of 41 receptors/enzymes, including major cytochrome P450s and a number of nuclear receptors, showed cross reactivity with only a minor CYP (2C19) and one L-type calcium channel.

3.4.9 HNF4α antagonism induces loss of HNF4α expression and hepatic steatosis in vivo

BI6015, injected into the peritoneum of mice once per day for 5 days at a dose of 30 mg/kg, was well tolerated. A VetScan Liver Panel revealed that blood chemistries, including ALT, were unaffected by BI6015 treatment, suggesting an
absence of hepatocellular death in the normal liver (Figure 3.7). Because of the role HNF4α is known to play in the liver, it was harvested and examined, along with the intestine and kidney (other sites of HNF4α expression).

Consistent with the effects of HNF4α antagonist BIM5078 on HNF4α expression in vitro, BI6015 induced a substantial loss of HNF4α protein in the liver but not in the intestine or kidney (Figure 3.8A). It is possible that the lack of effect in the intestine and kidney was due to a high degree of hepatic first pass metabolism following IP administration, as hepatic microsomal stability studies on BI6015 demonstrated that only 22% of the compound remained after a 60-minute incubation in vitro. Furthermore, no difference in Ki67 levels was observed in liver, intestine or kidney with BI6015, as compared with vehicle.

Although there was no evidence of hepatocellular death, the hepatocytes from livers of mice injected with BI6015 exhibited marked steatosis (Figure 3.8B), similar to the effect of genetic deletion of HNF4α in the liver [41]. Steatosis and vesicular changes were dose-dependent between 10 and 30 mg/kg of BI6015 administered daily or every other day, as evidenced by increased Oil Red O staining. Similarly, BI6015 induced steatosis in primary mouse hepatocytes in vitro, again without evidence of cell death (Figure 3.8B).

3.4.10 HNF4α antagonism induces fatty acid oxidation defects

Given the known role of HNF4α in fatty acid metabolism [45, 46] and the fact that our synthetic HNF4α antagonists induced hepatic steatosis, we studied the effect of the antagonists on fatty acid metabolism. Acylcarnitine profiles of T6PNE cells
treated with BI6015 revealed elevated levels of acetylcarnitine and decreased levels of longer chain acylcarnitines, particularly C18, C18:1, C16, and C14 (Figure 3.8C). These results indicate that hepatic steatosis resulting from HNF4α inhibition cannot be explained by impaired fatty acid oxidation. A similar phenomenon has been reported in mice fed the conjugated trans-10, cis-12-isomer of linoleic acid, in which lipid accumulation is observed despite increases in fatty acid oxidation [107].

3.4.11 HNF4α antagonists are cytotoxic to human hepatocellular carcinoma (HCC)

While studying BIM5078 and BI6015 in vitro, we unexpectedly noted that they were potently toxic to a number of different tumor cell lines but not to cultured primary cells. To further characterize this phenomenon, we examined the effects of BI6015 on a variant of the human hepatoma cell line Hep3B, which is engineered to express a luciferase transgene under the control of the CMV promoter, in parallel with primary hepatocytes. BI6015 treatment was markedly toxic to Hep3B cells (Fig. 7b) but spared primary hepatocytes (Figure 3.9A). These observations coincided with cell cycle arrest, as evidenced by decreased BrdU (green) incorporation after 48 hours of treatment (Figure 3.9B). Consistent with what we and others have observed with respect to HNF4α antagonism, perturbations to hepatocytes by BI6015 resulted in characteristic steatosis in the absence of cell death (Figure 3.9D). To determine the scope of tumor cytotoxicity, BI6015 was submitted to the developmental therapeutics program (DTP) at the NCI/NIH, where it was screened in vitro against a panel of 60 human tumor cell lines. Slowed growth and/or toxicity were observed with BI6015
treatment on all cancer cell lines tested (Figure 3.10). Similarly, siRNA to HNF4α also resulted in toxicity to HepG2 cells and to a greater extent to T6PNE cells (Figure 3.9G).

To extend the in vitro finding of cytotoxicity to tumor cells in vivo, we studied the effects of BI6015 treatment in a human orthotopic xenograft model, in which luciferase-expressing Hep3B cells were injected directly into the liver parenchyma [108]. Once the tumor was established, as defined by a doubling of luciferase counts on 3 consecutive reads by bioluminescent imaging, the animals were injected IP with 30 mg/kg of BI6015 daily or every other day, as tolerated. After 20-57 days of treatment, the tumors and normal liver samples were harvested and analyzed. There was a marked induction of Oil Red O staining in both the normal liver and tumor samples, although the background level of Oil Red O in the tumor cells was very high (Figure 3.9D). Interestingly, the distribution of the steatotic cells was limited, being restricted to areas around the hepatic blood vessels (Figure 3.9D).

Within the tumor, but not in the normal liver, there were regions with high levels of apoptotic cells positive for TUNEL and cleaved caspase-3 (Figure 3.9E,F). While some treated animals exhibited decreased luciferase counts, these results were not consistent and could not be pursued further because higher doses could not be delivered due to limited solubility of the compound. We believe that the limited efficacy could have been due to a high degree of hepatic metabolism, thereby limiting the amount of drug that penetrated into the tumor, which would be consistent with steatosis in both normal liver and tumor samples being localized to the region around blood vessels (Figure 3.9D). Poor tissue penetration of the compound could also have
played a role. Our results suggest that strategies to improve the pharmacokinetic properties of benzimidazole-based HNF4α antagonists would increase the therapeutic index and therefore tumor size and survival.

### 3.5 Discussion

HNF4α is an attractive target for pharmacologic manipulation for a number of reasons. Not only is it at the center of multiple complex feedback loops that maintain differentiated function in the pancreas and the liver, but HNF4α has also been implicated in a number of disease states, including diabetes, inflammatory bowel disease, cancer, and others. Here, we report the discovery of a novel class of small molecule inhibitors of HNF4α that help interrogate and clarify complex processes driven by HNF4α-dependent pathways.

Until recently, HNF4α was considered to be an orphan receptor. Studies by Shoelson and colleagues identified tightly bound fatty acids in the ligand binding pocket of HNF4α, which existed in a mixture of active and inactive conformations, suggesting the fatty acids were playing a structural role, rather than inducing a specific conformational change [51, 53]. However, in 2009 Yuan et al. used affinity isolation followed by mass spectrometry (AIMS) to reveal that linoleic acid (LA, C18:2ω6) was bound to HNF4α in the livers of fed but not fasted mice, suggesting a ligand could reversibly bind to the ligand-binding pocket [55]. This was the first piece of evidence that suggested the ligand-binding pocket could be targeted by an exogenous ligand, lending itself to pharmacologic manipulation. These results were also consistent with findings in *Drosophila* demonstrating that a GAL4-dHNF4 ligand sensor could be
activated by starvation or exogenous long chain fatty acids [46]. However, the study by Yuan et al. did not find evidence of a significant effect of ligands on HNF4α transactivation. In contrast, we found that MCFAs and LCFAs, including the polyunsaturated fatty acid linoleic acid, antagonized the insulin promoter, while shorter chain fatty acids, which do not bind to HNF4α, did not. The difference between our results and those of previous investigators is not surprising, as we have found that the T6PNE cells are highly sensitive to modulation of HNF4α compared to other cell lines (unpublished results). The striking correlation between fatty acids that have been found to be bound to HNF4α and those that have a repressive effect on the insulin promoter assay suggests a role for HNF4α in β-cell lipotoxicity.

The exquisite sensitivity of the T6PNE cells to modulation of HNF4α activity is further exemplified by comparison with the effect of genetic deletion of HNF4α in pancreatic β-cells, which did not result in loss of insulin gene expression [47]. Thus, we attribute the successful discovery of HNF4α antagonists to the characteristics of the T6PNE assay that render it sensitive to modulation of HNF4α by exogenous ligands. We have shown that HNF4α not only appears to act on a number of target genes through effects on E47, which we induced at a submaximal level in T6PNE for the screen, but it is also induced by E47 expression, suggesting that HNF4α and E47 form a multicomponent regulatory network involving complex feedback loops, similar to what has been described by Odom et al. for the interaction between HNF4α and HNF1α [45]. This relationship between E47 and HNF4α renders T6PNE cells susceptible to HNF4α manipulation and afforded us a unique opportunity to discover HNF4α ligands. The fact that the vast majority of HNF4α is synthesized de novo...
following the induction of E47 (Figure 3.3c) and the addition of compound may also have facilitated access of synthetic ligands to the ligand binding pocket, which would otherwise have been occupied by endogenous fatty acids.

The effect of HNF4α antagonists on E-box containing genes was unexpected, as the vast majority of these genes are not known to contain HNF4α binding sites and do not bind to HNF4α in ChIP-Chip assays [45]. Classically, nuclear receptors exert their effects by binding to highly conserved DNA binding elements. However, it is known that nuclear receptors can also act indirectly with target genes through a mechanism known as tethering to affect transactivation [109]. This tethering of nuclear receptors to affect gene promoters has been reported to occur through a number of DNA binding transcription factors, including members of the bHLH class such as E47 [110] and the homeodomain class [111]. On the human insulin promoter, which does not contain an HNF4α binding site and does not bind to HNF4α in ChIP-Chip assays [45], we found that binding of both E47 and PDX-1 was inhibited by an HNF4α antagonist. This suggests a model in which the effects observed on E-box containing genes, including the insulin promoter, is through a mechanism that involves ligand dependent recruitment of HNF4α by a tethering mechanism to influence binding of direct transcriptional activators.

Finally, while studying the effect of the compounds in vitro, we noticed marked toxicity in a variety of tumor cell lines but not to cells cultured from primary tissue. HNF4α has previously been described to have a role in tumor pathogenesis, but the studies are conflicting, with both upregulation and downregulation of HNF4α expression being reported in association with tumor progression. Knockdown of
HNF4α mRNA by siRNA has been shown to inhibit growth and proliferation of colorectal cancer cells \textit{in vitro} [112]. This is consistent with reports that have shown that HNF4α is upregulated in human hepatocellular carcinoma [61]; however, others have shown downregulation of HNF4α promotes tumorigenesis in hepatocellular and other cancers [63].

Several lines of evidence have suggested that the effects of BI6015 on transformed cells are the result of on-target effects mediated through HNF4α. First, we detected marked toxicity with selective HNF4α knockdown by siRNA. Furthermore, the concentration dependence of cytotoxicity is very similar to that for effects on downstream targets of HNF4α, including its repressive effects on the insulin promoter. This would not be expected if the effects on cancer cells were caused by an off-target effect.

\textit{In vivo}, BI6015 caused dose-dependent hepatic steatosis in normal hepatocytes and in the Hep3b xenografted cells. This is consistent with the effects of genetic ablation of HNF4α [41]. Importantly, the induction of steatosis provided a biomarker for where the compound was acting, which appeared to be restricted to regions surrounding blood vessels. Apoptosis was induced in Hep3b cells but not in primary hepatocytes, consistent with the \textit{in vitro} results. Of note, apoptotic cells also were distributed in a perivascular pattern, mimicking that of steatosis. The lack of activity of BI6015 distal to the vessels suggests either poor tissue penetration, as has been shown for a number of chemotherapeutics including doxorubicin [113], or extensive hepatic metabolism that limits that amount of active compound to regions surrounding vessels where it first encounters hepatocytes. The \textit{in vitro} pharmacology with hepatic
microsomes suggests that extensive hepatic metabolism is likely to be occurring in vivo, suggesting the need for additional medicinal chemistry. Nonetheless, our data indicate that HNF4α antagonists are promising and highly novel targets for cancer therapy. In addition, they provide powerful tools for studying the function of HNF4α.

3.6 Acknowledgements

The OTC-luciferase reporter plasmid was a kind gift from Professor Frank J. Gonzalez (National Institute of Health). Roy Williams (Informatics Core) and Kang Liu (Genomics Core) provided support for the microarray studies. This work was supported by the Sanford Children’s Health Research Center (FL) and grants from the UCSD Genetics Training Grant (AK, TC), the UC Systemwide Biotechnology Research and Education Program Graduate Research and Education in Adaptive Biotechnology (GREAT) Training Program (AK), and the American Diabetes Association Clinical Scientist Training Award (AK).

The initial screen was conducted by Suzette Farber-Katz and Paul Bushway. In silico docking studies and medicinal chemistry were performed by Mao Ye, Ph.D. and Marcia I. Dawson, Ph.D. Binding studies were performed by Heather A. Hostetler, Ph.D. and Friedhelm Schroeder, Ph.D. All pharmacokinetic studies were done by Layton Smith, Ph.D., at the Sanford-Burnham Institute for Medical Research at Lake Nona. Acylcarnitine studies were performed by the Sanford-Burnham Institute for Medical Research at Lake Nona. Orthotopic xenograft mouse model studies and primary hepatocyte isolations were done in collaboration with Seung-Hee Lee, Ph.D. in Dr. Fred Levine’s lab and Mingjun Zhang, Ph.D. in Dr. Yong Cang’s lab.
Chapter 3, in part, is currently being prepared for submission for publication of the material as it may appear in Nature Chemical Biology, 2011, with co-authors Alice Kiselyuk, Seung-Hee Lee, Suzette Farber-Katz, Mingjun Zhang, Sonalee Athavankar, Tom Cohen, Anthony B. Pinkerton, Mao Ye, Paul Bushway, Adam Richardson, Heather A. Hostetler, Li Huang, Benjamin Spangler, Jennifer Higginbotham, John Cashman, Pamela Itkin-Ansari, Marcia I. Dawson, Friedhelm Schroeder, Yong Cang, Mark Mercola, and Fred Levine. The dissertation author was the primary investigator and author of this paper.
3.7 Figures

Figure 3.1: Identification of BIM5078, a regulator of insulin gene transcription

(A) A subset of the Chembridge DiverSet library was screened for activity on the insulin promoter in T6PNE cells. (B) T6PNE cells were exposed to the indicated concentration of BIM5078 for 48 hours. Effects of BIM5078 on the exogenous insulin promoter in T6PNE is reported as percent GFP+ cells. Endogenous insulin promoter activity was probed through RT-PCR for insulin and GAPDH mRNA. Insulin mRNA values are reported as normalized to GAPDH to control for nonspecific compound effects. Values represent the mean ± SE, n=3. (C) Structural homology between BIM5078 and known PPARγ agonist, FK614. (E) PPRE Activation. HeLa cells were co-transfected with a PPRE reporter plasmid and a PPARγ expression vector. BIM-5078 activated the PPRE in HeLa cells and activation was enhanced by co-transfection with a PPARγ expression vector. B16015 was also identified as a PPARγ agonist. Values represent the mean ± SE, n=3.
Figure 3.2: Only BIM5078 passed all confirmatory and counter screens

(A) Compounds found to modulate the insulin-eGFP transgene in the primary screen were confirmed in a dose-response assay. Three potential activators and four potential repressors were identified. (B) T6PNE contains E47 MER, which translocates to the nucleus and becomes transcriptionally active in the presence of 4-hydroxytamoxifen. Compounds found to reproducibly increase activity of the insulin promoter driving eGFP in the secondary screen (A) were tested for their ability to bind a multimerized E-box and drive luciferase expression. One of three potential agonists tested was found to modulate E47 MER. (C) Similarly, compounds found to decrease the activity of the insulin-eGFP transgene (A) were tested for their ability to prevent the translocation of E47 MER. All potential compounds, with the exception of BIM5078, were found to act through E47 MER. (D) Compounds that did not interact with E47 MER were tested for their ability to modulate endogenous insulin mRNA. Of those tested, only BIM5078 acted on the endogenous insulin promoter, and was thus chosen for further evaluation.
Figure 3.3: BIM5078 represses insulin promoter activity through HNF4α antagonism

(A) Direct binding based on quenching of HNF4α aromatic amino acid fluorescence emission. Full-length HNF4α protein (100nM) in 2mls PBS was titrated with increasing concentrations of BIM5078 and FK614. The data are presented as the change in fluorescence intensity (F₀ – F) plotted as a function of ligand concentration. Values represent the mean ± SE, n=3. Insets, linear plot of the binding curve from each panel. (B) In silico docking of C12FA, BIM5078 and FK614 into the rat HNF4α ligand binding domain. C12FA, BIM-5078 and FK614 were docked into HNF4α receptor (PDB Code: 1m7w) with the program BioMedCache. Superimposition of the docking pose with the original ligand fatty acid shows that the oxygen of the nitro group forms a stronger salt bridge ion interaction and hydrogen bond with the residue Arg 226 and Gly 237. Moreover, the chloro group of BIM 5078 extrudes into a wide pocket inside the active site, and forms a hydrophobic interaction with residue Val 178. (C) Induction of E47 by 4-hydroxytamoxifen, as previously described[96], induces HNF4α expression more than 40-fold in T6PNE, which have a relatively low level of HNF4α expression at baseline. (D) HNF4α siRNA represses exogenous and endogenous insulin promoter activity in T6PNE cells 15- and 120-fold, respectively.
Figure 3.4: BIM5078 disrupts binding of E47 and PDX-1 to the human insulin promoter

(A and B) BIM5078 decreases association of E47 and PDX-1 to the human insulin promoter. T6PNE cells were treated with 1µM 4-hydroxytamoxifen and DMSO or 5µM BIM5078 for 48 hours followed by ChIP with anti-E47 antibody and anti-PDX-1 antibody. Q-PCR was performed with primers targeting E-box and A-box elements in the insulin promoter. (A) At 48 hours, treatment with BIM5078 significantly decreased binding of E47 to the distal E-box (E2), but not to the proximal E-box (E1). (B) Similarly, BIM5078 decreased binding of PDX-1 to all A-box elements tested (A1, A3, and A5) on the insulin promoter. Differences in binding between vehicle (DMSO) and BIM5078 were measured by a student t-test (*p<0.05, **p<0.01).
Figure 3.5: BIM5078 and BI6015 affect the expression of known and novel HNF4α target genes

(A and B) HNF4α antagonists, BIM5078 and BI6015, regulate gene expression of HNF4α. T6PNE or HepG2 cells were cultured for 48 hours in the presence of BIM5078 or BI6015. HNF4α mRNA values are reported as normalized to 18s to control for nonspecific compound effects. (C) HNF4α antagonists, BIM5078 and BI6015, reduce insulin and Kip2 mRNA in T6PNE cells. Insulin and Kip2 mRNA values are reported as normalized to 18s to control for nonspecific compound effects. (D) Effect of BI6015 on OTC promoter. pGL3/mOTC-235, a plasmid containing the HNF4α-responsive OTC promoter driving the firefly luciferase gene, was co-transfected into HepG2 and CV-1 cells with a plasmid encoding full length human HNF4α. Cells were treated for DMSO or 1µM BI6015 for 48 hours. HepG2 cells had greater baseline stimulation of the OTC-luciferase transgene than CV-1 cells and BI6015 potently reduced this baseline activity in both cell lines. Values represent the mean ± SE, n=3.
Figure 3.6: Reported HNF4α ligands have variable effect on insulin promoter activity

(A) T6PNE cells were treated with either fatty acid or vehicle (DMSO) for 48 hours. Medium and long chain fatty acids (MCFAs and LCFAs, respectively), but not short chain fatty acids (SCFAs) inhibited insulin expression. Insulin mRNA values are reported as normalized to 18s to control for nonspecific compound effects. Values represent the mean ± SE, n=3. (B and C) T6PNE cells were treated with bezafibrate, Medica-16 or a nitro-naphthofuran derivative ("Compound 5") for 48 hours in the presence of either 0.5µM (B) or 1µM (C) 4-hydroxytamoxifen. Effects on the exogenous insulin promoter transgene in T6PNE are reported as percent GFP+ cells. Values represent the mean ± SE, n=3.
Figure 3.7: Blood chemistry of mice treated with BI6015

Mice (NONcNZO10/LtJ or ICR) were injected IP with vehicle (DMSO) or BI6015 once daily for 5 days. Prior to sacrifice, blood was drawn and analyzed using a VetScan blood analyzer, measuring alkaline phosphatase (ALP, IU/L), alanine aminotransferase (ALT, IU/L), gamma glutamyl transferase (GGT, IU/L), bile acids (BA, µmol/L), total bilirubin (TBIL, mg/dL), albumin (ALB, g/dL), blood urea nitrogen (BUN, mg/dL), and cholesterol (CHOL, mg/dL). Five groups of mice were studied: normal mice injected with DMSO (Normal DMSO, n=4), normal mice injected with BI6015 at a dose of 30mg/kg/day (Normal BI6015H, n=4) or 10mg/kg/day (Normal BI6015L, n=4), mice injected with the hepatocellular carcinoma (HCC) cell line Hep3B and treated with DMSO (HCC DMSO, n=5) for 13-29 days, and mice injected with Hep3B and treated with BI6015 at a dose of 30mg/kg/2day (HCC BI6015, n=4) for 29-36 days.
Figure 3.8: HNF4α antagonism induces hepatic steatosis independent of fatty acid oxidation defects

(A) BI6015 decreases expression of HNF4α in the liver, but not in the intestine or kidney. (b and c) BI6015 induces hepatic steatosis in murine hepatocytes in vitro and in vivo. (B) Primary murine hepatocytes were exposed to DMSO or BI6015 (5mM) for 3 days, followed by fixation and staining with Oil Red O. Mag 100X (C) BI6015 was injected IP once per day at either 10 or 30 mg/kg/day. Mice were sacrificed for analysis of organ histology after 5 days. Liver sections were stained with H&E and Oil Red O. Mag. 400X. (D) Acylcarnitine analysis by GC/MS was performed on cellular extracts of T6PNE cells. Biological duplicates were analyzed. Values represent the mean. Inlay depicts the same data beginning with C4-OH on the x-axis on a different scale to show less abundant species.
Figure 3.9: HNF4α antagonists are cytotoxic to hepatocellular carcinoma

(A) Toxicity was not observed in primary hepatocytes treated with BI6015 in vitro. (B) BI6015 was toxic to Hep3B-Luc cells, as measured by bioluminescence. Fat accumulation (Oil Red O, red) was enhanced with BI6015, while cell division (BrdU, green) was blocked. (C) Bioluminescence of (B) was quantified and correlated with cell counts. (D) Hep3B-Luc cells were injected into the liver parenchyma Cof nude mice (HCC mice). Animals were injected IP with 30 mg/kg of BI6015 daily, as tolerated. Treatment with BI6015 resulted in induction of Oil Red O staining in the normal liver and tumor region. (E and F) Apoptosis was evaluated by immunostaining for cleaved caspase-3 (red) and TUNEL (green). High levels of cleaved caspase-3 were found in tumor regions, but not in normal liver of BI6015-treated HCC mice as. Similar results were found with TUNEL positive cells. (G) Transfection with HNF4α siRNA resulted in marked toxicity to both T6PNE and HepG2 cells.
Figure 3.10: NCI Panel
Data by the NCI panel were generated by the NCI, as described [114].
### 3.8 Tables

Table 3.1: Summary of small molecule screening data

<table>
<thead>
<tr>
<th>Category</th>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assay</td>
<td>Type of assay</td>
<td>Cell-based</td>
</tr>
<tr>
<td></td>
<td>Target</td>
<td>Human insulin promoter</td>
</tr>
<tr>
<td></td>
<td>Primary measurement</td>
<td>Detection of green intensity (green fluorescence protein driven by the human insulin promoter)</td>
</tr>
<tr>
<td></td>
<td>Key reagents</td>
<td>4-hydroxy4-hydroxytamoxifen (Sigma-Aldrich; MO, USA)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAPI (Invitrogen; CA, USA)</td>
</tr>
<tr>
<td></td>
<td>Assay protocol</td>
<td></td>
</tr>
<tr>
<td>Library</td>
<td>Library size</td>
<td>8,064</td>
</tr>
<tr>
<td></td>
<td>Library composition</td>
<td>Diverse small molecules</td>
</tr>
<tr>
<td></td>
<td>Source</td>
<td>ChemBridge DiverSet library (ChemBridge Corporation; CA, USA)</td>
</tr>
<tr>
<td>Screen</td>
<td>Format</td>
<td>384-well tissue culture plates (Greiner Bio-One; NC, USA)</td>
</tr>
<tr>
<td></td>
<td>Concentration(s) tested</td>
<td>5µM compound, 0.025% DMSO</td>
</tr>
<tr>
<td></td>
<td>Plate controls</td>
<td>0.025% DMSO (negative control), 4µM 4-hydroxytamoxifen (positive control)</td>
</tr>
<tr>
<td></td>
<td>Reagent/compound dispensing system</td>
<td>BiomekFX (Beckman Coulter; CA, USA)</td>
</tr>
<tr>
<td></td>
<td>Detection instrument and software</td>
<td>GE/Amersham InCell 1000 high-throughput microscopy system</td>
</tr>
<tr>
<td></td>
<td>Assay validation/QC</td>
<td>Z’ = 0.2-0.6</td>
</tr>
<tr>
<td></td>
<td>Correction factors</td>
<td>Due to lamp variations throughout the imaging process, shade correction and plate-to-plate normalization techniques were utilized by equilibrating plate and global medians. Images were also corrected for compound autofluorescence (imaging of the red channel). Green pixel intensity contained within cytoplasmic masks were normalized to the total number of cells per well to give a percent GFP positive cells quotient.</td>
</tr>
<tr>
<td></td>
<td>Normalization</td>
<td></td>
</tr>
<tr>
<td>Post-HTS analysis</td>
<td>Hit criteria</td>
<td>Hits were determined by applying a threshold intensity gates on the green channel and surveying the number of cells above or below that gate in a given well by analyzing the average green pixel intensity under the cytoplasmic mask.</td>
</tr>
<tr>
<td></td>
<td>Hit rate</td>
<td>0.2%</td>
</tr>
<tr>
<td></td>
<td>Primary confirmatory assay</td>
<td>Primary confirmatory assay (retesting initial hits over a range of doses)</td>
</tr>
<tr>
<td></td>
<td>Additional assay(s)</td>
<td>Counterscreen for estrogenic activity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Secondary assay for endogenous insulin mRNA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BIM5078 (1-(2-chloro-5-nitrobenzenesulfonyl)-2-methylbenzimidazole) is commercially available from Princeton Bio (NJ, USA), CAS 337506-43-1, Cat# OSSL290849. Sample purity of &gt;95% was confirmed by $^1$H NMR and LC/MS.</td>
</tr>
<tr>
<td></td>
<td>Confirmation of hit purity and structure</td>
<td></td>
</tr>
<tr>
<td>ID</td>
<td>INITIAL SCREEN</td>
<td>IC50 OR EC50</td>
</tr>
<tr>
<td>---------</td>
<td>----------------</td>
<td>--------------</td>
</tr>
<tr>
<td>5120083</td>
<td>ACTIVATOR</td>
<td>1.15</td>
</tr>
<tr>
<td>5122762</td>
<td>REPRESSOR</td>
<td>X</td>
</tr>
<tr>
<td>5157189</td>
<td>ACTIVATOR</td>
<td>2.10</td>
</tr>
<tr>
<td>5185182</td>
<td>ACTIVATOR</td>
<td>X</td>
</tr>
<tr>
<td>5359535</td>
<td>REPRESSOR</td>
<td>0.15</td>
</tr>
<tr>
<td>5544524</td>
<td>ACTIVATOR</td>
<td>1.93</td>
</tr>
<tr>
<td>BIM5078</td>
<td>REPRESSOR</td>
<td>0.18</td>
</tr>
<tr>
<td>5811370</td>
<td>REPRESSOR</td>
<td>X</td>
</tr>
<tr>
<td>5816458</td>
<td>REPRESSOR</td>
<td>0.21</td>
</tr>
<tr>
<td>5849200</td>
<td>REPRESSOR</td>
<td>0.15</td>
</tr>
<tr>
<td>5856345</td>
<td>REPRESSOR</td>
<td>X</td>
</tr>
<tr>
<td>5867913</td>
<td>REPRESSOR</td>
<td>0.11</td>
</tr>
<tr>
<td>5994286</td>
<td>REPRESSOR</td>
<td>X</td>
</tr>
<tr>
<td>6058051</td>
<td>ACTIVATOR</td>
<td>X</td>
</tr>
<tr>
<td>6069474</td>
<td>REPRESSOR</td>
<td>X</td>
</tr>
<tr>
<td>6123012</td>
<td>REPRESSOR</td>
<td>X</td>
</tr>
<tr>
<td>6139941</td>
<td>ACTIVATOR</td>
<td>X</td>
</tr>
<tr>
<td>6238880</td>
<td>ACTIVATOR</td>
<td>X</td>
</tr>
<tr>
<td>6239484</td>
<td>REPRESSOR</td>
<td>X</td>
</tr>
<tr>
<td>6240540</td>
<td>REPRESSOR</td>
<td>X</td>
</tr>
</tbody>
</table>
4 CONCLUSIONS

4.1 Summary of work

The major finding of this dissertation is the discovery of a novel and potent class of ligands for the nuclear receptor HNF4α through a high-throughput screen for modulators of the insulin promoter. The utility of such ligands is vast, and in this dissertation we reveal one potential pharmacologic application in hepatocellular carcinoma.

We report the development and application of a high-throughput screening assay based on a highly engineered cell line derived from human β-cells and adapted to express the human insulin gene in a regulated and reproducible manner. HTS using promoter reporter assays is a powerful approach to identifying molecules capable of modulating central signaling pathways involved in organ function and disease pathogenesis. However, identifying the precise molecular target of hits validated from these assays can be challenging, particularly in the case of promoters under the
complex control of a number of signaling pathways and transcription factors, such as the human insulin promoter.

First, we describe how we applied this HTS assay to indentify the insulin promoter as a previously unrecognized target of a structurally diverse group of antipsychotics. Because the insulin promoter is a known target of diabetogenic stimuli [66], we hypothesized that HTS for modulators could reveal valuable information about the pathogenesis of diabetes. It has previously been documented that antipsychotics are strongly associated with metabolic side effects, including obesity, insulin resistance, and type II diabetes. Our finding that chronic exposure of a number of antipsychotics inhibits insulin gene expression in T6PNE cells is thus consistent with metabolic profile observed clinically.

We then expanded the HTS assay to screen a diverse synthetic chemical library and identified a structurally distinct inhibitor of insulin gene expression. Through a cheminformatic analysis and a direct binding assay, we identified the molecular target of this compound to be HNF4α. Our success in identifying the precise molecular target of a small-molecule hit discovered in a cell-based assay is a rare example of successful hit-to-lead discovery.

Once the target was identified, we explored the utility of the HNF4α bioactive ligand as a therapeutic for hepatocellular carcinoma. In addition to its ability to repress insulin gene expression in T6PNE and modulate HNF4α target genes, this class of HNF4α ligands exhibits selective toxicity to transformed cells in vitro, including in HepG2 and Hep3b cells. Apoptotic signaling cascades by these ligands were activated in an orthotopoic xenograft model of hepatocellular carcinoma in vivo, but no survival
advantage was observed. This was likely the result of high levels of hepatic metabolism and inadequate tissue penetration of the compound. Further optimization of these ligands is warranted for pharmacologic intervention.

4.2 Therapeutic implications

In addition to contributing to our knowledge of the physiological function of HNF4α, these small molecules introduce a platform for pharmacologic intervention in diseases governed by HNF4α, including diabetes and cancer. Given the broad influence of the nuclear receptor family of transcription factors, including HNF4α, on tissue development and homeostasis, they have long been considered amongst the most druggable targets for a large number of pathological states. Of the 48 nuclear receptors that have been identified, 17 are targets of currently marketed drugs, which is particularly striking because nearly half of these receptors are “orphans”, for which there are no known endogenous ligands or even synthetic chemical tool with which gene-regulatory function can be studied [115]. Thus, our discovery of novel ligands directed at a previously untargeted nuclear receptor, provides an extraordinary opportunity for drug development.

At the translational level, HNF4α unmistakably plays a central role in many disease processes. In MODY1, HNF4α is directly mutated and in type II diabetes it is strongly downregulated [116]. In other conditions such as liver disease and cancer, HNF4α has an apparent, but less-defined role. Not only will the discovered ligands targeting HNF4α provide a tool with which to probe the physiology of normal
signaling pathways implicated in receptor regulation and validate its role in disease processes, but they open the door to developing therapeutics for these disorders.

4.3 Future directions

Further development of compounds structurally related to those discussed in this dissertation could yield pharmacologic interventions for diseases in which HNF4α plays a role, including diabetes and cancer, as well as providing a robust tool with which to study the physiological role of HNF4α. The utility of those compounds already identified in our assays in studying in vivo processes is grossly limited by their pharmacokinetic properties, especially with respects to hepatic metabolism and tissue penetration.

Compelling evidence for incomplete drug delivery into tumor tissue comes from studies with doxorubicin, as its intrinsic fluorescence has facilitated direct detection of drug penetration [113]. Similarly, we have the unique ability to use steatosis as a biomarker for the penetration of HNF4α antagonists. As described in the dissertation, lipid accumulation in regions immediately surrounding hepatic blood vessels suggests that our prototypic ligand is unable to penetrate deeper into the tissue. Methods to overcome this are reported in the literature, and include targeting the tumor’s blood supply using ligand-targeted liposomes, as has been done with doxorubicin [117]. Additional work is warranted to evaluate whether this would be an effective mechanism to enhance distribution of the benzimidazole-based HNF4α antagonists and thereby prolong survival in the setting of hepatocellular carcinoma.
Also of great consequence is the very high level of hepatic metabolism observed \textit{in vitro} in pharmacokinetic studies with these small molecules. Metabolism of the active compound into pharmacologically inactive metabolites may limit penetration of parent drug deeper into the tissue. Developing analog with less hepatic conversion may also enhance active drug distribution and thereby efficacy. Nonetheless, our data demonstrate proof of principle that HNF4α antagonists are cytotoxic to hepatocellular carcinoma, and other transformed cell lines.

Finally, decreased levels of HNF4α in the setting of diabetes suggest that the development of HNF4α-selective agonists may be a valuable pharmacologic intervention. We hypothesize that HNF4α is protecting the β-cell from lipotoxic stress by stabilizing transcriptional machinery on the insulin promoter. This is consistent with disruption of PDX-1 and E47 binding to the insulin promoter in the presence of the HNF4α antagonist, BIM5078.

Future experiments using structure-activity relationship (SAR) are warranted to develop HNF4α bioactive agonists derived from the antagonists we have already described throughout this dissertation. The \textit{in silico} docking studies of the antagonists to the HNF4α ligand binding pocket (LBD) that we have described should guide substituent replacement to yield molecules with tighter binding and reduced steric clashes in the LBD, resulting in an HNF4α agonist.
References


23. Pearson ER, Boj SF, Steele AM, Barrett T, Stals K, Shield JP, Ellard S, Ferrer J, Hattersley AT. Macrosomia and hyperinsulinaemic hypoglycaemia in


118


