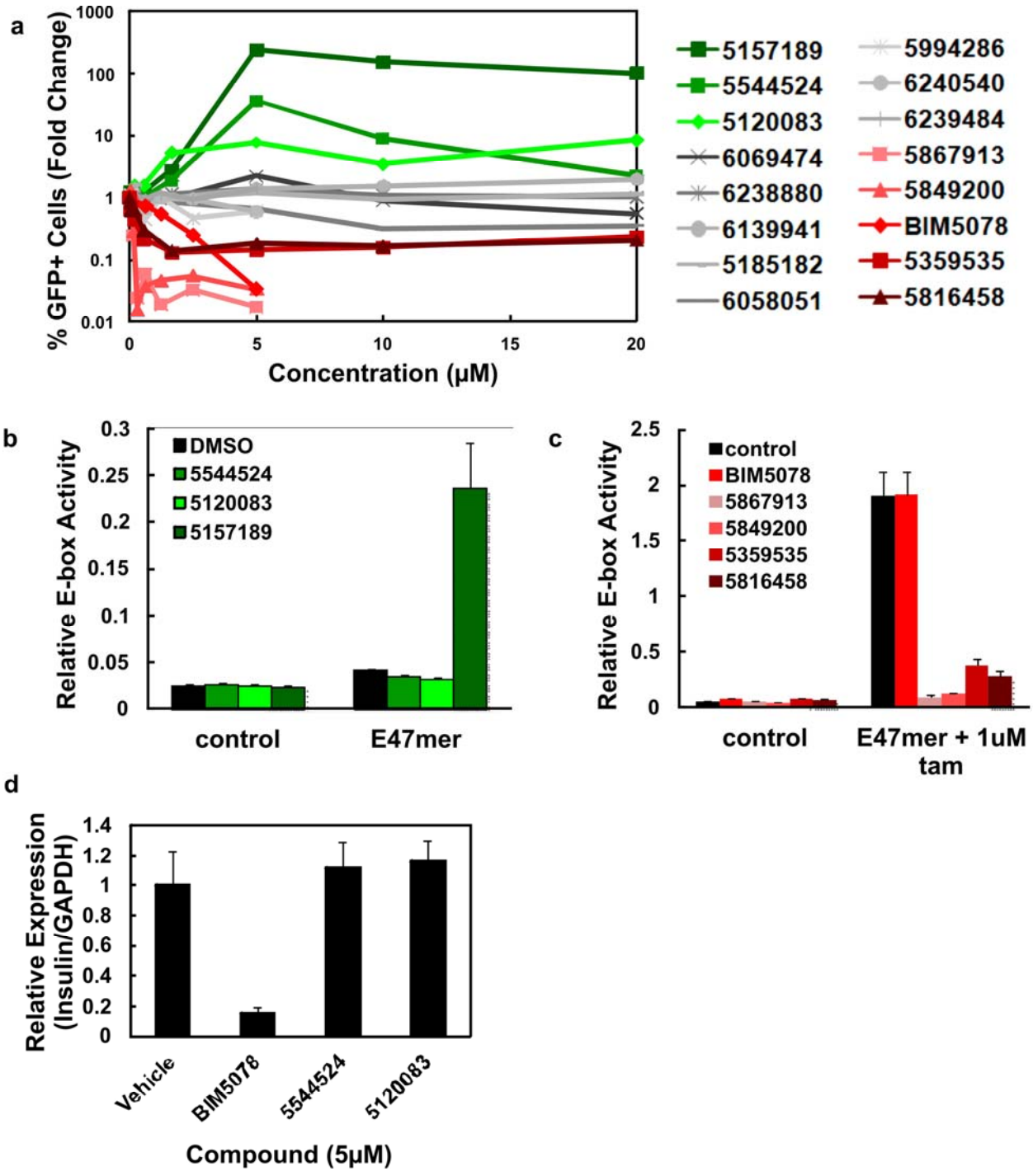
Chemistry & Biology, Volume *19*

Supplemental Information

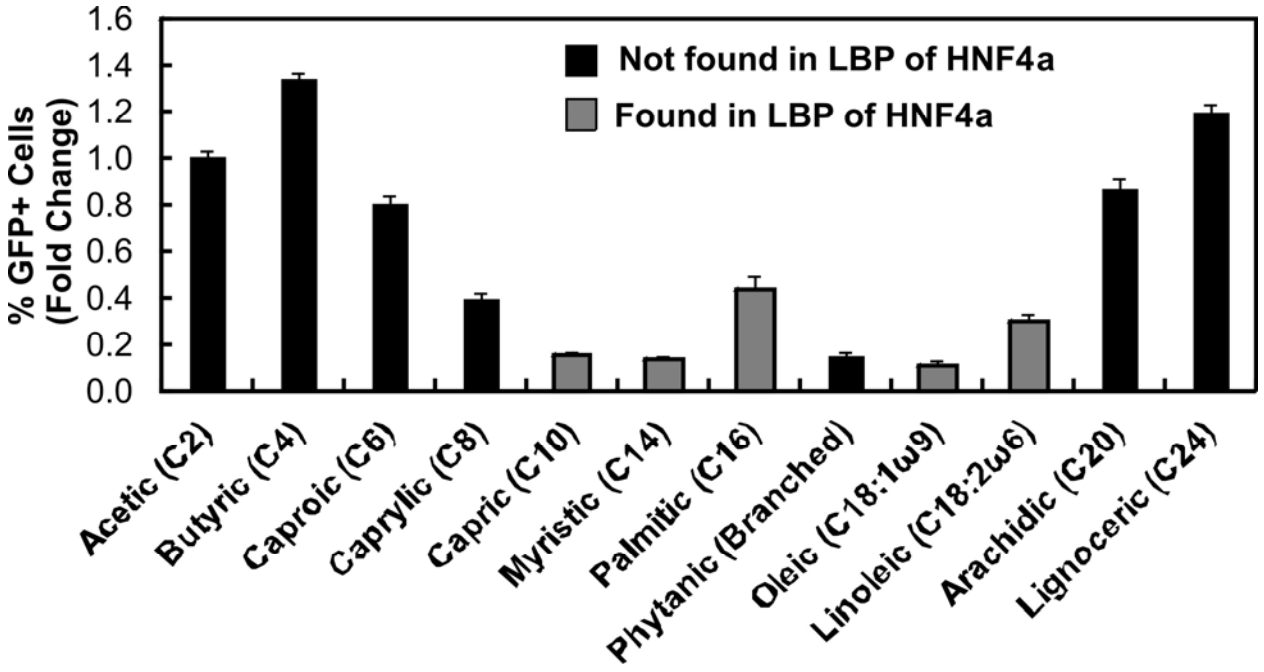
HNF4α Antagonists Discovered by a High-Throughput Screen for Modulators of the Human Insulin Promoter

Alice Kiselyuk, Seung-Hee Lee, Suzette Farber-Katz, Mingjun Zhang, Sonalee Athavankar, Tom Cohen, Anthony B. Pinkerton, Mao Ye, Paul Bushway, Adam D. Richardson, Heather A. Hostetler, Mariam Rodriguez-Lee, Li Huang, Benjamin Spangler, Layton Smith, Jennifer Higginbotham, John Cashman, Hudson Freeze, Pamela Itkin-Ansari, Marcia I. Dawson, Friedhelm Schroeder, Yong Cang, Mark Mercola, and Fred Levine

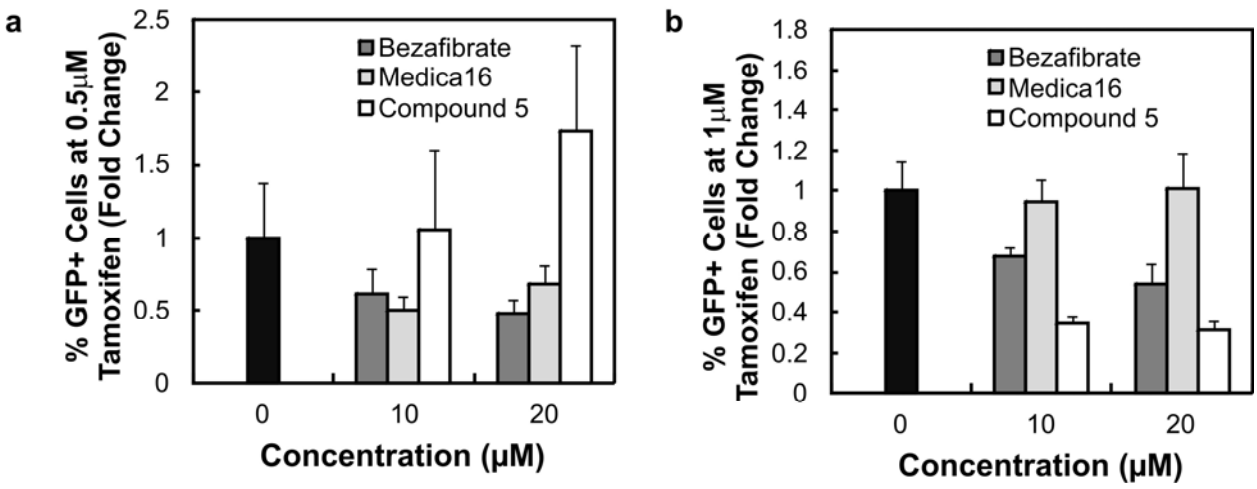
Inventory of Supplemental Information



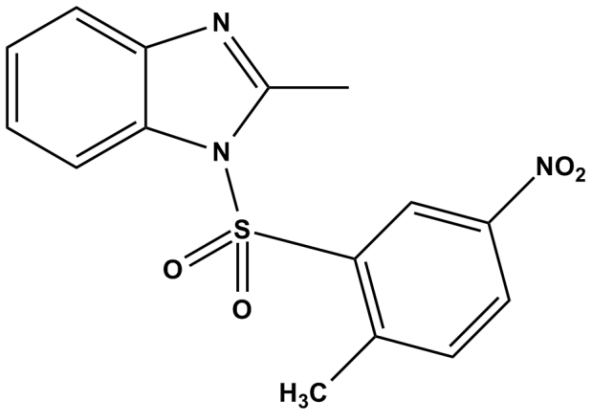
Supplemental Figure 1



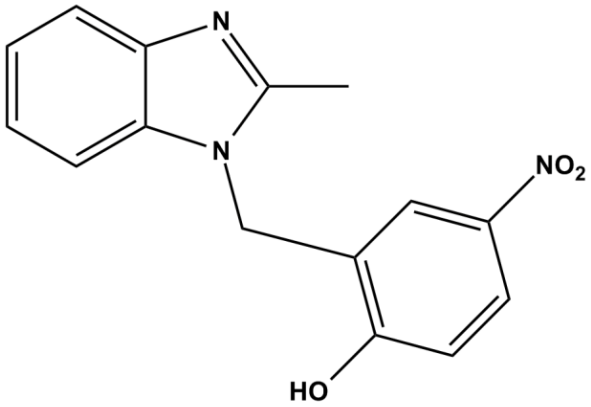
Supplemental Figure 2



Supplemental Figure 3

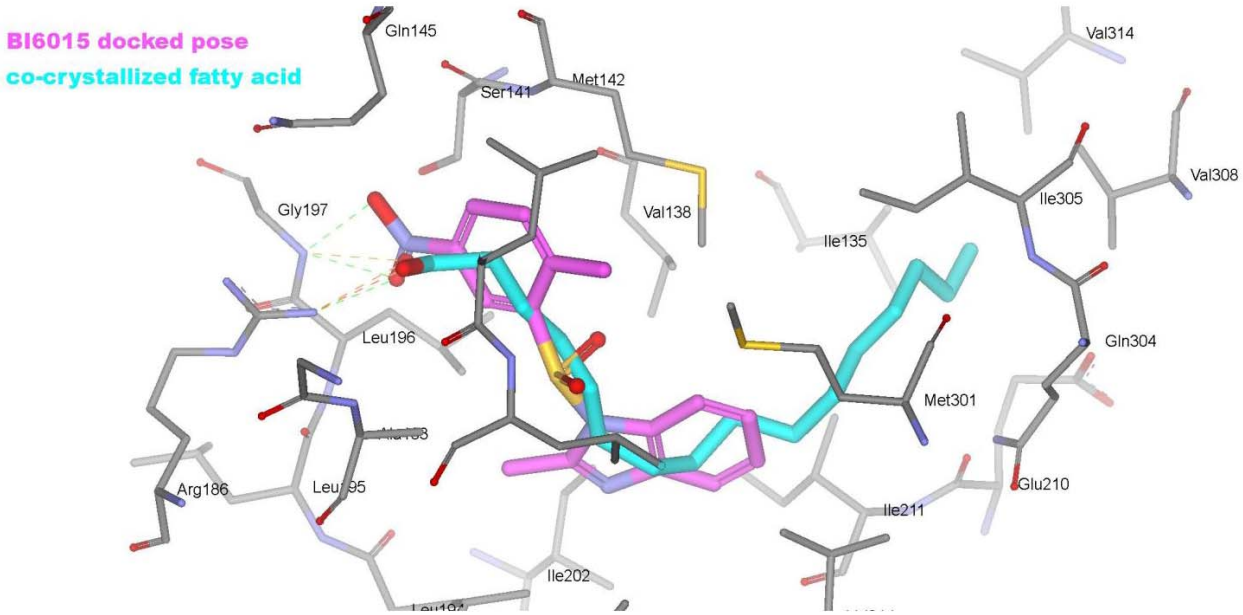


BI6015

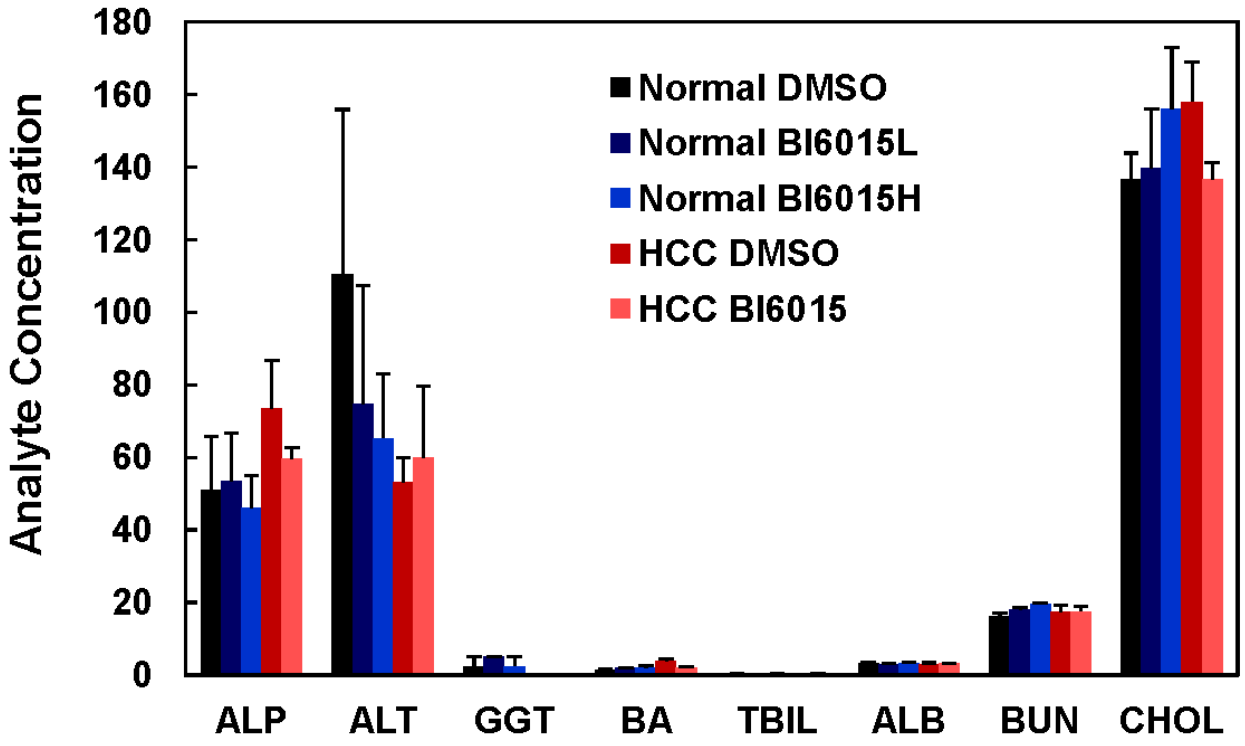


BI6018

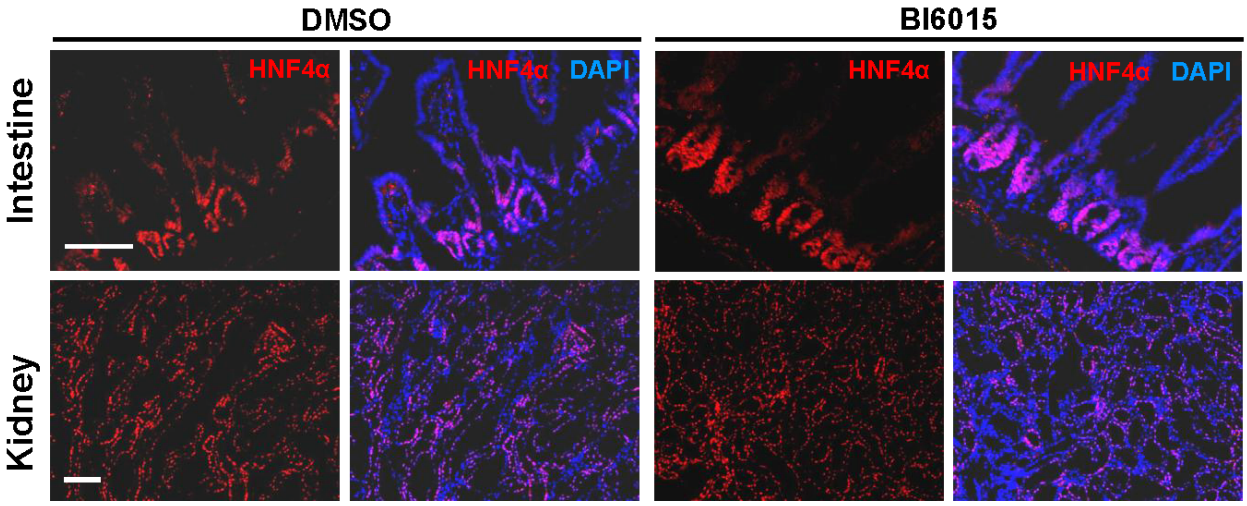
Supplemental Figure 4



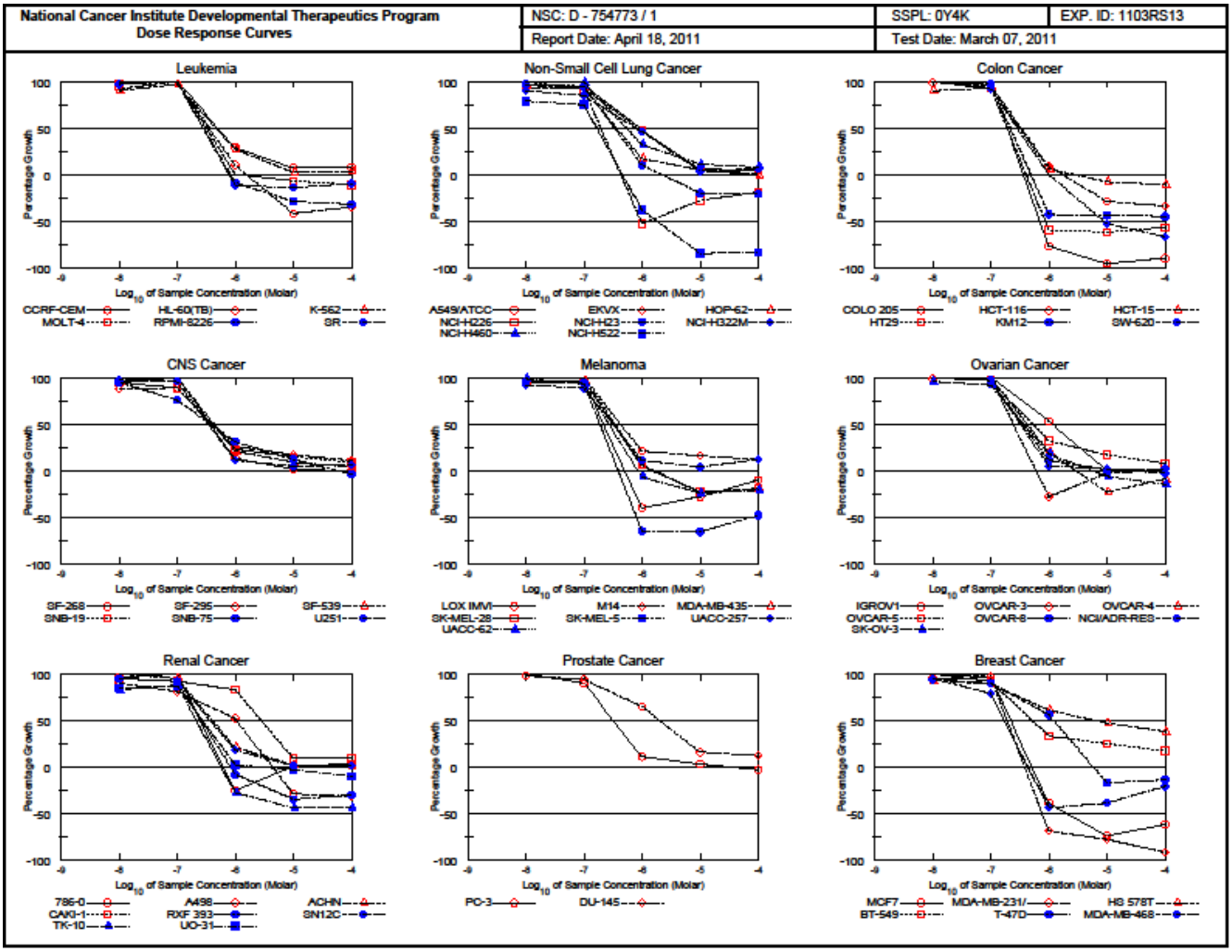
Supplemental Figure 5



Supplemental Figure 6



Supplemental Figure 7



Supplemental Figure 8

Supplemental Figure 1 | BIM5078 passed confirmatory and counter screens. (**a**) Compounds found to modulate the insulin-eGFP transgene in the primary screen were confirmed in dose-response assays. 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 tamoxifen. 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. Compound 5157189 was found be estrogenic. (**c**) Similarly, compounds found to decrease the functional activity of the insulin-eGFP transgene (**a**) were tested for their ability to prevent the translocation of $E47^{MER}$. Compounds 5867913, 5849200, 5359535 and 5816458 were found to be estrogenic antagonists. (**d**) Compounds that did not interact with $E47^{MER}$ were tested for their ability to modulate endogenous insulin mRNA. BIM5078, but not 5544524 or 5120083, had activity on the endogenous insulin promoter. Values represent the mean ± SE.

Supplemental Figure 2 | Fatty acids bound to HNF4α preferentially inhibit exogenous insulin promoter activity. T6PNE cells were treated with 0.12 mM fatty acids for 48 hours. Effects on the exogenous insulin promoter in T6PNE is reported as percent GFP+ cells, as determined by imaging the green channel and normalizing to the total number of cells per well. Values represent the mean ± SE, n=8.

Supplemental Figure 3 | Effect of reported HNF4α ligands on insulin promoter activity. T6PNE cells were treated with bezafibrate, Medica-16 or a nitro-naphthofuran derivative ("Compound 5") for 48 hours in the presence of either 0.5 μM (**a**) or 1 μM (**b**) tamoxifen. Effects on the exogenous insulin promoter transgene in T6PNE are reported as percent GFP+ cells. Values represent the mean ± SE, n=3.

Supplemental Figure 4 | Structures of BI6015 and BI6018.

Supplemental Figure 5 | Docking of BI6015 (purple) in the LBD of HNF4γ with linoleic acid (cyan) crystallized. The nitro group of BI6015 forms hydrogen bonds with the backbone N of Gly197 and the side chain guanadinium group of Arg 186. The remaining parts of BI6015 primarily make hydrophobic interactions with the binding pocket, which is for the most part hydrophobic in nature. The empirical docked score (GOLD Fitness score) for this pose is 40.06, similar to that for HNF4α.

Supplemental Figure 6 | 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 30 mg/kg/day (Normal BI6015H, n=4) or 10 mg/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 30 mg/kg/2day (HCC BI6015, n=4) for 29-36 days. Values represent the mean ± SE.

Supplemental Figure 7 | HNF4α expression does not change in the intestine or kidney in mice receiving IP injection of BI6015. HNF4α expression was assessed as described in the Methods. Scale bar, 100 μm.

Supplemental Figure 8 | NCI Panel. Data were generated by the NCI Developmental Therapeutics Program, as previously described (Shoemaker, 2006).

Supplemental Table 1 | Small molecule screening data

Category	Parameter	Description
Assay	Type of assay	Cell-based
	Target	Human insulin promoter
	Primary measurement	Detection of green intensity (green fluorescence protein driven by the human insulin promoter)
	Key reagents	tamoxifen (Sigma-Aldrich; MO, USA)
		DAPI (Invitrogen; CA, USA)
	Assay protocol	Kiselyuk, A. et al. Phenothiazine neuroleptics signal to the human insulin promoter as revealed by a novel high-throughput screen. *J Biomol Screen* **15**, 663-70
Library	Library size	8,064
	Library composition	Diverse small molecules
	Source	ChemBridge DiverSet library (ChemBridge Corporation; CA, USA)
Screen	Format	384-well tissue culture plates (Greiner Bio-One; NC, USA)
	Concentration(s) tested	5μM compound, 0.025% DMSO
	Plate controls	0.025% DMSO (negative control), 4μM tamoxifen (positive control)
	Reagent/ compound dispensing system	BiomekFX (Beckman Coulter; CA, USA)
	Detection instrument and software	GE/Amersham InCell 1000 high-throughput microscopy system
	Assay validation/QC	Z' = 0.2-0.6
	Correction factors	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).
	Normalization	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.
Post-HTS analysis	Hit criteria	Hits were determined by applying a threshhold 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.
	Hit rate	0.2%
	Additional assay(s)	Primary confirmatory assay (retesting initial hits over a range of doses)
		Counterscreen for estrogenic activity
		Secondary assay for endogenous insulin mRNA
	Confirmation of hit purity and structure	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 >95% was confirmed by ^{1}H NMR and LC/MS.

Supplemental Table 2 | List of genes affected both by genetic deletion of HNF4α *in vivo* and pharmacologic inhibition of HNF4α *in vitro*. 36% of the genes altered by genetic deletion of HNF4α in the mouse *in vivo* are either identical or closely related to genes modulated by pharmacologic inhibition of HNF4α in the human cell line, T6PNE.

Exact Match

Rnase4	*Mgst1*	*Mt2*	*Sdcbp*
Sez6l	*Slc23a2*	*PENK*	*Csrp1*
Syne1	*Rps6ka5*	*Lgmn*	*Cd24a*
Comt	*Smarcd1*	*Nucb2*	*Tat*
Dapp1	*S100a11*	*Srp14*	*CTRB*

Family Match

Ugtb	*Fbxo*	*Mrpl*	*Stgal*	*Nrxn*	*Stab*
Igf	*Rabb*	*Sdcbp*	*Prkarb*	*Alscr*	*Igfbp*
Slca	*Raba*	*Lsm*	*Tmem*	*Fxyd*	*Rdh*
Nkx2-2	*Bgalt*	*Cda*	*Prss*	*Mt*	*Mal*
Scarf	*Tnfrsfa*	*Ndufb*	*Galnt*	*St*	*Tbcd*
Rnase	*Fkbp*	*Anxa*	*Pdia*	*Ccnd*	*Sa*

Supplemental Table 3 | List of genes affected by both pharmacologic inhibition of of HNF4α and induction of E47. Analysis was restricted to genes altered at least two-fold. The expression of 214 genes were significantly altered by BIM5078 at least two-fold. Of these, 67 were also modulated by E47 induction through 4-hydroxytamoxifen. This association was enhanced when only the genes containing E-boxes were compared. 96 E-box containing genes were altered by BIM5078. 42 of these were also altered by E47 induction

Overlap of genes altered by BIM5078 and E47-responsive genes

PRSS2	*RAB26*	*DBH*	*CHRNA3*	*PVALB*	*SPRY1*	*DUSP6*
AQP1	*C10orf10*	*INMT*	*FYN*	*LRIG1*	*FATE1*	*SOCS2*
CALY	*SERPINI1*	*GAL*	*BCL6*	*MME*	*EFHD1*	*CDKN2A*
TNNT2	*CALML3*	*IGFBP1*	*CALML5*	*ATP6V0A4*	*D4S234E*	*C19orf33*
CHRNA1	*DCN*	*DUSP4*	*CREG1*	*HMP19*	*HTR3C*	*PRKCDBP*
CABP7	*CELA3A*	*FGFBP1*	*SULF2*	*FXYD2*	*FAM174B*	*AXL*
ACTG2	*ANGPT2*	*CAMK2B*	*FGD5*	*PPP1R16B*	*TEX14*	*DOCK2*
BMP4	*MOXD1*	*NPTX2*	*GNG8*	*HCP5*	*APOD*	
MYBPH	*BMP7*	*TIPARP*	*RASL11B*	*HBEGF*	*VCX*	
C6orf15	*MAN1C1*	*KIAA1199*	*MAL*	*SPRED2*	*S100A8*	

Overlap of genes altered by BIM5078 and E47-responsive genes containing E-boxes

PRSS2	*DCN*	*BCL6*	*PPP1R16B*	*TNNT2*	*TUBB2B*
AQP1	*CELA3A*	*FYN*	*TMEM178*	*ACTG2*	*S100A8*
CALY	*BMP7*	*GNG8*	*SPRED2*	*MYBPH*	*ANGPT2*
CHRNA1	*NPTX2*	*HMP19*	*AXL*	*DBH*	*FXYD2*
BMP4	*DUSP4*	*LRIG1*	*FAM174B*	*MAN1C1*	*SOCS2*
RAB26	*CAMK2B*	*HBEGF*	*XIRP1*	*PVALB*	*RASL11B*
SERPINI1	*ACOT11*	*ATP6V0A4*	*DUSP6*	*APOD*	*INMT*

Supplemental Data | Ricerca Pharmacology Data Report

Ricerca Biosciences, LLC Pharmacology Data Report

On Compound SBR-4, BI6015 For Sanford-Burnham Medical Research Institute

Work Order Number:	1-1040466-0	**Services Being Reported:**	Combination Screen
Study Number:	AA98499	**Alternative Work Order No:**	
Quote No:	21424-1	**Purchase Order Number:**	125555
Compound Information:		**Total # of Assays:**	43
Compound Code:	SBR-4		
Alternative Code 1:	BI6015		
Alternative Code 2:			
Alternative Code 3:			
Ricerca Internal #:	1141740		
	331.54		

Sponsor:
Sanford-Burnham Medical Research Institute
10901 N Torrey Pines Rd
La Jolla, CA 92037
United States

Undertaken at:
Ricerca Biosciences, LLC
Pharmacology Laboratories
158 Li-Teh Road, Peitou
Taipei, Taiwan 112
Taiwan R.O.C.

Date of Study: September 17, 2010 - September 30, 2010

Study Directors: Kun-Yuan Lin, Ricerca Biosciences, LLC
Kuo-Hsin Chen, Ricerca Biosciences, LLC

Distribution: Sanford-Burnham Medical Research Institute

"This study was conducted according to the procedures described in this report. All data presented are authentic, accurate and correct to the best of our knowledge."

Kun-Yuan Lin
Study Director for Animal Assays

Kuo-Hsin Chen
Study Director for Biochemical Assays

Peter Chiu, Ph.D
Technical Director

TABLE OF CONTENTS

SUMMARY

STUDY OBJECTIVE

To evaluate, in Enzyme, and Radioligand Binding assays, the activity of compound BI6015 (SBR-4, PT# 1141740).

METHODS

Methods employed in this study have been adapted from the scientific literature to maximize reliability and reproducibility. Reference standards were run as an integral part of each assay to ensure the validity of the results obtained. Assays were performed under conditions described in the accompanying "Methods" section of this report. The literature reference(s) for each assay are in the "Literature References" section. If either of these sections were not originally requested with the accompanying report, please contact us at the number below for a printout of either of these report sections.

Where presented, IC_{50} values were determined by a non-linear, least squares regression analysis using MathIQ™ (ID Business Solutions Ltd., UK). Where inhibition constants (K_i) are presented, the K_i values were calculated using the equation of Cheng and Prusoff (Cheng, Y., Prusoff, W.H., Biochem. Pharmacol. 22:3099-3108, 1973) using the observed IC_{50} of the tested compound, the concentration of radioligand employed in the assay, and the historical values for the K_D of the ligand (obtained experimentally at **Ricerca Biosciences, LLC**). Where presented, the Hill coefficient (n_H), defining the slope of the competitive binding curve, was calculated using MathIQ™. Hill coefficients significantly different than 1.0, may suggest that the binding displacement does not follow the laws of mass action with a single binding site. Where IC_{50}, K_i, and/or n_H data are presented without Standard Error of the Mean (SEM), data are insufficient to be quantitative, and the values presented (K_i, IC_{50}, n_H) should be interpreted with caution.

RESULTS

A summary of results meeting the significance criteria is presented in the following sections. Complete results are presented under the section labeled "Experimental Results". Individual responses, if requested, are presented in the appendix to this report.

SUMMARY/CONCLUSION

Significant results are displayed in the following table(s) in rank order of potency for estimated IC_{50} and/or K_i values.

SUMMARY OF SIGNIFICANT PRIMARY RESULTS

Biochemical assay results are presented as the percent inhibition of specific binding or activity throughout the report. All other results are expressed in terms of that assay's quantitation method (see Methods section).

- *For primary assays, only the lowest concentration with a significant response judged by the assays' criteria, is shown in this summary.*
- *Where applicable, either the secondary assay results with the lowest dose/concentration meeting the significance criteria or, if inactive, the highest dose/concentration that did not meet the significance criteria is shown.*
- *Unless otherwise requested, primary screening in duplicate with quantitative data (e.g., IC50± SEM, Ki ± SEM and nH) are shown where applicable for individual requested assays. In screening packages, primary screening in duplicate with semi-quantitative data (e.g., estimated IC50, Ki and nH) are shown where applicable (concentration range of 4 log units); available secondary functional assays are carried out (30 μM) and MEC or MIC determined only if active in primary assays >50% at 1 log unit below initial test concentration.*
- *Please see Experimental Results section for details of all responses.*

Significant responses (≥ 50% inhibition or stimulation for Biochemical assays) were noted in the primary assays listed below:

PRIMARY TESTS

CAT. #	PRIMARY BIOCHEMICAL ASSAY	SPECIES	CONC.	% INH.	IC_{50} *	K_i	n_H
118070	CYP450, 2C19	hum	10 μM	94			
214600	Calcium Channel L-Type, Dihydropyridine	rat	10 μM	83			

‡ Partially soluble in *in vitro* test solvent.
* A standard error of the mean is presented where results are based on multiple, independent determinations.
bov=bovine; ham=hamster; hum=human

EXPERIMENTAL RESULTS - BIOCHEMICAL ASSAYS

Cat. #	TARGET	BATCH*	SPP.	n=	CONC.	†% INHIBITION	IC_{50}	K_i	n_H	R
						% -100 -50 0 50 100				
118050	CYP450, 1A2	280597	hum	2	10 µM	7				
♦ 118070	CYP450, 2C19	280599	hum	2	10 µM	94				
118060	CYP450, 2C9	280598	hum	2	10 µM	7				
118080	CYP450, 2D6	280600	hum	2	10 µM	-26				1
118090	CYP450, 3A4	280635	hum	2	10 µM	-11				
200510	Adenosine A_1	280524	hum	2	10 µM	18				
200610	Adenosine A_{2A}	280525	hum	2	10 µM	8				
203100	Adrenergic α_{1A}	280526	rat	2	10 µM	19				
203200	Adrenergic α_{1B}	280705	rat	2	10 µM	-2				
203620	Adrenergic α_{2A}	280528	hum	2	10 µM	11				
204010	Adrenergic β_1	280530	hum	2	10 µM	7				
204110	Adrenergic β_2	280531	hum	2	10 µM	2				
285010	Androgen (Testosterone) AR	280625	rat	2	10 µM	13				
♦ 214600	Calcium Channel L-Type, Dihydropyridine	280706	rat	2	10 µM	83				
217030	Cannabinoid CB_1	280536	hum	2	10 µM	17				
219500	Dopamine D_1	280539	hum	2	10 µM	29				
219700	Dopamine D_{2S}	280540	hum	2	10 µM	17				
226010	Estrogen ERα	280571	hum	2	10 µM	4				
226050	Estrogen ERβ	280572	hum	2	10 µM	-5				
226600	$GABA_A$, Flunitrazepam, Central	280544	rat	2	10 µM	13				
226500	$GABA_A$, Muscimol, Central	280543	rat	2	10 µM	-16				
232020	Glucocorticoid	280649	hum	2	10 µM	9				
233000	Glutamate, NMDA, Phencyclidine	280545	rat	2	10 µM	12				
239610	Histamine H_1	280546	hum	2	10 µM	17				
241000	Imidazoline I_2, Central	280548	rat	2	10 µM	6				
252710	Muscarinic M_2	280550	hum	2	10 µM	-4				
252810	Muscarinic M_3	280551	hum	2	10 µM	2				
258590	Nicotinic Acetylcholine	280557	hum	2	10 µM	-11				
258700	Nicotinic Acetylcholine α1, Bungarotoxin	280558	hum	2	10 µM	9				
260410	Opiate µ (OP3, MOP)	280556	hum	2	10 µM	-3				
264500	Phorbol Ester	280560	mouse	2	10 µM	8				

* Batch: Represents compounds tested concurrently in the same assay(s). ‡ Partially soluble in *in vitro* test solvent.
♦ Denotes item meeting criteria for significance
† Results with ≥ 50% stimulation or inhibition are highlighted.
R=See Remarks (if any) at end of this section.
bov=bovine; ham=hamster; hum=human

EXPERIMENTAL RESULTS - BIOCHEMICAL ASSAYS

Cat. #	TARGET	BATCH*	SPP.	n=	CONC.	†% INHIBITION	IC_{50}	K_i	n_H	R
						% -100 -50 0 50 100				
265600	Potassium Channel [K_{ATP}]	280871	ham	2	10 µM	-4				
265900	Potassium Channel hERG	280563	hum	2	10 µM	31				
268000	Progesterone	280700	bov	2	10 µM	-14				
268420	Prostanoid EP_4	280565	hum	2	10 µM	11				
269500	Retinoid X Receptor RXRα	280659	hum	2	10 µM	19				
270000	Rolipram	280566	rat	2	10 µM	21				
271700	Serotonin (5-Hydroxytryptamine) 5-HT_{2B}	280567	hum	2	10 µM	15				
278110	Sigma σ_1	280568	hum	2	10 µM	7				
279510	Sodium Channel, Site 2	280924	rat	2	10 µM	17				
285900	Thyroid Hormone	280624	rat	2	10 µM	7				
204410	Transporter, Norepinephrine (NET)	280704	hum	2	10 µM	35				
288000	Vitamin D_3	280811	hum	2	10 µM	5				

* Batch: Represents compounds tested concurrently in the same assay(s). ‡ Partially soluble in *in vitro* test solvent.
♦ Denotes item meeting criteria for significance
† Results with ≥ 50% stimulation or inhibition are highlighted.
R–See Remarks (if any) at end of this section.
bov–bovine; ham–hamster; hum–human

SUPPLEMENTAL METHODS

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% CO_2, 37°C. To induce E47 activity, 0.5 or 1 μM 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% CO_2, 37°C. HeLa cells were grown at 10% CO_2, 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.

Compound library screening in T6PNE: A collection of 8,064 chemically diverse small molecules, a subset of the ChemBridge DiverSet library (ChemBridge Corporation; CA, USA) was screened in a high-throughput assay for functional inhibition of the insulin promoter in T6PNE expressing a human insulin promoter-eGFP transgene, as previously described (Kiselyuk et al., 2010). 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 tamoxifen to induce a submaximal level of insulin expression. Compound addition (active compound in DMSO or vehicle alone) occurred 24 hours after tamoxifen 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 (Kiselyuk et al., 2010).

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.

Counterscreen for estrogenic activity. As previously reported (Kiselyuk et al., 2010), 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 $E47^{MER}$. 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, $E47^{MER}$ 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 tamoxifen and/or compound or vehicle (DMSO). Cells were then lysed and assayed for luciferase activity using the Promega Dual Luciferase kit (Promega Corp., Madison, Wisconsin, USA), and luminescence was measured using the Veritas Microplate Luminometer (Turner Biosystems, Sunnyvale, California, USA). Data were normalized to *Renilla* luciferase (pRL-TK) and are expressed as fold-change over vehicle alone.

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 conducted 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:

Insulin	FWD: CTACCTAGTGTGCGGGGAAC REV: GCTGGTAGAGGGAGCAGATG
HNF4α	FWD: CAGAATGAGCGGGACCGGATC REV: CAGCAGCTGCTCCTTCATGGAC
Kip2	FWD: GGCGATCAAGAAGCTGTCC REV: GGGCTCTTTGGGCTCTAAAT
18S rRNA	FWD: GATATGCTCATGTGGTGTTG REV: AATCTTCTTCAGTCGCTCCA
GAPDH	FWD: ACAGTCAGCCGCATCTTCTT REV: AATGAAGGGGTCATTGATGG

Compounds. BIM5078 (1-(2´-chloro-5´-nitrobenzenesulfonyl)-2-methylbenzimidazole) is commercially available from Princeton Bio (NJ, USA), CAS 337506-43-1, Cat# OSSL290849. BIM5078 was also synthesized in larger quantity; 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, and is now commercially available from Princeton Bio (NJ, USA), CAS 93987-29-2, Cat# OSSL290848. Sample purity of >95% was confirmed by ^{1}H NMR and HPLC. 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 according to the procedure published by Le Guevel *et. al.*(2009) (Le Guevel et al., 2009). 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. Linoleic 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.
FK614 was prepared as described (Yamasaki et al., 1997). Spectral data (1H NMR and LCMS) matched that of the literature, and purity was determined to be >95% by LCMS.

PPRE-Luc reporter assays. PPRE assays were performed as previously described (Dawson et al., 2009).

***In silico* docking studies.** Docking was performed using GOLD from the Cambridge Crystallographic Data Centre (Cambridge, UK). Several poses were found that were within 10% of their empirical fitness scores, with the pose in **Fig. 2b** chosen as the best fit and overlap with the fatty acid.

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 (Petrescu et al., 2002). Full length HNF4α protein was recombinantly expressed and purified as previously described (Petrescu et al., 2002). Briefly, 100 nM of full-length HNF4α protein was titrated with increasing concentrations of each compound. A low concentration of HNF4α protein was used because full length HNF4α has a high tendency to form aggregates. Fluorescence emission spectra were obtained at 24°C with a PC1 photon counting spectrofluorometer (ISS Inc.; IL, USA), corrected for background (protein only and ligand only), and maximal intensities were measured at 330 nm upon excitation at 280 nm. The data were fitted to the rectangular hyperbolic equation: $F = F_{max}[L]/(EC_{50} + [L])$, where F is the fluorescence intensity changes corresponding to each [L] ligand concentration in the sample, F_{max} is the asymptotic value for maximal fluorescence intensity, EC_{50} is the ligand concentration resulting in half-maximal fluorescence change.

DARTS assay. HepG2 cells were treated with DMSO, BI6018, FK614, BIM5078 or BI6015 at a concentration of 20μM for 24hr.Total cell protein was extracted, measured by BCA protein assay (Thermo scientific).Each sample was split into three aliquots for proteolysis without (-) or with(+) Subtilisin (Sigma) and Coomassie (InstantBlue, expedeon) staining. Twenty μg of cell lysate was incubated with or without protease (20ng/ml subtilisin) for 30 minutes at room temperature. Western blot was then performed with primary anti-HNF4α polyclonal antibody (1:1000 dilution, Santa Cruz, 54kDa) and secondary HRP conjugated anti-goat IgG (1:2000 dilution, Jackson Immuno), detected with chemiluminescence ECL kit (Thermo Scientific). One of aliquot of each sample was loaded onto SDS/PAGE gels (Invitrogen) for InstantBlue staining (M : protein ladder), N=3.

siRNA studies: Ambion silencer siRNA (*HNF4α* siRNA ID#4013) was administered to 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% CO_2, followed by the addition of 1 µM tamoxifen. Transfected cells were incubated at 37°C for an additional 48 hours with tamoxifen. 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 was done as described above.

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:

hINS +40	FWD: GTAGGGGAGATGGGCTCTG REV: GCTTGATGGCCTCTTCTGAT
hINS -173	FWD: AGCTGTGAGCAGGGACAG REV: CTTGGTCGTCAGCACCTCTT
hINS –300	FWD: GGACAGCAGCGCAAAGAG REV: GGGCCACTTTCCACATTAGA

Gene expression profiling.
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). Downstream data analysis was performed using GeneSpring (Agilent; CA, USA).
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.

NextBio. Genes altered by ≥2-fold by BIM6078 treatment of 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. For consistency, analysis was restricted to the Molecular Signatures Database (MSigDB) containing motif gene sets annotated by the Broad Institute (Xie et al., 2005).

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 (Inoue et al., 2002) and 0.0625 µg/well of pRL-TK using Lipofectamine 2000 reagent (Life Technologies; 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.

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).

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.

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 neat propylene glycol 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.

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 (2×10^6 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. All studies with mice were approved by the SBMRI IACUC.

Mouse primary hepatocyte 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% CO_2 incubator on collagen coated plates.

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.

***In vitro* analysis of ADME properties.**

Assays were performed by the SBMRI Exploratory Pharmacology Core Facility.

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.

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).

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:

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

All reactions were run in triplicate, except negative controls (no NADPH), which were done as single incubations. 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.
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 acetronitrile (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).
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. Alsp, 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:

$$\% \text{ of free parent compound} = \frac{\text{amount of compound in receiver chamber}}{\text{amount of compound in donor chamber}} \times 100$$

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

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.

References

Dawson, M.I., Ye, M., Cao, X., Farhana, L., Hu, Q.Y., Zhao, Y., Xu, L.P., Kiselyuk, A., Correa, R.G., Yang, L., et al. (2009). Derivation of a retinoid X receptor scaffold from peroxisome proliferator-activated receptor gamma ligand 1-Di(1H-indol-3-yl)methyl-4-trifluoromethylbenzene. ChemMedChem *4*, 1106–1119.

Inoue, Y., Hayhurst, G.P., Inoue, J., Mori, M., and Gonzalez, F.J. (2002). Defective ureagenesis in mice carrying a liver-specific disruption of hepatocyte nuclear factor 4alpha (HNF4alpha): HNF4alpha regulates ornithine transcarbamylase in vivo. J. Biol. Chem. *277*, 25257–25265.

Kiselyuk, A., Farber-Katz, S., Cohen, T., Lee, S.H., Geron, I., Azimi, B., Heynen-Genel, S., Singer, O., Price, J., Mercola, M., et al. (2010). Phenothiazine neuroleptics signal to the human insulin promoter as revealed by a novel high-throughput screen. J. Biomol. Screen. *15*, 663–670.

Le Guével, R., Oger, F., Lecorgne, A., Dudasova, Z., Chevance, S., Bondon, A., Barath, P., Simonneaux, G., and Salbert, G. (2009). Identification of small molecule regulators of the nuclear receptor HNF4alpha based on naphthofuran scaffolds. Bioorg. Med. Chem. *17*, 7021–7030.

Petrescu, A.D., Hertz, R., Bar-Tana, J., Schroeder, F., and Kier, A.B. (2002). Ligand specificity and conformational dependence of the hepatic nuclear factor-4alpha (HNF-4alpha). J. Biol. Chem. *277*, 23988–23999.

Shoemaker, R.H. (2006). The NCI60 human tumour cell line anticancer drug screen. Nat. Rev. Cancer *6*, 813–823.

Xie, X., Lu, J., Kulbokas, E.J., Golub, T.R., Mootha, V., Lindblad-Toh, K., Lander, E.S., and Kellis, M. (2005). Systematic discovery of regulatory motifs in human promoters and 3′ UTRs by comparison of several mammals. Nature *434*, 338–345.

Yamasaki, N., Imito, T., Murai, Y., Hiramura, T., Oku, T., and Sawada, K. (1997). Preparation of benzimidazole derivatives as drugs, P.I. Appl, ed. (Tokyo, Japan: Fujisawa Pharmaceutical Co., Ltd.).