DISCOVERY, OPTIMIZATION, AND UNDERSTANDING THE MECHANISM OF

ACTION OF SMALL MOLECULES THAT RESTORE E-CADHERIN EXPRESSION

By

Sydney Lear Stoops

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

In partial fulfillment of the requirements

For the degree of

DOCTOR OF PHILOSOPHY

in

Pharmacology

May, 2012

Nashville, Tennessee

Approved:

Professor Craig Lindsley

Professor Brian Wadzinski

Professor Robert Coffey

Professor Stephen Fesik

Professor Lawrence Marnett

Professor Albert Reynolds

To My Family

Especially My Parents Steve and Lisa Glassman and Steve and Liz Stoops

ACKNOWLEDGEMENTS

There are many individuals that I would like to acknowledge for giving me the opportunity and support to pursue my Ph.D in Pharmacology. I would like to thank my undergraduate research mentor, Dr. Susan Hoffman, for allowing me to spend time in the lab. Had she not given me this opportunity, I may never have considered pursing a summer research internship and ultimately applying to graduate school. Additionally, I would also like to thank my professors at Miami University, especially Dr. John Stevenson, who taught me to love research and to think critically. I would like to thank Dr. Jack Hall for giving me the opportunity to be an R&D Intern at Viracor-IBT Laboratories the summer after my junior year of college as well as Dr. Michelle Altrich who was an excellent mentor. My time at Viracor-ICT Laboratories was short, but it was my first experience working in the lab on my own project. Thank you to Vanderbilt University, the IGP Program, and most importantly the Department of Pharmacology from which I am receiving my Ph.D.

I would like to thank my Ph.D. mentor, Craig Lindsley, for all he has taught me during my time at Vanderbilt University. Your knowledge on medicinal chemistry, drug discovery, and translational research is endless, and I am fortunate to have trained under such an intelligent, passionate, and hard-working individual. Craig, you have taught me more then I could have imagined during my Ph.D. in regards to medicinal chemistry and drug discovery. Thank you for giving me, someone with no chemistry experience, an opportunity to learn chemistry and work on a project that allowed me to do both medicinal chemistry and cancer biology research. In addition, I appreciate the

opportunities you have allowed me to pursue in regards to my career ambitions. Being able to intern at TriStar and attend the SIE program at Stanford have given me the confidence and ambition to pursue my career goals in the future.

As well, to all of the collaborators on my project – thank you. Dan Beauchamp you have been like a second mentor, and I appreciate you allowing me to be apart of your laboratory while working on my thesis project. You provided an abundance of guidance throughout my time at Vanderbilt University in regards to cancer biology research. Alex Waterson, you have been wonderful to work with on my thesis project, and I appreciate all of the chemistry work you have done as well as guiding me through my own chemistry library syntheses. Connie Weaver, you were truly a pleasure to work with on a daily basis! Thank you for all of the time spent working with me on my research project, troubleshooting assay protocols, and coming up with great ideas.

Thank you to my committee members. You are all the most intelligent, creative thinking, and driven individuals that I have come to know during my time at Vanderbilt University. I appreciate tremendously the time you have taken to be on my committee and guide me through the Ph.D. process; it has been priceless.

Finally, I would like to thank my family and friends. You have unconditionally supported me through the last 4.5 years while I have been working on my Ph.D. at Vanderbilt University. I thank you for all of the times you encouraged me to persevere when I thought I was crazy for wanting to get my Ph.D. You stood by my side through the roller coaster of failed experiments and glimmers of success – a feat, I imagine, was not easy at times. I look forward to moving on to the next chapter of my life and am comforted to know that I will always have y'all by my side.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter III

Figure

Chapter IV

Figure

Chapter V

Figure

LIST OF ABBREVIATIONS

°C	Degrees Celsius
^{13}C NMR	Carbon-13 Nuclear Magnetic Resonance
^{1}H NMR	Proton Nuclear Magnetic Resonance
5-FU	5-Flourouracil
AJ	Adherens Junction
APC	Adenomatous Polyposis Coli
bp	Base Pair
BrdU	Bromodeoxyuridine
BSA	Bovine Serum Albumin
CAF	Cancer-Associated Fibroblast
CCC	Cytoplasmic Cell-Adhesion Complex
ChIP	Chromatin Immunoprecipitation
CO_2	Carbon Dioxide
Da	Dalton
DCM	Dicholormethane
DIEA	N,N-Diisopropylethylamine
DMF	Dimethylformamide
DMPK	Drug Metabolism and Pharmacokinetics
DMSO	Dimethyl Sulfoxide
ECL	Enhanced Chemiluminescence
EDC	Ethylene dichloride

EGFR	Epithelial Growth Factor Receptor
ELSD	Evaporative Light Scattering Detector
EMT	Epithelial to Mesenchymal Transition
Epithelial-Cadherin	E-Cadherin
EPLIN	Epithelial Protein Lost in Neoplasm
FBS	Fetal Bovine Serum
GSK-3β	Glycogen Synthase Kinase 3β
H$_3$R	Histamine 3 Receptor
HAV	Histidine-Alanine-Valine Motif
HDAC	Histone Deactylase
HDACi	Histone Deactylase Inhibitor
HPLC	High Performance Liquid Chromatography
HRMS	High Resolution Mass Spectrometry
HTS	High-Throughput Screen
ICW	In Cell Western
Ig-CAM	Immunoglobulin-like Cell Adhesion Molecule
kbp	kilobase pairs
kDa	kiloDaltons
LCMS	Liquid Chromatography Mass Spectometry
MET	Mesenchymal to Epithelial Transition
MMP	Matrix Metalloprotease
NFATc1/c2	Nuclear Factor of Activated T-Cells (c1/c2)
NSCLC	Non-Small Cell Lung Cancer

PBS	Phosphate-Buffered Saline
PBS-T	Phosphate-Buffered Saline with Tween
PCR	Polymerase Chain Reaction
qPCR	Quantitative Polymerase Chain Reaction
SAR	Structure-Activity Relationship
TSA	Trichostatin A

LIST OF PUBLICATIONS

Sydney L. Stoops, Alex G. Waterson, R. Daniel Beauchamp, Craig W. Lindsley. 'Elucidating the Mechanism of Action of Small Molecules that Restore E-cadherin Expression'. 2012. *In Preparation.*

John T. Brogan, **Sydney L. Stoops**, Brenda C. Crews, Lawrence J. Marnett, Craig W. Lindsley. 'Total Synthesis (+)-7-Bromotrypargine and Unnatural Analogues: Biological Evaluation Uncovers Activity at CNS Targets of Therapeutic Relevance'. *ACS Chemical Neuroscience*. **November 16 2011**, 2(11) 633-639.

Margrith E. Mattmann, **Sydney L. Stoops**, and Craig W. Lindsley. 'Inhibition of Akt with Small Molecules and Biologics: Historical Perspective and Current Status of the Patent Landscape'. *Expert Opinion on Therapeutic Patents*. **June 2 2011**. *ASAP.*

Sydney L. Stoops, A. Scott Pearson, Connie Weaver, Alex G. Waterson, Emily Days, Chris Farmer, Suzanne Brady, C. David Weaver, R. Daniel Beauchamp, Craig W. Lindsley. 'Identification and Optimization of Small Molecules that Restore E-cadherin Expression and Reduce Invasion in Colorectal Carcinoma Cells'. *ACS Chemical Biology*. **May 20 2011**, 6(5): 452-65.

Leslie N. Aldrich, **Sydney L. Stoops,** Brenda S. Crews, Lawrence J. Marnett, Craig W. Lindsley. 'Total Synthesis and Biological Evaluation of Tambjamine K and a Library of Unnatural Analogs'. *Bioorganic & Medicinal Chemistry Letters*. **September 1 2010**, 20(17): 5207-5211

Jana A. Lewis, Sarah A. Scott, Robert Lavieri, Jason R. Buck, Paige E. Selvy, **Sydney L. Stoops**, Michelle D. Armstrong, H. Alex Brown, Craig W. Lindsley. 'Design and synthesis of isoform-selective phospholipase D (PLD) inhibitors. Part 1: Impact of alternative halogenated privileged structures for PLD1 specificity'. *Bioorganic & Medicinal Chemistry Letters*. **April 1 2009**, 19(4): 1916-1920.

CHAPTER I

INTRODUCTION

Current estimates from the American Cancer Society and International Union Against Cancer indicated that an estimated 12 million cases were diagnosed in 2008, with 7 million deaths occurring worldwide. Cancer is defined as a group of diseases characterized by the uncontrolled growth and spread of abnormal cells as a result of genetic alterations to the DNA. These alterations include deletions, inversions, amplifications, repressions, and chromosomal translocations, all of which can lead to genetic instability and tumor development [1, 2]. Furthermore, these alterations can be potentiated by both external risk factors (tobacco, chemicals, and infection) and internal risk factors (inherited mutations, random mutations, and age). Traditionally, no single cause, but multiple causal factors acting together or in sequence to initiate and promote carcinogenesis. Lack of diagnosis or proper treatment can result in spread of the cancer within the body and eventual death [3, 4].

The majority of human cancers, 80-90%, arise from epithelial cells, which are tightly held together through several junction structures: tight junctions, adherens-type junctions, and desmosomes. Development of malignant tumors, in particular the transition from benign to invasive, metastatic cancer, is often characterized by a tumor cell's ability to overcome cell-to-cell adhesion and to invade the surrounding tissue, lymph system, and the circulatory system. During the transition from a normal epithelial cell to a highly malignant (mesenchymal-like) cell, expression of some of these junction

molecules is drastically reduced or switched off [5]. This is often referred to as the epithelial to mesenchymal (EMT) transition, and is believed to play a prominent role in invasion, extravasion, and colonization during metastasis (**Figure 1**) [6].
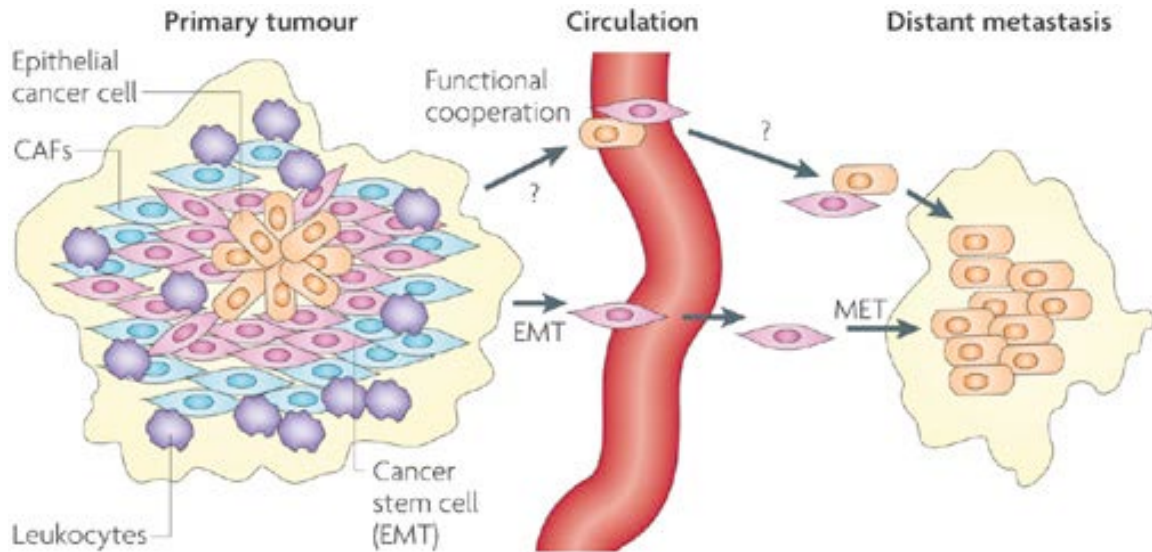
Figure 1. Primary tumor cells undergo epithelial to mesenchymal transition (EMT) and mesenchymal to epithelial transition (MET) which contributes to tumor heterogeneity and can influence the ability of select cells to metastasize. Interactions with surrounding stromal cells including leukocyes and cancer-associated fibroblasts (CAF), may induce EMT resulting in 'migratory cancer stem cells' which can be detected in circulation and micrometastases. However, distance metastasis may include differentiated epithelial cancer cells, a result of cells undergoing MET from local microenvironment selective pressures in order to survive (Polyak and Weinberg, 2009).

Through the advancement of cancer therapy in radiation, surgical resection procedures, and targeted chemotherapeutic development, it has been seen that patients do not die from complications from the primary tumor, but from the metastatic lesions. Common locations for metastatic tumors to arise are the lungs, brain, and bone, which are areas that are more difficult to treat or have more detrimental side effects from treatment

[7].   Through micromestastasis that can lay dormant and go undetected for years to decades, metastatic tumors can arise after the primary tumor has been confirmed to be completely removed from the patient.   There are a variety of theories that have been postulated within the cancer research community, one of popularity and growing support is the idea of cancer stem cells. It is believed that 'migratory cancer stem cells' may seed the metastatic microenvironments after undergoing EMT at the invasive front of the primary tumor and migrating into the surround tissue and circulatory system (**Figure 1**). Thus EMT plays a crucial role for the transformation of these cells to enable cellular detachment, dissemination, and finally metastasis [7-9].

As stated previously, many different cell-adhesion molecules are implicated in human carcinogenesis and the EMT process, and recently much attention has been directed towards Epithelial (E)-cadherin [10].   E-cadherin is a single-spanning transmembrane domain protein that forms homodimers at the cell surface membrane and interacts with homodimers of neighboring cells. Aside from cell-to-cell adhesion, E-cadherin is a key component in cell polarity induction and epithelium organization. The loss of E-cadherin function elicits active signals that support tumor-cell migration, invasion, and metastatic dissemination. This loss of expression during tumor progression can be caused by both genetic and epigenetic mechanisms; however, E-cadherin expression is most commonly downregulated at the transcriptional level [10].

Background and Significance

Colorectal cancer is the third most commonly diagnosed cancer in the United States and the third leading cause of cancer related deaths even though colorectal cancer incidence and mortality have been on the decline the past two decades. The decrease in incidence and mortality can both be attributed to improvement in early detection through colorectal cancer awareness and colonoscopy exams, thus allowing for the removal of polyps before they become malignant [11]. Even so, colorectal cancer still remains a major leading cause of cancer related deaths, which is related to the difficulty in treating colorectal cancer after it has spread from the primary tumor. In recent years, colorectal cancer has become a model for studying multistage carcinogenesis including the identification of sequential mutations, which lead to the development of colonrectal cancer (**Figure 2**) [12, 13]. This multistep process is observed in which the normal epithelial cells of the colon progress through a series of premalignant lesions to invasive and metastatic cancer with loss of E-cadherin being observed as a late stage step during this progression [14].
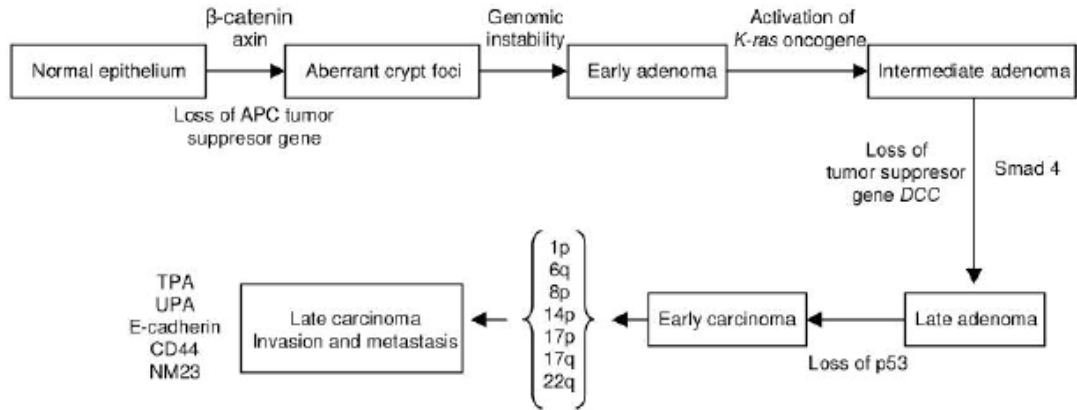
Figure 2. Genetic model of colorectal carcinogenesis.

E-cadherin was one of the first identified cadherins, and therefore has been thoroughly characterized in normal and pathological conditions, especially the role of E-cadherin in metastasis. Although E-cadherin expression can still be found in well-differentiated solid tumors in which cell adhesion is still intact, there is an inverse correlation between E-cadherin levels, tumor grade, and patient mortality rates. Loss of E-cadherin is commonly seen in poorly differentiated cells at the invasive front of the primary tumor, as well as in metastatic nodules [15-17]. One mode by which E-cadherin expression is lost is epigenetic silencing of *CDH1* gene, which encodes the E-cadherin protein. Genetic analysis of tumor samples has shown that the *CDH1* gene remains functional, suggesting that alterations in the regulatory system are one cause for loss of E-cadherin expression. For this reason, much emphasis has been put on determining which regulatory mechanisms are involved in silencing of *CDH1* transcription and if this process is reversible [5, 16, 18].

Structure, Function, and Localization of E-cadherin Protein

Cadherins are the principal component of Adherens Junctions (AJ) and desmosomes and the cluster at sites of cell-cell contact in most solid tissues. E-cadherin, named for its presence primarily in epithelial cells, is a Type I, or 'classical' cadherin, and is involved in calcium-dependent cell-cell adhesion [10, 19]. The mature E-cadherin protein is comprised of 728 amino acids that form a single transmembrane domain, a cytoplasmic domain of about 150 amino acids, and an ectodomain of about 550 amino acids. The ectodomain is comprised of 5 tandemly repeated domains separated by calcium binding motifs: 4 domains are 'extracellular cadherin repeats' and a 5th domain is characterized by four conserved cysteines [16, 20]. The cytoplasmic domain can be divided into two subdomains: the membrane proximal cytoplasmic conserved domain and the β-catenin binding domain. Each cytoplasmic subdomain contains a sequence motif of about 30-35 amino acid residues, which are conserved among all 'classic' cadherins [21]. The cytoplasmic β-catenin binding domain, as it suggests, is the location for β-catenin binding as well as formation of the cytoplasmic cell-adhesion complex (CCC). The cytoplasmic cell-adhesion complex is composed of β-catenin as well as α-catenin, ϒ-catenin, and p120-catenin (**Figure 3**) [10, 22-25].
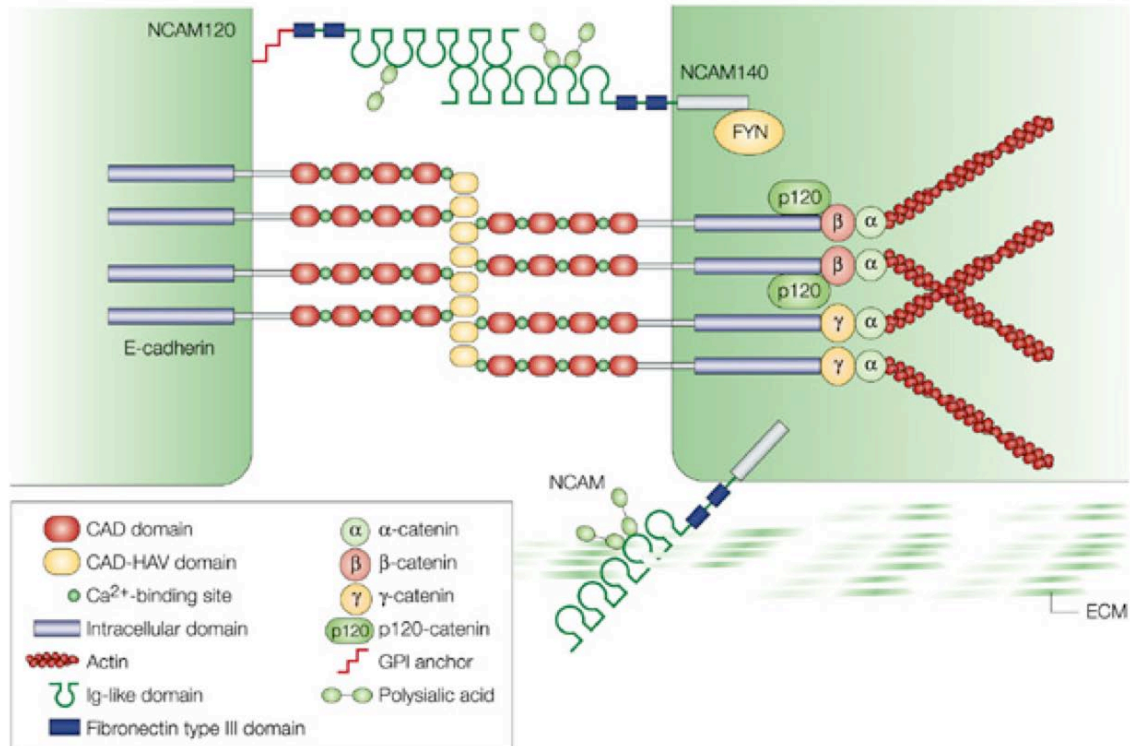
Figure 3. E-Cadherin molecules are expressed at the plasma membrane and interact with adjacent cells through an HAV motif in the amino-terminal cadherin domain. The cytoplasmic cell-adhesion complex (CCC) consists of β-catenin, α-catenin (on occasion ϒ-catenin) and p120-catenin. Through α-catenin, the CCC is linked to the actin cytoskeleton and plays a role in cell polarity and epithelial organization (Cavallaro and Christofori, 2004).

Both the extracellular domain and the cytoplasmic domain of the E-cadherin protein play important roles in the function of E-cadherin and cell-cell adhesion. The extracellular domain of E-cadherin is directly involved in cell-cell adhesion, which is accomplished by homophilic protein-protein interactions between two E-cadherin molecules at the surface of neighboring cells. This interaction is believed to be mediated between the most amino-terminal cadherin domains of E-cadherin proteins from adjacent cells. These distal domains contain a histidine-alanine-valine (HAV) motif located between tryptophan residues and hydrophobic pockets, which is responsible for the

interactions between E-cadherin molecules [16]. Based on structural analysis using electron tomography it is believed that the E-cadherin molecules interact through their tips (HAV motifs) in *cis* and in *trans* in a highly flexible manner. It was also shown that E-cadherin molecules localize in groups at the cell membrane suggesting that the molecular interactions occur in a zipper-like fashion (**Figure 3**) [26]. Through knockdown of E-cadherin or disruption of cadherin-mediated adhesion, it has been shown that both disturb or delay the functional assembly of other adhesion complexes including tight junctions and desmosomes [27, 28]. Although, it is worth noting that in some models these delays in adhesion complex formation can be overcome through alternative signaling pathways within the cell. In general, research has shown that E-cadherin in conjunction with other cellular junctions define the physiological function of the cell; or how the cell will be integrated in functional structures, such as organ epithelia or stroma [10].

As stated previously, the intracellular portion of the E-cadherin protein is the backbone for the cytoplasmic cell-adhesion complex, which is comprised of β-catenin, α-catenin, and p120-catenin, and mediates the association with the actin cytoskeleton [23, 24]. Recently, it was confirmed through use of biochemical analysis and quantitative microscopy that only α-catenin homodimers bind actin. Additionally, α-catenin could bind efficiently to the E-cadherin/β-catenin complex, but only in its monomeric form. Thus α-catenin alone cannot mediate the contact of the E-cadherin/β-catenin (CCC) complex to the F-actin cytoskeleton [29-32]. Therefore, it is believed that there is a linker protein present that engages the CCC to the F-actin cytoskeleton. To date, the strongest candidate for such a link between α-catenin in the CCC and the F-actin is

Epithelial Protein Lost in Neoplasm (EPLIN) protein. Research has shown that EPLIN localizes to the apical cortical F-actin cytoskeleton in epithelial cells, has at least two actin binding sites, and binds to the C-terminal domain of the monomeric α-catenin in the CCC. This interaction was confirmed when depletion of EPLIN from the cell resulted in disorganization of the actin belt, but did not affect non-junctional actin fiber formation [33-35].

In conclusion, without an intact CCC that is engaged with the actin cystoskeleton, cadherin-mediated strong cell-cell adhesion is compromised. Conversely, without cell-cell adhesion the CCC will not form, suggesting a co-dependent relationship between cell-cell adhesion and cellular cytoskeleton stability via the E-cadherin protein. In combination, both cell-cell adhesion and CCC interaction with the actin cytoskeleton aid in the development of epithelial cell polarization as well as regulating signaling towards the formation of other junctions [15].

Importance of E-cadherin Protein in Cell-Cell Adhesion

E-cadherin is involved in adherens cell-cell junctions between epithelial cells and plays a critical role in the establishment and maintenance of cell polarity and cell society. The cell-cell adherens junctions are specialized regions of the plasma membrane connected to the cytoskeletal actin filaments via the E-cadherin-mediated CCC and are believed to play a role in tissue morphogenesis [17]. These E-cadherin-mediated interactions in epithelial cells are important for establishing and maintaining polarity, preserving epithelial cell survival, and controlling proliferation.

When the cytoplasmic cell-adhesion complex is disrupted, intracellular alterations, such as release of β-catenin and γ-catenin into the cytoplasm occurs. There are two pathways for β-catenin when present in the cytoplasm. In normal epithelial cells, β-catenin and γ-catenin are rapidly phosphorylated by APC-GSK-3β complex and subsequently targeted for ubiquitinated proteasomal degradation [36, 37]. The second pathway for β-catenin arises during tumor progression in which mutations can occur in the tumor suppressor APC rendering it non-functional, as commonly seen in colon cancer (**Figure 1**), or GSK-3β can be blocked by activated Wnt signaling leading to accumulation of high levels of β-catenin in the cytoplasm [10]. Activated Wnt signaling can arise from both a positive feedback loop as well as release of Wnt ligand for surrounding cells thus allowing free β-catenin in the cytoplasm to localize to the nucleus. As a result, β-catenin activates TCF/LEF1 mediated transcription, resulting in upregulation of repressor transcription factors such as Twist1, Snail, and Slug [38]. Additionally, the TCF/LEF1 family of transcription factors modulate the expression of c-MYC, cyclin D1, fibronectin, MMP7, ID2, CD44, conductin, TCFI, and other genes all of which are implicated in cell proliferation and tumor progression (**Figure 4**) [17].
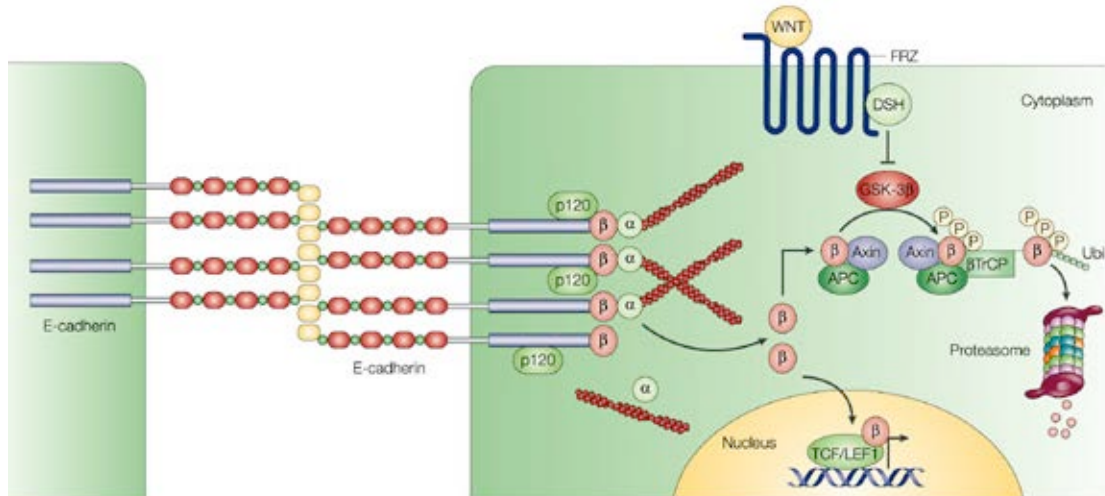
Figure 4. Loss of E-cadherin alters Wnt signaling. After loss of E-cadherin function and disassembly of the cytoplasmic cell-adhesion complex, catenins are released and accumulate in the cytoplasm. β-Catenin is then sequestered by the adenomatous polyposis coli (APC)–axin–glycogen synthase kinase 3β (GSK-3β) complex and phosphorylated by GSK-3β. Phosphorylated β-catenin is ubiquitinated for rapid proteosomal degradation. However, on activation of the Wnt signaling pathway, which is commonly seen in cancer through increased receptor activity or activating mutations within the signaling pathway, GSK-3β is repressed and β-catenin is no longer phosphorylated. It translocates to the nucleus where, together with the TCF/LEF1 transcription factors, it modulates the expression of several target genes that are known to be involved in cell proliferation and tumour progression. (Cavallaro and Christofori, 2004)

Therefore, in normal tissue, in which E-cadherin is expressed, cell-cell adhesion remains intact creating epithelial cell sheets. Additionally, expression of E-cadherin, and the presence of the cytoplasmic cell-adhesion complex, aids in the maintenance of cell polarity and cell structure. However, in cancer cells where E-cadherin expression is suppressed, cells may overcome adhesions to adjacent cells and dissociate from the primary tumor, leading to invasion of neighboring tissues and metastasis [39, 40].

Epithelial to Mesenchymal Transition in Tumor Invasion and Metastasis

Epithelial to mesenchymal transition (EMT) is a process by which epithelial cells lose many of their epithelial characteristics and take on properties that are typical of mesenchymal cells. This transition requires complex changes in cell architecture and behavior. Epithelial cells are characterized by forming layers of cells that are closely adjoined by specialized membrane structures such as tight junctions, adherens junctions, desmosomes, and gap junctions. Additionally, epithelial cells have apical-basolateral polarization, which manifests itself through localized distribution of the adhesion molecules, the polarized organization of the actin cytoskeleton, and the presence of basal laminal at the basal surface. Under normal conditions, the epithelial cells may become motile, detach, and move away from their nearest neighbors while remaining within the epithelial layer; however, they do not detach and move away from the epithelial layer [41, 42]. Conversely, mesenchymal cells do not form an organized cell layer, nor do they have the same apical-basolateral organization and polarization of the cell-surface molecules and the actin cytoskeleton as epithelial cells. Mesenchymal cells make contact focally with other cells, are not typically associated with basal lamina, and migrate either in chains or as individual cells. However, without mesenchymal cells during morphogenesis, tissue and organs would never be formed. Thus, epithelial to mesenchymal transition is an indispensible mechanism during development [41, 43, 44].

Consequently, EMT is also thought to play a central role in tumor progression. During progression to metastatic competence, tumor cells acquire mesenchymal gene expression patterns and properties, resulting in changed adhesive properties, activation of

proteolysis, and increased motility [41, 45]. Nascent mesenchymal tumor cells detach from tightly organized epithelial sheets by breaking down their tight associations or adherens junctions with adjacent epithelial cells. Mesenchymal tumor cells no longer form sheets, display poorly organized adhesive junctions, and exhibit more fibroblastic characteristics (**Figure 5**). Activation of proteolysis, as a result of production of extracellular matrix degrading enzymes, facilitates invasion of the basement membrane by the mesenchymal-like tumor cells, leading to entrance into the circulatory or lymph systems and recolonization at distant organs [46, 47].
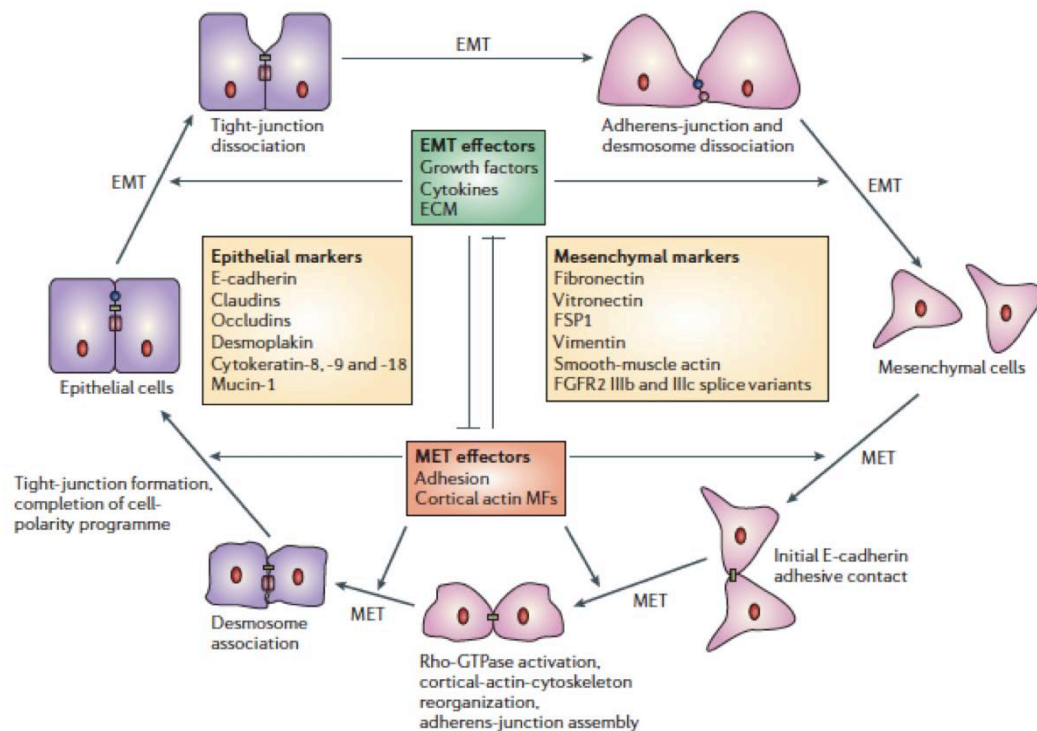
Figure 5. Regulation of the tight junctions and adherens junctions between adjacent cells affects the progression of EMT and MET. The different stages during epithelial-mesenchymal transition (EMT) and the reverse process mesenchymal-epithelial transition (MET) are regulated by effectors of EMT and MET, which influence each other. (Thiery and Sleeman, 2006).

Much of the research conducted to confirm EMT involvement in epithelial tumor metastasis has been done *in vitro*, where epithelial carcinoma cells are induced to undergo EMT by growth factors, chemokines, alterations to transcriptional regulation, or deletion of critical adherens molecules. Unfortunately, EMT involvement in epithelial tumors *in vivo* has been poorly documented. This is most likely because EMT is thought to be a dynamic or reversible process *in vivo*, making it difficult to pinpoint in live tissues. Specifically, migrating carcinoma cells are endowed with great cellular plasticity, allowing them to change shape and to revert back and forth from an epithelial to mesenchymal morphology in response to regulatory signals form the surrounding environment [46, 48]. Development of novel small molecules to target the EMT process, either through reversion or blocking EMT, would be an excellent tool to further elucidate the role of EMT during tumor progression within *in vivo* models.

Loss of E-cadherin in Tumorigenesis

The loss of E-cadherin function during tumor progression can be caused by various genetic or epigenetic mechanisms. One mode in which E-cadherin function can be lost is mutation in the *CDH1* gene, leading to expression of a non-functional protein. Mutations in the *CDH1* gene have been seen in diffuse gastric cancer, lobular breast cancer, and a lower incidence in thyroid, bladder, and gynecological cancers [49, 50]. Somatic gene mutations are accompanied by loss of heterozygosity of the remaining E-cadherin allele, aiding in tumor progression, while germline mutations may act early in the natural history of the disease [50-52]. *CDH1* gene mutations include exon skipping,

frame shift deletions, and insertions. Subsequent mutations can also arise in β-catenin and α-catenin (a very rare event) leading to disruption of the cytosplasmic cell-adhesion complex, preventing interaction with the actin cytoskeleton and eventual loss of cell-cell adhesion [53-55].

Another mechanism by which E-cadherin mediated cell-cell adhesion can be ablated is proteolytic degradation of E-cadherin by matrix metalloproteases (MMP). In such a case, E-cadherin expression levels would remain normal within the cell, but extracellular alterations would prevent functional cell-cell adhesion establishment, aiding in tumor progression. A soluble 80 kDa form of E-cadherin, produced by degradation of the full-length protein, is frequently found in cultured tumor cell lines and in tumor biopsy samples. It has been shown that the soluble form of E-cadherin, the cleaved 80 kDa ectodomain, promotes tumor progression, and specifically tumor-cell invasion, by further upregulating MMPs [10, 55].

The most common mechanism by which E-cadherin expression is downregulated is at the transcriptional level. Repressor transcription factors Snail, Slug, and SIP1 as well as the helix-loop-helix transcription factor E12/E47 have been found to bind to the E2 boxes in the promoter of the E-cadherin gene and actively repress transcription. As a direct consequence of transcriptional inactivation, the E-cadherin locus is epigenetically silenced by hypermethylation and deacetylation as seen in Figure 6 [56-59]. DNase I hypersensitive site mapping indicated the loss of transcription factor binding, resulting in chromatin rearrangement in the regulatory region of the E-cadherin gene. It was shown through cloning and sequencing of the E-cadherin gene promoter that CpG methylation around the promoter region of the E-cadherin gene was present in cell lines that lacked E-

15

cadherin expression. Additionally, E-cadherin expression could be restored in these cell lines upon treatment with the DNA methyltransferase inhibitor 5-azacytidine [60, 61].
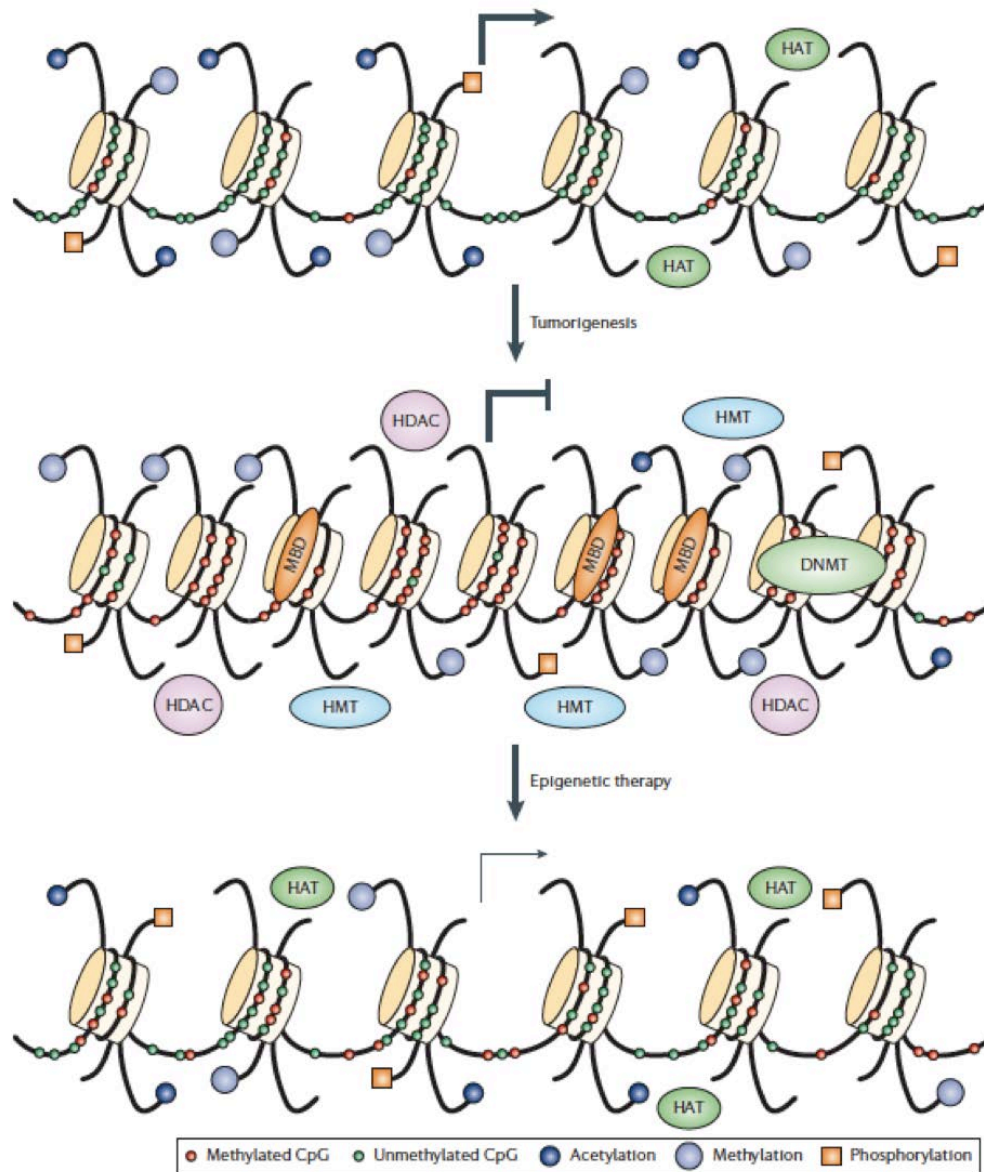
Figure 6. Epigenetic silencing during tumorigenesis. A transcriptionally active gene is marked by DNA hypomethylation of the promoter region of the gene and acetylation of the lysine residues on histone tails, both of which promote the formation of euchromatin. During tumorigenesis, DNA methylation accumulates in the promoter region, which attracts methyl-binding repressive proteins as well as promotes deacetylation of histone lysine residues resulting in compacted chromatin and transcriptional repression. (Yoo and Jones, 2006)

Additionally, deacetylation of lysine residues of histones by histone deacteylase (HDAC) enzymes results in chromatin compaction and inactivation of genes (**Figure 6**). Deactylation has also been shown to occur around the E-cadherin gene promoter region by a repressor complex comprised of Snail/HDAC1/HDAC2. It has been shown that Snail preferentially binds the E2 box in the promoter region, while directly binding to HDAC2, and indirectly to HDAC1, in a complex. Concurrently, treatment of cell lines with reduced E-cadherin expression with Trichostatin A (TSA), a Class I and Class II histone deacetylase inhibitor, leads to restored expression of E-cadherin in these cell lines [62, 63].

Targeting E-cadherin Expression Through Small Molecule Modulation

E-cadherin restoration has been observed with the treatment of histone deactylase inhibitors (HDACi) within a variety of cancer cell lines including colon cancer, lung cancer, and breast cancer [64, 65]. HDACi are known to promote hyperacetylation, whereby they reactivate suppressed genes, such as E-cadherin, leading to activation of cellular differentiation programs, inhibition of the cell cycle, and induction of apoptosis [66].

Further, research conducted preliminarily in non-small cell lung cancer (NSCLC) has shown that EMT is associated with erlotinib sensitivity. Erlotinib is a reversible tyrosine kinase inhibitor to the epithelial growth factor receptor (EGFR) [67-69]. It was observed that lung cancer cell lines expressing epithelial proteins, such as E-cadherin and α- and γ-catenins, were sensitive to growth inhibition by erlotinib, whereas those cancer

cell lines that had undergone an EMT-like transition from an epithelial phenotype (expression of E-cadherin) to a mesenchymal phenotype (expression of fibronectin, vimentin, and the transcription factor Zeb1) were less sensitive to EGFR inhibition [68]. Similarly, a retrospective analysis of a phase III trial in NSCLC for erlotinib plus chemotherapy compared with chemotherapy alone showed that E-cadherin expression was a significant predictive marker for efficacy of EGFR inhibition by erlotinib, as measured by progression-free survival [69].

Haley and coworkers found a similar relationship between sensitivity to EGFR inhibition and the formation of adhesion junctions in colorectal and pancreatic cancer cell lines. They show that epithelial colorectal and pancreatic cancer cell lines were more sensitive than those that had undergone EMT and had a mesenchymal phenotype. Collectively, the data presented showed that colorectal cancer cell lines expressing E-cadherin and β-catenin and morphologically showing an epithelial phenotype were sensitive to erlotinib, whereas those that had lost an epithelial phenotype, either through mutation or by EMT, were less sensitive as measured in maximal growth inhibition after a 10 μM treatment of erlotinib for 72 hours [67].

Combining this knowledge, HDACi restoring E-cadherin expression and E-cadherin expression being a marker for EGFR inhibitor sensitivity, it seemed logical for researchers to investigate the synergistic effects of HDACi and EGFR inhibitors given in combination. This was done by Paul Bunn, Jr. and coworkers, wherein they first transfected E-cadherin into gefitinib, an EGFR inhibitor, resistant cell lines and observed restoration of gefitinib sensitivity. They followed up this data, by pre-treating the cells with an HDACi, which induced E-cadherin along with EGFR and led to growth-

inhibitory and apoptotic effects by gefitinib similar to those seen in gefitinib-sensitive NSCLC cell lines [70].

Therefore, the data suggests that E-cadherin expression; specifically restoration of E-cadherin expression, may be a potential therapeutic target both for the prevention and reversion of EMT as well as resensitization of cancer cells to alternative therapies already in the clinic (e.g. EGFR inhibitors). Further, several groups have demonstrated that re-establishing the functional E-cadherin complex, such as through forced expression of E-cadherin, resulted in a reversion from an invasive, mesenchymal phenotype to a benign, epithelial phenotype of cultured cancer cells [16, 40]. Thus, it seems appropriate to hypothesize that developing small molecules to restore E-cadherin expression would be therapeutically relevant for the treatment of cancer.

Notes on Figures, Tables, and Compound Numbering

Throughout the remaining Chapters, all figure and table numbering is specific to each Chapter. Put differently, each Chapter's figure and table numbering begins with "1". Similarly, compound numbering formats and nomenclature is specific for Chapters II-IV and Chapter V, separately. Arbitrary numbering (e.g. Compound **1** or (**1**)) is used interchangeably throughout each Chapter. These are used in place of the VU registration codes (e.g. VU0075630) or the R/Y Lab Code (e.g. R21) in order to increase the ease and readability and to reduce space. Table 1 in the Appendix 2 contains the Thesis numbering, VU registration code, and Y/R lab code listed for each compound (by Thesis numbering) to allow for ease in requesting compounds from universal storage or referencing laboratory notebooks.

References

1. Aggarwal, B.B., et al., *Models for prevention and treatment of cancer: problems vs promises.* Biochem Pharmacol, 2009. **78**(9): p. 1083-94.

2. Nambiar, M., V. Kari, and S.C. Raghavan, *Chromosomal translocations in cancer.* Biochim Biophys Acta, 2008. **1786**(2): p. 139-52.

3. Garcia, M., et al., *Global Cancer Facts & Figures*, A.C. Society, Editor. 2007: Atlanta, GA.

4. *Cancer Facts & Figures 2009*, A.C. Society, Editor. 2009: Atlanta, GA.

5. Christofori, G. and H. Semb, *The role of the cell-adhesion molecule E-cadherin as a tumour-suppressor gene.* Trends Biochem Sci, 1999. **24**(2): p. 73-6.

6. Steeg, P.S., *Metastasis suppressors alter the signal transduction of cancer cells.* Nat Rev Cancer, 2003. **3**(1): p. 55-63.

7. Brabletz, T., et al., *Opinion: migrating cancer stem cells - an integrated concept of malignant tumour progression.* Nat Rev Cancer, 2005. **5**(9): p. 744-9.

8. Savagner, P., *Leaving the neighborhood: molecular mechanisms involved during epithelial-mesenchymal transition.* Bioessays, 2001. **23**(10): p. 912-23.

9. Thiery, J.P., *Epithelial-mesenchymal transitions in tumour progression.* Nat Rev Cancer, 2002. **2**(6): p. 442-54.

10. Cavallaro, U. and G. Christofori, *Cell adhesion and signalling by cadherins and Ig-CAMs in cancer.* Nat Rev Cancer, 2004. **4**(2): p. 118-32.

11. *Cancer Facts & Figures 2011*, A.C. Society, Editor. 2011: Atlanta, GA.

12. Tsanou, E., et al., *The E-cadherin adhesion molecule and colorectal cancer. A global literature approach.* Anticancer Res, 2008. **28**(6A): p. 3815-26.

13. Fearon, E.R. and B. Vogelstein, *A genetic model for colorectal tumorigenesis.* Cell, 1990. **61**(5): p. 759-67.

14. Fearnhead, N.S., J.L. Wilding, and W.F. Bodmer, *Genetics of colorectal cancer: hereditary aspects and overview of colorectal tumorigenesis.* Br Med Bull, 2002. **64**: p. 27-43.

15. van Roy, F. and G. Berx, *The cell-cell adhesion molecule E-cadherin.* Cell Mol Life Sci, 2008. **65**(23): p. 3756-88.

16.	Birchmeier, W. and J. Behrens, *Cadherin expression in carcinomas: role in the formation of cell junctions and the prevention of invasiveness.* Biochim Biophys Acta, 1994. **1198**(1): p. 11-26.

17.	Hirohashi, S., *Inactivation of the E-cadherin-mediated cell adhesion system in human cancers.* Am J Pathol, 1998. **153**(2): p. 333-9.

18.	Bracke, M.E., F.M. Van Roy, and M.M. Mareel, *The E-cadherin/catenin complex in invasion and metastasis.* Curr Top Microbiol Immunol, 1996. **213 (Pt 1**): p. 123-61.

19.	Fujita, Y., et al., *Hakai, a c-Cbl-like protein, ubiquitinates and induces endocytosis of the E-cadherin complex.* Nat Cell Biol, 2002. **4**(3): p. 222-31.

20.	Nollet, F., P. Kools, and F. van Roy, *Phylogenetic analysis of the cadherin superfamily allows identification of six major subfamilies besides several solitary members.* J Mol Biol, 2000. **299**(3): p. 551-72.

21.	Rimm, D.L. and J.S. Morrow, *Molecular cloning of human E-cadherin suggests a novel subdivision of the cadherin superfamily.* Biochem Biophys Res Commun, 1994. **200**(3): p. 1754-61.

22.	Giehl, K. and A. Menke, *Microenvironmental regulation of E-cadherin-mediated adherens junctions.* Front Biosci, 2008. **13**: p. 3975-85.

23.	Kemler, R., *From cadherins to catenins: cytoplasmic protein interactions and regulation of cell adhesion.* Trends Genet, 1993. **9**(9): p. 317-21.

24.	Gumbiner, B.M., *Cell adhesion: the molecular basis of tissue architecture and morphogenesis.* Cell, 1996. **84**(3): p. 345-57.

25.	Hartsock, A. and W.J. Nelson, *Adherens and tight junctions: structure, function and connections to the actin cytoskeleton.* Biochim Biophys Acta, 2008. **1778**(3): p. 660-9.

26.	He, W., P. Cowin, and D.L. Stokes, *Untangling desmosomal knots with electron tomography.* Science, 2003. **302**(5642): p. 109-13.

27.	Tunggal, J.A., et al., *E-cadherin is essential for in vivo epidermal barrier function by regulating tight junctions.* EMBO J, 2005. **24**(6): p. 1146-56.

28.	Lewis, J.E., et al., *Cross-talk between adherens junctions and desmosomes depends on plakoglobin.* J Cell Biol, 1997. **136**(4): p. 919-34.

29.	Drees, F., et al., *Alpha-catenin is a molecular switch that binds E-cadherin-beta-catenin and regulates actin-filament assembly.* Cell, 2005. **123**(5): p. 903-15.

30.   Yamada, S., et al., *Deconstructing the cadherin-catenin-actin complex.* Cell, 2005. **123**(5): p. 889-901.

31.   Weis, W.I. and W.J. Nelson, *Re-solving the cadherin-catenin-actin conundrum.* J Biol Chem, 2006. **281**(47): p. 35593-7.

32.   Rimm, D.L., et al., *Alpha 1(E)-catenin is an actin-binding and -bundling protein mediating the attachment of F-actin to the membrane adhesion complex.* Proc Natl Acad Sci U S A, 1995. **92**(19): p. 8813-7.

33.   Abe, K. and M. Takeichi, *EPLIN mediates linkage of the cadherin catenin complex to F-actin and stabilizes the circumferential actin belt.* Proc Natl Acad Sci U S A, 2008. **105**(1): p. 13-9.

34.   Maul, R.S., et al., *EPLIN regulates actin dynamics by cross-linking and stabilizing filaments.* J Cell Biol, 2003. **160**(3): p. 399-407.

35.   Gates, J. and M. Peifer, *Can 1000 reviews be wrong? Actin, alpha-Catenin, and adherens junctions.* Cell, 2005. **123**(5): p. 769-72.

36.   Bienz, M. and H. Clevers, *Linking colorectal cancer to Wnt signaling.* Cell, 2000. **103**(2): p. 311-20.

37.   Polakis, P., *Wnt signaling and cancer.* Genes Dev, 2000. **14**(15): p. 1837-51.

38.   Yang, J. and R.A. Weinberg, *Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis.* Dev Cell, 2008. **14**(6): p. 818-29.

39.   Behrens, J., et al., *Dissecting tumor cell invasion: epithelial cells acquire invasive properties after the loss of uvomorulin-mediated cell-cell adhesion.* J Cell Biol, 1989. **108**(6): p. 2435-47.

40.   Vleminckx, K., et al., *Genetic manipulation of E-cadherin expression by epithelial tumor cells reveals an invasion suppressor role.* Cell, 1991. **66**(1): p. 107-19.

41.   Thiery, J.P. and J.P. Sleeman, *Complex networks orchestrate epithelial-mesenchymal transitions.* Nat Rev Mol Cell Biol, 2006. **7**(2): p. 131-42.

42.   Schock, F. and N. Perrimon, *Molecular mechanisms of epithelial morphogenesis.* Annu Rev Cell Dev Biol, 2002. **18**: p. 463-93.

43.   Thompson, E.W., D.F. Newgreen, and D. Tarin, *Carcinoma invasion and metastasis: a role for epithelial-mesenchymal transition?* Cancer Res, 2005. **65**(14): p. 5991-5; discussion 5995.

44. Friedl, P., *Prespecification and plasticity: shifting mechanisms of cell migration.* Curr Opin Cell Biol, 2004. **16**(1): p. 14-23.

45. Sleeman, J.P., *The lymph node as a bridgehead in the metastatic dissemination of tumors.* Recent Results Cancer Res, 2000. **157**: p. 55-81.

46. Agiostratidou, G., et al., *Differential cadherin expression: potential markers for epithelial to mesenchymal transformation during tumor progression.* J Mammary Gland Biol Neoplasia, 2007. **12**(2-3): p. 127-33.

47. Hay, E.D., *The mesenchymal cell, its role in the embryo, and the remarkable signaling mechanisms that create it.* Dev Dyn, 2005. **233**(3): p. 706-20.

48. Wang, W., et al., *Identification and testing of a gene expression signature of invasive carcinoma cells within primary mammary tumors.* Cancer Res, 2004. **64**(23): p. 8585-94.

49. Strathdee, G., *Epigenetic versus genetic alterations in the inactivation of E-cadherin.* Semin Cancer Biol, 2002. **12**(5): p. 373-9.

50. Guilford, P., et al., *E-cadherin germline mutations in familial gastric cancer.* Nature, 1998. **392**(6674): p. 402-5.

51. Jeanes, A., C.J. Gottardi, and A.S. Yap, *Cadherins and cancer: how does cadherin dysfunction promote tumor progression?* Oncogene, 2008. **27**(55): p. 6920-9.

52. Berx, G., et al., *E-cadherin is inactivated in a majority of invasive human lobular breast cancers by truncation mutations throughout its extracellular domain.* Oncogene, 1996. **13**(9): p. 1919-25.

53. Oda, T., et al., *E-cadherin gene mutations in human gastric carcinoma cell lines.* Proc Natl Acad Sci U S A, 1994. **91**(5): p. 1858-62.

54. Becker, K.F., et al., *E-cadherin gene mutations provide clues to diffuse type gastric carcinomas.* Cancer Res, 1994. **54**(14): p. 3845-52.

55. Nawrocki-Raby, B., et al., *Upregulation of MMPs by soluble E-cadherin in human lung tumor cells.* Int J Cancer, 2003. **105**(6): p. 790-5.

56. Batlle, E., et al., *The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells.* Nat Cell Biol, 2000. **2**(2): p. 84-9.

57. Cano, A., et al., *The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression.* Nat Cell Biol, 2000. **2**(2): p. 76-83.

58.   Comijn, J., et al., *The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion.* Mol Cell, 2001. **7**(6): p. 1267-78.

59.   Hajra, K.M., D.Y. Chen, and E.R. Fearon, *The SLUG zinc-finger protein represses E-cadherin in breast cancer.* Cancer Res, 2002. **62**(6): p. 1613-8.

60.   Hennig, G., et al., *Progression of carcinoma cells is associated with alterations in chromatin structure and factor binding at the E-cadherin promoter in vivo.* Oncogene, 1995. **11**(3): p. 475-84.

61.   Yoshiura, K., et al., *Silencing of the E-cadherin invasion-suppressor gene by CpG methylation in human carcinomas.* Proc Natl Acad Sci U S A, 1995. **92**(16): p. 7416-9.

62.   Yoo, C.B. and P.A. Jones, *Epigenetic therapy of cancer: past, present and future.* Nat Rev Drug Discov, 2006. **5**(1): p. 37-50.

63.   von Burstin, J., et al., *E-cadherin regulates metastasis of pancreatic cancer in vivo and is suppressed by a SNAIL/HDAC1/HDAC2 repressor complex.* Gastroenterology, 2009. **137**(1): p. 361-71, 371 e1-5.

64.   Krishnan, M., et al., *HDAC inhibitors regulate claudin-1 expression in colon cancer cells through modulation of mRNA stability.* Oncogene. **29**(2): p. 305-12.

65.   Wu, Y., A. Starzinski-Powitz, and S.W. Guo, *Trichostatin A, a histone deacetylase inhibitor, attenuates invasiveness and reactivates E-cadherin expression in immortalized endometriotic cells.* Reprod Sci, 2007. **14**(4): p. 374-82.

66.   Johnstone, R.W., *Histone-deacetylase inhibitors: novel drugs for the treatment of cancer.* Nat Rev Drug Discov, 2002. **1**(4): p. 287-99.

67.   Buck, E., et al., *Loss of homotypic cell adhesion by epithelial-mesenchymal transition or mutation limits sensitivity to epidermal growth factor receptor inhibition.* Mol Cancer Ther, 2007. **6**(2): p. 532-41.

68.   Thomson, S., et al., *Epithelial to mesenchymal transition is a determinant of sensitivity of non-small-cell lung carcinoma cell lines and xenografts to epidermal growth factor receptor inhibition.* Cancer Res, 2005. **65**(20): p. 9455-62.

69.   Yauch, R.L., et al., *Epithelial versus mesenchymal phenotype determines in vitro sensitivity and predicts clinical activity of erlotinib in lung cancer patients.* Clin Cancer Res, 2005. **11**(24 Pt 1): p. 8686-98.

70.   Witta, S.E., et al., *Restoring E-cadherin expression increases sensitivity to epidermal growth factor receptor inhibitors in lung cancer cell lines.* Cancer Res, 2006. **66**(2): p. 944-50.

CHAPTER II

IDENTIFICATION AND OPTIMIZATION OF SMALL MOLECUES THAT

RESTORE E-CADHERIN EXPRESSION

Introduction

As stated previously, most human cancers arise from epithelial cells in which E-cadherin mediated cell-cell adhesion is lost concomitantly with progression towards tumor malignancy. Although E-cadherin expression can still be found in well differentiated tumors, there is an inverse correlation between E-cadherin expression levels and tumor grade [1-3]. Research has shown that inhibition of E-cadherin expression aids in the epithelial to mesenchymal transition in both *in vivo* and *in vitro* models as well as increased metastatic capabilities in *in vivo* models. Currently, research looking at restoration of E-cadherin expression *in vitro* involves the use of small molecules such as HDAC inhibitors and DNA methyltransferase inhibitors. HDAC and DNA methyltransferase inhibitors are known to restore expression of E-cadherin in cell lines that have a repressed functional gene [4-7]. However, upregation of E-cadherin expression is merely a side effect of HDAC or DNA methyltransferase inhibition by these small molecules, but not the targeted phenotypic response [7]. Therefore, to further

understand the role of E-cadherin in the metastatic process, it is necessary to develop novel small molecules that are optimized to restore E-cadherin expression.

A high-throughput screening assay can be developed to find small molecules that directly bind the target of interest (target-based approach) or can be developed to find small molecules that elicit a desired phenotypic response (phenotypic approach). While target-based drug discovery has become most popular in the large biopharmaceutical companies, it only accounts for 23% of first-in-class drugs between 1999 – 2008 as compared to 37% derived using a phenotypic approach [8]. In this case, the goal was to identify small molecules that restored E-cadherin expression. Therefore, a preliminary high-throughput screening assay was developed to discover small molecules that elicited a desired phenotypic response: restoration of E-cadherin expression. The main disadvantage with a phenotypic readout is that the target(s) for the small molecules identified is unknown. This presents a challenge for optimizing the molecular properties of the candidate small molecules without design parameters provided by a priori knowledge of the molecular mechanism of action [8]. Additionally, if several small molecules are indentified; potentially there could be several targets and signaling pathways by which these small molecules are restoring E-cadherin expression. Conversely, a major advantage to using a phenotypic screen is that it does not require prior understanding of the molecular mechanism of action, thus an assay developed for a phenotypic screen is more effective at translating a given disease state then target-based approaches, which are often more artificial [8].

To date, there have been no high throughput screens or drug discovery programs focused directly on developing small molecules that specifically upregulate E-cadherin

expression as the primary response. For that reason, the high-throughput screening assay developed and utilized in our preliminary research, confirms a novel mechanism by which four screening hits were discovered to elicit restored E-cadherin function.

Additionally, an optimized approach to synthesize small molecule analogs and screen activity was used to drive aggressive preliminary structure-activity-relationships (SAR) surrounding the lead hit molecule. Basic medicinal chemistry was conducted in an iterative fashion using parallel libraries. The reactions employed straightforward chemistry approaches in order to maximize synthetic efficiency and library size. Each analog was isolated by mass-directed HPLC (Agilent Technologies) to >98% purity, diluted in a barcoded vial to 10 mM DMSO, and registered with a unique VU number that serves as an identifier for compound management. Analogs in each library iteration were first screened at a single point concentration (10 μM) in two assays using the Odyssey Infrared Imaging System: standard Western blot assay and in-cell western assay. DMSO was used as a negative control and TSA, a histone deacetylase inhibitor, was used as a positive control. The screening data was then used to drive synthesis of future library iterations while a subset of selected analogs were used for further biological evaluation (**Figure 1**).
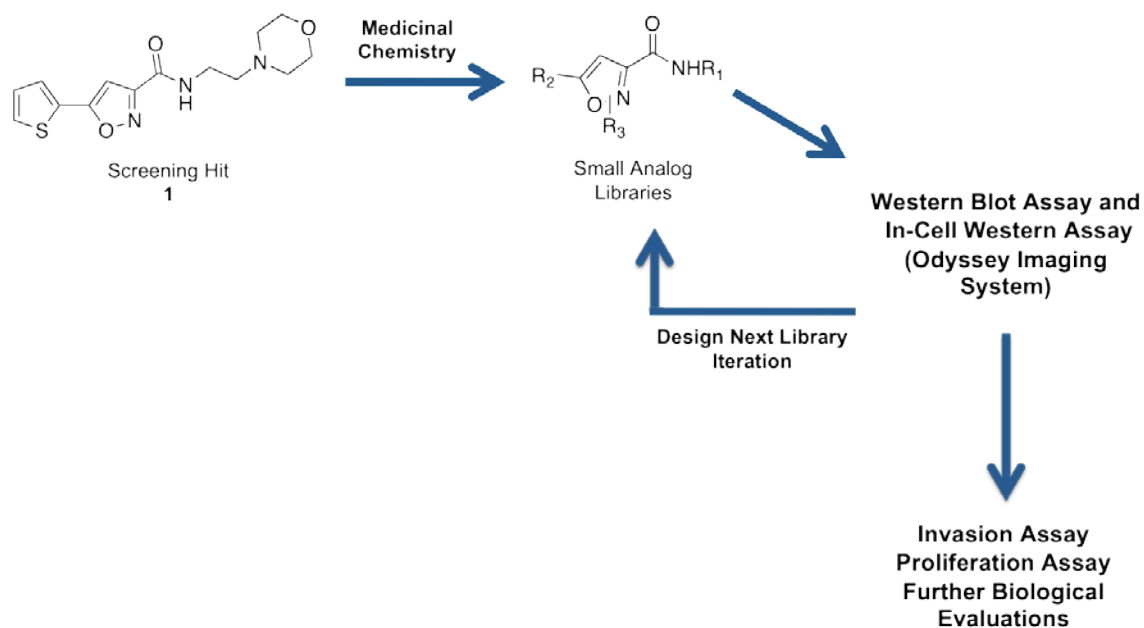
Figure 1. Optimized approach to efficiently synthesize and screen small analog libraries. The data collected from each screen was used to develop the next library iteration, resulting in a relatively efficient method to probe preliminary SAR around the screening hit, compound **1**.

Thus, we developed and used a high-throughput screen to discover lead compounds that restored E-cadherin expression in a metastatic colon adenocarcinoma cell line, SW620, which exhibited low levels of E-cadherin expression. We were able to develop preliminary SAR around the lead compounds using a Western blot-based E-cadherin restoration assay (**Figure 1**). In addition, we were able to develop an in-cell western assay with the Odyssey Imaging System to quantify EC_{50} values for select compounds.

Methods and Materials

Cell Culture

A colorectal adenocarcinoma cell line, SW620, and a squamous cell lung carcinoma cell line, H520, were obtained from the American Type Culture Collection (Manassas, VA) and maintained in a humidified atmosphere of 5% CO_2 in air at 37 °C. The cells were routinely cultured in RPMI 1640 supplemented medium with 10% fetal bovine serum (FBS) (Atlanta Biologicals, GA) and 100 µg/mL penicillin-streptomycin.

High-Throughput Screen

The metastatic human colorectal adenocarcinoma cell line SW620 was selected for this cell-based assay based on its low level of E-cadherin expression [9]. The high-throughput screening assay was accomplished by modifying the E-cadherin immunoassay from a microscope-based protocol to an automated, HTS-compatible assay in 384 well plates. Variables such as cell concentration and media content, plate type, antibody concentrations, incubation times, and fluorescent indicators, were systematically tested to achieve the best overall signal-to-noise ratio and uniformity across the wells of the assay plates. The optimal conditions, which best reproduced previous microscopy data showing Trichostatin A (TSA)-induced increases in E-cadherin staining, were determined as follows: groups of 10-20 batches of cell plates were seeded on 384 black-walled, clear bottom plates [10] at 4,000 cells/20 µL/well in RPMI (Invitrogen), 10% heat-inactivated FBS (Invitrogen), 1% penicillin-streptomycin (Invitrogen) for 24 hours followed by media exchange and 16-18 hour pretreatment of TSA control or test compounds

(ChemDiv, ChemBridge) in serum-free RPMI (Invitrogen) using a Vprep liquid handler (Velocity 11/Agilent). Compound preparation was conducted by transferring 10 mM DMSO stocks of test compounds with an ECHO 550 (Labcyte) into dry polypropylene plates. Compounds were diluted to 10 μM in RPMI using a MicroFill (BioTek). Overnight compound treatment was performed at 37 °C with 5% CO_2 in a Cytomat II Incubator (Thermo Fisher). The process of fixation and antibody staining was conducted using a Vprep, an ELx405 plate washer (BioTek), and Multidrop (ThermoFisher). The process was integrated and time locked using a F3 robotic arm (Thermo Fisher) run with a Polara (Thermo Fisher) scheduler. To decrease the time required during the screening process, steps from the initial microslide staining protocol were consolidated during the HTS validation by combining the permeabilization and blocking steps and the secondary antibody and propidium iodide counterstain steps. Each plate was removed from the Cytomat incubator in a time controlled stagger and the fixation step was completed on the Vprep by removing compound treatment medium and washing one time with Phosphate-buffered Saline (PBS) (Invitrogen) followed by room temperature methanol for 15 minutes. Next, the cell plate was transferred to the ELx405 washer to remove methanol and washed for two cycles with 50 μL/well PBS. 20 μL per well of a mouse monoclonal anti-human E-cadherin antibody (BD Transduction Laboratories) was diluted to a concentration of 625 ng/mL in ice-cold PBS containing 0.2% (v/v) Triton X-100, 1% (w/v) BSA and added to the cell plates using the Multidrop dispenser. The plates were then lidded and incubated for 60 minutes at room temperature. Cell plates were washed for two cycles on the ELx405 washer with PBS. Following the PBS wash, cell plates received 20 μL/well of Alexa Fluor® 488-labeled goat anti-mouse IgG (Invitrogen)

diluted 1:4000 in ice-cold PBS containing 0.2% (v/v) Triton X-100, 1% (w/v) BSA, and 0.1 µg/mL propidium iodide (BD Biosciences). The plates were lidded and incubated at room temperature for 45 minutes. Finally, the cell plates were washed on the ELx405 for 2 cycles with PBS.

The cell plates were immediately imaged on an Isocyte (Blueshift Biotechnologies/Molecular Devices) at 10 mm resolution. The excitation wavelength was 488 nm. Alexa Fluor® 488 and propidium iodide signals were acquired simultaneously using a 510-545 nm band-pass filter and a 600 nm long-pass filter respectively. Images we analyzed using BlueImage 2.0 by first using the propidium iodide channel to define the regions occupied by cells and then interrogating the pixel intensity in this region in the Alexa Fluor® 488 channel. Background fluorescence was subtracted from the Alexa Fluor® 488 channel by subtracting values from local pixels outside the area defined by the propidium iodide channel mask. The data were compiled in two ways: either using a custom-built application to compile plate data into a final batch format of 3,200-6,400 test samples per batch or extracted from Microsoft Excel files into summary records using Pipeline Pilot (Accelrys) protocols. Hits were selected by calculating Z-scores based on the 320 test compound wells on each plate. Wells with Z-scores > 3 were called hits. Hits were reordered from ChemBridge and ChemDiv and retested in duplicate. Confirmed hits were resynthesized and tested in triplicate on a concentration series covering a range of concentrations from 30 µM to 1.5 nM. Compounds that produced concentration-dependent effects on E-cadherin expression were confirmed by mass spectrometry and taken forward for more extensive testing detailed below.

Control wells were used as quality indicators for assay performance. Controls on each plate included 16 wells treated with DMSO (negative control), quadruplicate wells of cells treated with a dose range from 0.078 mM - 5 μM of TSA, and 16 wells plated with SW620$^{sicla-1}$ cells which express high levels of E-cadherin promoted by stable siRNA knockdown of Claudin-1 [11]. For wells treated with the TSA concentration series that hit threshold ($Z > 3$) on average corresponded to 40% of the maximum dose. A "checkerboard" plate in which every other well was treated with either DMSO (negative control) or 1 μM TSA (positive control) was used as the first plate on each screening day to assess response uniformity and to set parameters of image analysis for the entire batch. The suitability for HTS was assessed daily using the Z' statistic [12] with a Z' > 0.5 indicating acceptable data.

Compound Synthesis (1 – 285)

Amide formation from an acid chloride:

A mixture containing 0.1 mmol of the acid chloride, 0.11 mmol of the amine, 0.25 mmol of the appropriate base, such as MP-carbonate or diisopropyl ethyl amine (DIEA) and 2 mL of dichloromethane (DCM) was stirred or rotated for 12 hour or until reaction was judged complete by LCMS analysis. The reaction mixture was filtered if necessary to remove resins or insoluble impurities and the solvents were removed under reduced pressure. The residue was dissolved in a DMSO/methanol mixture and purified by mass-directed HPLC to generate analogs.

Amide formation from a carboxylic acid:

To a mixture containing 0.1 mmol of the carboxylic acid, 0.11 mmol of the appropriate amine, and 2 mL of DCM was added along with a sufficient amount of an

amide bond forming reagent. Typical conditions were adding 2 equivalents of PS-carbodiimide or 1.2 equivalents of EDCI, along with 1.2 equivalents of HOBt and 3 equivalents of DIEA. The reaction mixtures were allowed to stir or rotated for 12 hour or until reaction was judged complete by LCMS analysis. The reaction mixture was filtered if necessary to remove resins or insoluble impurities and the solvents were removed under reduced pressure. The residue was dissolved in a DMSO/methanol mixture and purified by mass-directed HPLC to generate analogs for testing.

General Experimental:

All NMR spectra were recorded on a Varian Inova 400 (400 MHz) spectrophotometer located in the Small Molecule NMR Facility at Vanderbilt University. [1]H chemical shifts (δ) are reported in ppm values downfield from TMS as the internal standard. Data are reported as follows: chemical shift, multiplicity, coupling constant (Hz), and integration. Splitting patterns describe apparent multiplicities and are designated as s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), and br (broad). [13]C chemical shifts are reported in δ values in ppm. Low-resolution mass spectra were obtained on an Agilent 1200 LCMS with electrospray ionization. High-resolution mass spectra were recorded on a Waters Qt of API-US plus Acquity system. Analytical thin layer chromatography was performed on 250 mM silica gel 60 F254 plates. Analytical HPLC was performed on an Agilent 1200 analytical LCMS with UV detection at 214 nm and 254 nm along with ELSD detection. Flash column chromatography was performed on silica gel (230-400 mesh, Merck) or using automated silica gel chromatography (Isco, Inc. 100sg Combiflash).

Compound Characterization

Compound 1
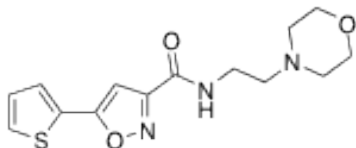
***N*-(2-Morpholinoethyl)-5-(thiophen-2-yl)isoxazole-3-carboxamide**: [1]**H NMR** (MeOD, 600MHz) δ (ppm): 2.55 (s, 4H), 2.61 (t, *J*=6.6Hz, 2H), 3.55 (t, *J*=6.6Hz, 2H), 3.71 (t, *J*=4.6Hz, 4H), 6.91 (s, 1H), 7.21 (q, *J*=3.7Hz, *J*=1.2Hz, 1H), 7.765 (dd,*J*=4.08Hz, *J*=2.76Hz, 2H); [13]**C NMR** (MeOD, 100MHz) δ (ppm): 37.23, 54.69, 58.36, 67.76, 99.27, 129.13, 129.46, 130.29, 160.58, 161.10, 168.00; **HRMS:** $C_{14}H_{18}N_3O_3S$, Calculated: [M+H]$^+$, 308.1069 Found: [M+H]$^+$, 308.1070.

Compound 54
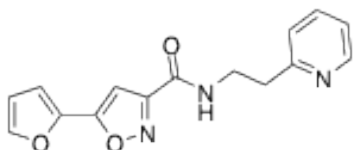
5-(Furan-2-yl)-*N*-(2-pyridin-2-yl)ethyl)isoxazole-3-carboxamide: [1]**H NMR** (MeOD, 400MHz) δ (ppm): 3.33 (m, 4H), 3.84 (t, J=6.5Hz, 2H), 6.65 (m, 1H), 6.81 (s, 1H), 7.06 (d, *J*=3.5Hz, 1H), 7.75 (s, 1H), 7.84 (t, *J*=6.6Hz, 1H), 7.91 (d, *J*=8.0Hz, 1H), 8.40 (t, *J*=7.8Hz, 1H), 8.72 (d, *J*=5.6Hz, 1H); [13]**C NMR** (MeOD, 100MHz) δ (ppm): 33.9, 38.1, 97.5, 111.0, 111.7, 124.5, 127.1, 142.2, 142.5, 144.8, 145.1, 155.3, 158.4, 159.7, 162.9; **HRMS:** $C_{15}H_{14}N_3O_3$, Calculated: [M+H]$^+$, 284.1035, Found: [M+H]$^+$, 284.1035.

Compound 57
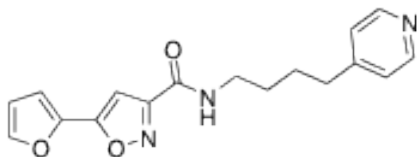
5-(Furan-2-yl)-*N*-(pyridine-4-yl)butyl)isoxazole-3-carboxamide: ^{1}**H NMR** (MeOD, 400MHz) δ (ppm): 1.72 (m, 2H), 1.83 (m, 2H), 3.0 (t, J=7.5Hz, 2H), 3.45 (t, *J*=6.7Hz, 2H), 6.65 (dd, *J*=1.76Hz, *J*=1.76Hz, 1H), 6.87 (s, 1H), 7.06 (d, *J*=3.5Hz, 1H), 7.75 (s, 1H), 7.91 (d, *J*=6.4Hz, 2H), 8.68 (d, *J*=6.4Hz, 2H) ^{13}**C NMR** (MeOD, 100 MHz) δ(ppm): 26.7, 28.32, 34.9, 38.4, 97.6, 111.0, 111.7, 126.8, 141.5, 142.3, 145.0, 158.8, 159.6, 162.8, 163.5; **HRMS:** $C_{17}H_{18}N_3O_3$, Calculated: [M+H]$^+$, 312.1348, Found: [M+H]$^+$, 312.1350.

Compound 73
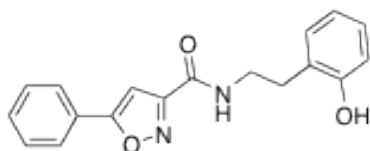
N-(2-Hydroxyphenethyl)-5-phenylisoxazole-3-carboxamide: ^{1}**H NMR** (MeOD, 400MHz) δ (ppm): 2.93 (t, *J*=7.1Hz, 2H), 3.62 (t, *J*=7.1Hz, 2H), 6.77 (m, 2H), 7.04 (m, 2H), 7.12 (d, *J*=7.4Hz, 1H), 7.51 (m, 3H), 7.87 (dd, *J*=2.2Hz, *J*=1.6Hz, 2H); ^{13}**C NMR** (MeOD, 150MHz) δ (ppm): 29.5, 38.8, 98.4, 114.5, 119.2, 125.1, 125.4, 126.7, 127.3, 128.9, 130.3, 130.4, 155.2, 159.2, 159.8, 171.3; **HRMS:** $C_{18}H_{17}N_2O_3$, Calculated: [M+H]$^+$, 309.1239, Found: [M+H]$^+$, 309.1239.

Compound 116
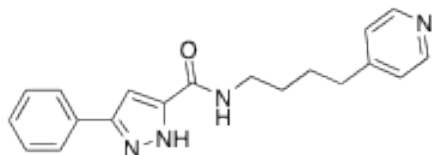
**5-Phenyl-*N*-(4-(pyridine-4-yl)butyl-1*H*-pyrrazole-3-carboxamide:** **^{1}H NMR** (MeOD, 400MHz) δ (ppm): 1.72 (m, 2H), 1.84 (m, 2H), 3.1 (t, J=7.5Hz, 2H), 3.45 (t, *J*=6.7Hz, 2H), 7.02 (s, 1H), 7.37 (t, *J*=7.36Hz, 1H), 7.45 (t, *J*=7.3Hz, 2H), 7.76 (d, *J*=7.3Hz, 2H), 7.92 (d, *J*=6.4Hz, 2H), 8.68 (d, *J*=6.4Hz, 2H); **^{13}C NMR** (MeOD,150MHz) δ (ppm): 26.8, 28.6, 35.0, 38.1, 101.9, 125.14,125.18, 126.8, 128.3, 128.7, 129.5, 141.5, 162.9, 163.7; **HRMS:** $C_{19}H_{21}N_4O$, Calculated: [M+H]$^+$, 321.1715, Found: [M+H]$^+$, 321.1715.

Compound 140
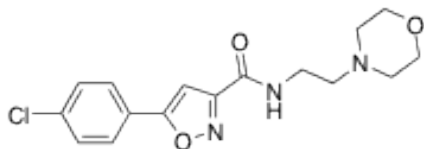
5-(4-Chlorophenyl)-*N*-(2-morpholinoethyl)isoxazole-3-carboxamide: ^{1}H NMR (DMSO-*d6*, 600MHz) δ (ppm): 3.1 (s, 2H), 3.55 (s, 2H), 3.65 (s, 4H), 4.0 (s, 2H), 7.53 (s, 1H), 7.65 (d, *J*=8.6Hz, 2H), 7.98 (d, *J*=8.6Hz, 2H), 9.0 (s, 1H), 9.7 (s, 1H); **^{13}C NMR** ((CD3)2S=O, 150MHz) δ (ppm): 34.01, 51.83, 55.55, 63.76, 101.02, 125.52, 128.12, 129.97, 136.12, 159.42, 159.86, 170.02; **HRMS:** $C_{16}H_{19}N_3O_3Cl$, Calculated: [M+H]$^+$, 336.1115, Found: [M+H]$^+$, 336.1117.

Compound Synthesis (286 – 298)
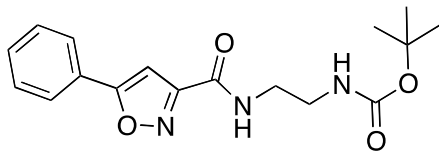
***tert*-Butyl (2-(5-phenylisoxazole-3-carboxamido)ethyl)carbamate:** To a solution containing 1.0 g (5.3 mmol) of the 5-phenylisoxazole acid and 15 mL of DCM was added 0.94 g (5.8 mmol) of carbonyldimidazole. The reaction mixture was allowed to stir at room temperature for 15 minutes and a solution containing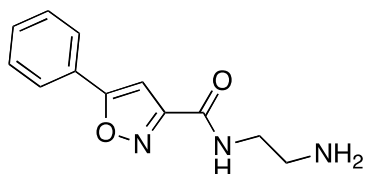 0.89 g (5.6 mmol) of the *tert*-butyl (2-aminoethyl)carbamate in 5mL of DCM was added, followed by 2 mL (11 mmol) of DIPEA. The reaction mixture was allowed to stir at room temperature overnight, quenched by the addition of water, and extracted with DCM. The combined organic layers were dried by passage through a phase separator cartridge and the solvent was removed under reduced pressure. The residue was subjected to silica gel chromatography to give 1.2 g (76%) of *tert*-butyl (2-(5-phenylisoxazole-3-carboxamido)ethyl)carbamate as a white solid. **LCMS:** 1.20 min, $m/z = 354.2$ [M+Na]$^+$.
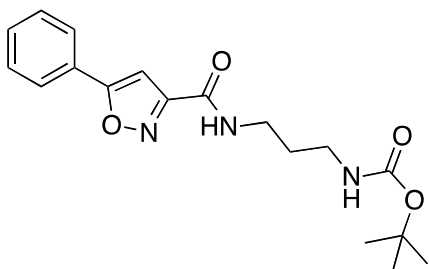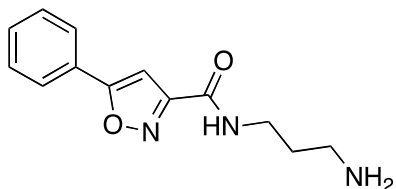

***tert*-Butyl (3-(5-phenylisoxazole-3-carboxamido)propyl)carbamate:** In a similar manner to above, 0.32 g (92%) of *tert*-butyl (3-(5-phenylisoxazole-3-carboxamido)propyl)carbamate was prepared from 0.38 g (2.0 mmol) of 5-phenylisoxazole acid and 0.2 g (1.1 mmol) of (3-aminopropyl)carbamate. **LCMS:** 1.22 min , $m/z = 368.2$ [M+K]$^+$.
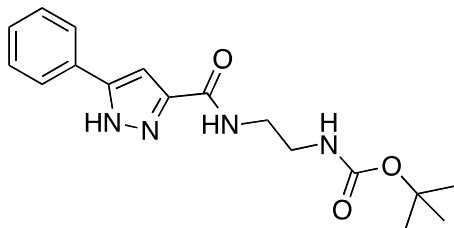
***tert*-Butyl (2-(5-phenyl-1*H*-pyrazole-3-carboxamido)ethyl)carbamate:** In a similar manner, 0.39 g (57 %) of *tert*-butyl (2-(5-phenyl-1*H*-pyrazole-3-carboxamido)ethyl)carbamate was prepared from 0.38 g (2.2 mmol) of 5-phenyl-1*H*-pyrazole-3-carboxylic acid and 0.35 g (2.2 mmol) of (2-aminoethyl)carbamate. **LCMS:** 1.06 min , *m/z* = 354.2 [M+Na]$^+$.

***N*-(2-Aminoethyl)-5-phenylisoxazole-3-carboxamide:** A solution containing 1.21 g (3.7 mmol) of *tert*-butyl (2-(5-phenylisoxazole-3-carboxamido)ethyl)carbamate, 15 mL of DCM and 3 mL of TFA was allowed to stir at rt overnight. The solvents were removed under reduced pressure to give a quantitative yield of *N*-(2-aminoethyl)-5-phenylisoxazole-3-carboxamide as its trifluoroacetic acid salt as a white solid. **LCMS:** 0.77 min, *m/z* = 232.2 [M+H]$^+$

***N*-(3-Aminopropyl)-5-phenylisoxazole-3-carboxamide:** In a similar manner, 0.21 g (99.8%) of *N*-(3-aminopropyl)-5-phenylisoxazole-3-carboxamide as its trifluoroacetate salt was obtained as a yellow oil from 0.31 g (1.44 mmol) of *tert*-butyl (3-(5-phenylisoxazole-3-carboxamido)propyl)carbamate: **LCMS:** 0.79 min, *m/z* = 246.3 [M+H]$^+$
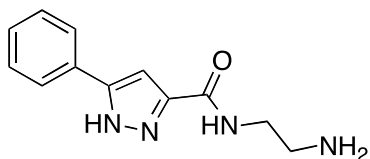
N-(2-Aminoethyl)-5-phenyl-1*H*-pyrazole-3-carboxamide: In a similar manner, 0.28 g (99.8%) of *N*-(2-aminoethyl)-5-phenyl-1*H*-pyrazole-3-carboxamide: as its trifluoroacetate salt was obtained as a yellow oil from 0.39 g (1.44 mmol) of *tert*-butyl (2-(5-phenyl-1*H*-pyrazole-3-carboxamido)ethyl)carbamate: **LCMS:** 0.67 min, *m/z* = 231.3 [M+H]$^+$.

General procedure for amide bond formation:

To a mixture containing 1.1 eq of the appropriate acid, 1mL of DCM, and 1 mL of DMF was added 1.5 eq of HATU or 1.5 eq of EDC and 1.5 eq of HOBt. The reaction mixture was allowed to stir at room temperature for 15 minutes and a solution containing 1.0 eq of the appropriate amine, 3 eq of Hunig's base, and 1 mL of DMF was added. The reaction mixture was allowed to stir overnight. The solvents were removed under reduced pressure and the residue was purified by HPLC.

Compound 289:
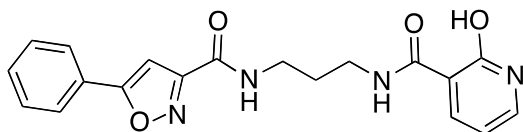
N-(3-(2-Hydroxynicotinamido)propyl)-5-phenylisoxazole-3-carboxamide: Following the general procedure starting with 0.15 mmol of *N*-(3-aminopropyl)-5-phenylisoxazole-3-carboxamide, 14 mg (23%) of *N*-(3-(2-hydroxynicotinamido)-propyl)-5-phenylisoxazole-3-carboxamide was prepared as a white solid: **^{1}H NMR** (400MHz, CD$_3$OD): δ 8.47 (dd, 1H, *J* = 11.2 and 2.0 Hz), 7.88-7.86 (m, 2H), 7.64 (dd, 1H, *J* = 6.4 and 2.0 Hz), 7.55-7.51 (m, 3H), 7.07 (s, 1H), 6.55 (t, 1H, *J* = 6.4 Hz), 3.54-3.48 (m, 4H), and 1.94-1.91 (m, 2H); **LCMS:** 0.98 min, *m/z* = 367.2 [M+H]$^+$.

Compound 290:
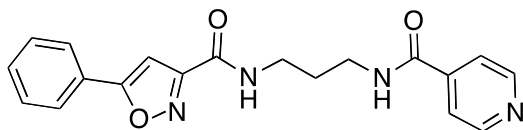
N-(3-(Isonicotinamido)propyl)-5-phenylisoxazole-3-carboxamide: Following the general procedure starting with 0.10 mmol of *N*-(3-aminopropyl)-5-phenylisoxazole-3-carboxamide, 18 mg (45%) of N-(3-(isonicotinamido)propyl)-5-phenylisoxazole-3-carboxamide was prepared as a white solid: LCMS: 0.95 min, *m/z* = 351.2 [M+H]$^+$.

Protein Expression – Western Blot Analysis

SW620 and H520 cells (7.5 x 10^5/mL) were seeded in 6 well plates for 24 hours prior to treatment. Cells were treated with a 10 μM concentration of synthesized compound for 24 hours in RPMI 1640 supplemented medium and 100 μg/mL penicillin-streptomycin. Total protein was isolated from cells with the use of RIPA lysis buffer with protease inhibitors. Lysates were sonicated for 10 seconds, incubated on ice, and centrifuged at 13,000 rpm for 15 minutes at 4 °C. Protein concentration was measured by absorbance at 595 nm using the Bradford assay reagent (Bio-Rad, Hercules, CA) with bovine albumin standards. 10 μg of proteins were loaded per sample onto 10% SDS-polyacrylamide gels that were run at 100V for approximately 1.5 hours. Proteins were electrophoretically transferred to polyvinylidene difluoride (Bio-Rad, Hercules, CA) membranes at 100V for 2 hours. After completion of transfer, the polyvinylidene difluoride membrane was incubated with 10 mL LI-COR blocking buffer (1:1 dilution in PBS; LI-COR Biosciences, Cambridge, UK) for 1 hour at room temperature with gentle agitation. To determine E-cadherin expression level, the membrane was incubated simultaneously with the anti-E-Cadherin (1:2000; BD Transduction Laboratories, San

Jose, CA) and anti-α-Tubulin (1:20,000; Abcam, Cambridge, MA) in 10 mL LI-COR blocking buffer (1:1 dilution in PBS) with gentle agitation overnight at 4 °C. The next day, the antiserum was removed and the membrane was washed 3 times in PBS with 0.1% Tween (PBS-T) before addition of secondary antibody conjugated to a fluorescent entity: IRDye 800-conjugated goat anti-mouse IgG (1:10,000) and IRDye-700-conjugated goat anti-rabbit IgG (1:10,000) in 10 mL LI-COR blocking buffer (1:1 dilution in PBS) with gentle agitation for 1 hour at room temperature. At the end of the incubation period, membranes were washed 3 times with PBS-T. The membrane was visualized and analyzed on the Odyssey IR imaging system (LI-COR Biosciences).

Protein Expression – In Cell Western Analysis

SW620 and H520 cells (10×10^4/100 μL) were seeded in a 96-well plate prior to treatment. Cells were treated with a 10 μM concentration of synthesized compound in quadruplicate for 24 hours in RPMI 1640 supplemented medium and 100 μg/mL penicillin-streptomycin. The cells were then fixed with 100% methanol for 20 minutes at 4 °C. The wells were then washed 2 times with PBS, permeabilized in 2% bovine serum albumin (BSA) and 0.2% Triton X-100 in PBS for 30 minutes at room temperature with gentle agitation, and blocked in LI-COR blocking buffer for 30 minutes at room temperature with gentle agitation. The cells were then incubated with the following primary antibodies: anti-E-Cadherin (1:500) and anti-α-Tubulin (1:1000) diluted in LI-COR blocking buffer (1:1 dilution in PBS) for 2 hours at room temperature with gentle agitation. The cells were washed 4 times in PBS-T for 5 minutes each, and then incubated with the following secondary antibodies conjugated to a fluorescent entity: IRDye 800-conjugated goat anti-rabbit IgG (1:2000) and IRDye 700-conjugated goat

anti-mouse IgG (1:2000) in LI-COR blocking buffer (1:1 dilution in PBS) with gentle agitation for 1 hour at room temperature. The cells were washed 4 times in PBS-T for 5 minutes each followed by a single wash with PBS. All liquid was removed from the wells and the plates were visualized and analyzed on the Odyssey IR imaging system (LI-COR Biosciences).

The assay was further optimized in the following manner. The cells were washed 2 times with PBS, fixed in 100% methanol at room temperature for 15 minutes, and again washed 2 times with PBS. The cells were then incubated with the following primary antibodies: anti-E-Cadherin (1:200) and anti-α-Tubulin (1:2000) diluted in ice cold 2% bovine serum albumin (BSA) and 0.2% Triton X-100 in PBS for 1 hour at room temperature with gentle agitation. The cells were washed 2 times in PBS and then incubated with the following secondary antibodies conjugated to a fluorescent entity: IRDye 800-conjugated goat anti-rabbit IgG (1:2000) and IRDye 700-conjugated goat anti-mouse IgG (1:2000) in ice cold 2% BSA and 0.2% Triton X-100 in PBS with gentle agitation for 45 minutes at room temperature. The wells were washed 2 times in PBS, dried, and analyzed as previously mentioned.

The assay was optimized a third, and final, time in the following manner. The cells were fixed in ice-cold 100% methanol at room temperature for 15 minutes and washed 2 times with PBS. The cells were blocked for 45 minutes in LI-COR blocking buffer (1:1 in PBS with 0.2% Triton X-100) then incubated with the following primary antibodies: anti-E-cadherin (1:500) and anti-α-tubulin (1:2000) diluted in LI-COR blocking buffer (1:1 dilution in PBS) for 1 hour at room temperature. The cells were washed 3 times in PBS (final wash for 10 minutes) and then incubated with the following

secondary antibodies conjugated to a fluorescent entity: Licor 800-conjugated goat anti-rabbit IgG (1:2000) and IRDye-700-conjugated goat anti-mouse IgG (1:2000) in LI-COR blocking buffer (1:1 dilution in PBS) for 45 minutes at room temperature. The wells were washed 3 times in PBS (final wash for 10 minutes), covered with 50 μL of PBS, and analyzed as previously mentioned.

Results

High-Throughput Screen to Discover Lead Compounds

A phenotypic assay suitable for high-throughput screening was developed to identify small molecules that restore E-cadherin expression in SW620 cells. The screen detects E-cadherin restoration at the cell membrane after a 16-18 hour incubation with test compounds via the use of an anti-E-cadherin monoclonal antibody and subsequent secondary antibody-based Alexa Fluor® 488 visualization. Images were obtained using a novel plate-based laser-scanning fluorometer, Blueshift Isocyte, that is capable of obtaining images at 2.5-10 μm resolution in all wells of a 384 well plate in 2-10 minutes/plate. For quantification, a second channel of data was obtained by staining cells with propidium iodide. The propidium iodide counterstain labels cellular nucleic acids, allowing E-cadherin expression levels, as judged by Alexa Fluor® 488 fluorescence, to be normalized to cell number on a per-well basis (**Figure 1**). The image analysis was performed using BlueImage 2.0 and data were compiled into reports using custom-written applications and Pipeline Pilot. In order to judge assay performance, positive and negative controls were run on each plate. The positive controls consisted of a small

molecule HDAC inhibitor, TSA, and SW620$^{sicld-1}$ cells, which express high levels of E-cadherin by virtue of siRNA knockdown of Claudin-1. After optimization, the screen was found to be suitable for screening in 384 well plates with a Z' averaging > 0.5 (**Figure 1**). The entire optimized assay protocol was automated on a robotic screening system built around an F3 robotic arm (Thermo Fisher) running on a Polara scheduler (Thermo Fisher).

The automated assay was used to conduct a high-throughput screen of 83,200 small molecules from the Vanderbilt Institute of Chemical Biology library, which was built using a diverse set of molecules available from the ChemDiv and Chembridge collections. The hit threshold was set at 3 standard deviations above the average fluorescent intensity calculated from the test compound wells. The screen revealed thirty confirmed hits, four of which produced concentration-dependent effects in the primary screening assay.

| | Claudin-1 KO | DMSO | 1 μM TSA | 10 μM 1 |
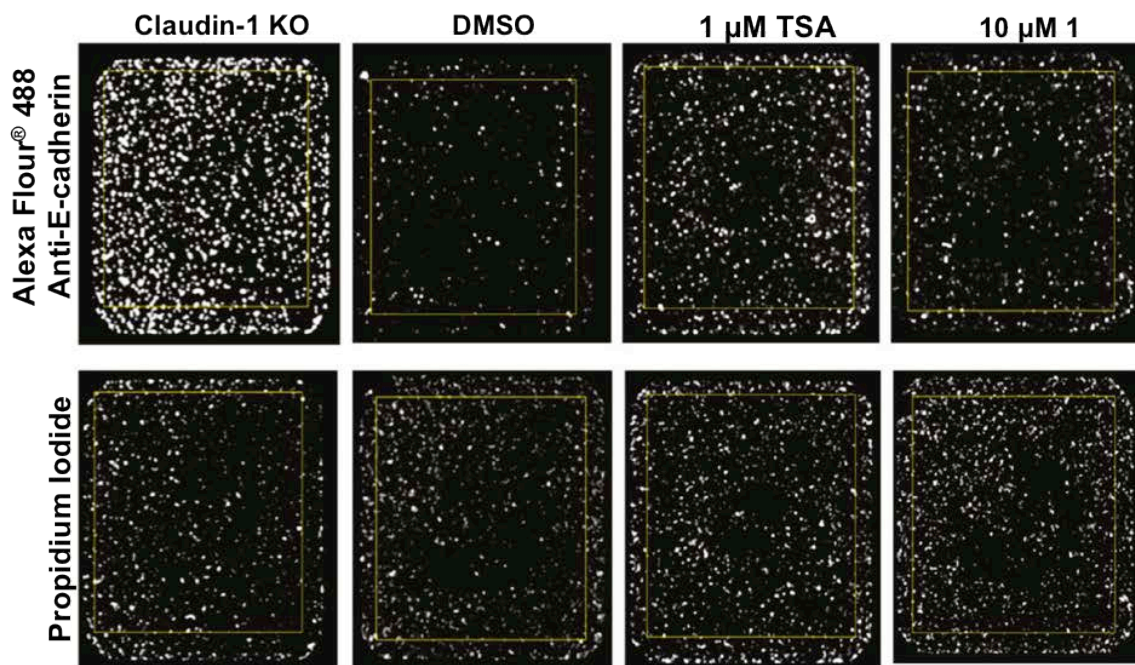
Z' for confirmation plates =0.63	Claudin-1 KO (Positive)	DMSO (negative)	TSA (positive)	VU0075630 (Selected HIT)
Plate 1 Mean Intensity Values	15500	2944 +/- 353	5538	3948
Plate 2 Mean Intensity Values	15542	2965 +/- 397	5570	4122

Figure 2. A high-throughput screen using immunofluorescent staining of human colorectal cancer cells (SW620) identified increased E-cadherin expression with treatment of compounds from the Vanderbilt Institute of Chemical Biology small molecule library. Dual-color acquisition images show the E-cadherin levels of four representative wells from a 384-well assay plate. The table shows average values for positive (TSA, Claudin-1 KO) and negative (DMSO) controls and a compound selected from the primary screen.

Preliminary SAR

Among the confirmed and validated hits from the initial E-cadherin restoration screen were four compounds with somewhat similar chemotypes (**Figure 2**). Each of the hit structures was re-synthesized using standard chemistry routes, and the re-synthesized molecules were tested to confirm activity in an E-cadherin restoration assay. As it is known that iron loading can play a role in the expression of E-cadherin [2], and that acyl hydrazone based on compounds **2**, **3**, and **4** structures can chelate iron effectively [3, 13], we elected to focus our optimization efforts on compound **1**. Employing a parallel synthesis library approach, we first explored the eastern amine tail of **1** while maintaining the western 2-thiophenylisoxazole core. This first generation library explored a wide range of functionalized amines as well as structural fragments from **2**, **3**, and **4**.
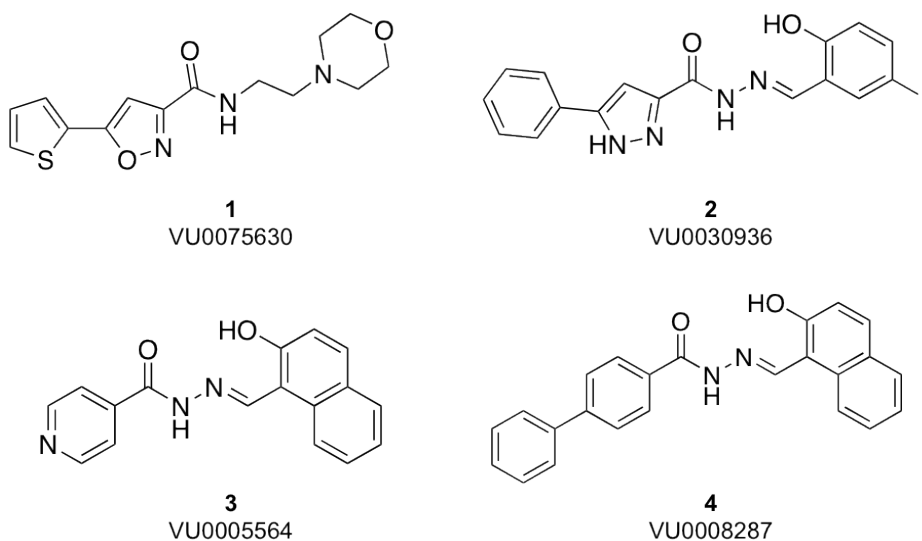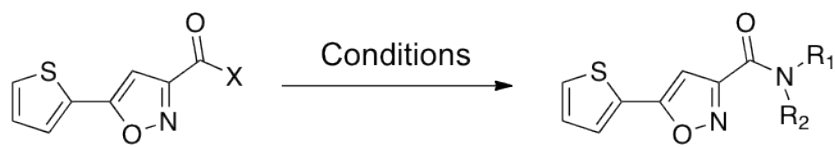

Figure 3. Selected hits **1-4** from the E-cadherin restoration screen.

Selected results from the first libraries are shown in Table 1, wherein compounds were synthesized using standard amide bond forming reactions using commercially available carboxylic acids or acid chlorides and amines. All compounds were evaluated initially at 10 μM in a Western blot assay measuring the ability of the compounds to restore E-cadherin expression in SW620 cells, using TSA as a positive control. The data are expressed as fold change in E-cadherin expression above the DMSO control treatment. Compound **1**, the resynthesized HTS hit, was confirmed with a 5.68-fold increase in E-cadherin expression. Elongating the linear linker that serves to attach the morpholine in compound **24** produced only a slight diminishment of activity (3.51-fold). However, removing the oxygen atom of the morpholine (**11**) led to a sharp decrease in potency (0.96-fold). Several additional attempts to tether a basic amine onto the structure (**8, 18, 20**) also were met with little success. Interestingly, replacement of the morpholine moiety in **1** with an ether (**9**) afforded a compound of only slightly reduced efficacy (4.3-fold). In addition, a number of heteroaromatic moieties (**6, 17, 21, 32**) were well tolerated as morpholine replacements (4.9- to 13.2-fold). In particular, the imidazole **21** (8.11-fold) and the pyridine **32** (13.25-fold) produced compounds more efficacious than compound **1**. Interestingly, substitution on the amide nitrogen is not tolerated, as **35**, the *N*-Me analog of **32** (13.2-fold), provides only a 1.63-fold increase in E-cadherin expression. Additional analogs with a substitution on or adjacent to the amide nitrogen, such as compounds **12-15**, also displayed sharply reduced potency. Overall, these preliminary libraries provided robust SAR and suggest that the presence of a hydrogen bond acceptor may be a key structural feature leading to enhanced E-cadherin expression.
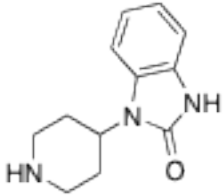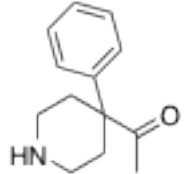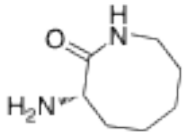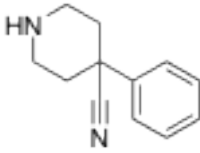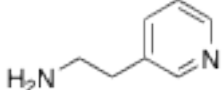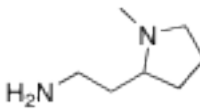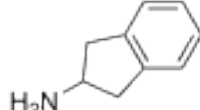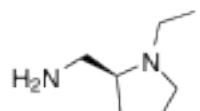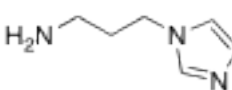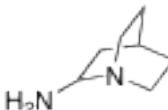
Table 1. Synthesis and evaluation of initial compound library. Variation to the eastern portion of compound **1**.

For X = OH: HNR$_1$R$_2$, EDC, HOBt, DIPEA, THF
For X = Cl: HNR$_1$R$_2$, DIPEA, DCM

Compound	R	WB[a] SW620	WB[a] H520	ICW[b] SW620	ICW[b] H520
1		5.68	5.55	2.19	4.79
5		7.30	10.06	5.03	
6		6.98	9.68	4.71	
7		5.05	10.01		
8		0.97	1.20		
9		4.29	8.05	3.88	
10		5.34	8.01	4.43	
11		0.96			
12		0.84			

49

13		0.92		
14		0.90		
15		0.75		
16		0.80		
17		4.97		5.41
18		0.86		
19		1.36	3.24	
20		1.13		
21		8.11	10.84	4.15
22		1.19		

#	Structure				
23		1.12			
24		3.51	6.10	2.71	
25		1.06			
26		1.25			
27		1.18			
28		1.13			
29		2.29	1.97	1.17	1.10
30		2.71	2.89	4.21	
31		12.90	4.55	4.49	
32		13.25	4.62	3.50	
33		2.50		0.72	1.65
34		8.91	9.20	4.44	

#	Structure				
35		1.63		0.91	0.83
36		8.44	5.38	3.96	4.59
37				1.19	2.12
38				1.17	0.56
39				3.81	3.84
40				3.10	4.16
41				2.58	2.28
42				0.14	1.76

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by Western blot for compound treatment at 10 μM. (TSA = 3.74).
[b] Data is fold change in E-cadherin restoration above a DMSO control as measured by In-Cell Western assay for compound treatment at 10 μM. (TSA = 2.37).

Optimization of Lead Compound 1

      With SAR in hand regarding the eastern amide moiety, we next decided to employ a matrix library approach to explore the western heterobiaryl fragment of **1**, using the eastern amines identified in the first generation library effort (**Table 2**). We also pursued expansions of the SAR around the amine tail groups of particular interest, especially the pyridine tails of **17** and **32**. Selected compounds from Table 1 are included for comparison. As with the morpholine, chain length extension proved successful with an ether-based tail (compare **9** and **10**). The 2-pyridyl congener **31** (12.9-fold) was more than 2-fold more potent compared with **17** (4.97-fold), being essentially the equal of the 4-pyridyl isomer **32** (13.2-fold). While a shortened tether was not successful (not shown), the longer butyl chain of **34** afforded good activity (8.9-fold).

Table 2. Optimization matrix library based on the initial compound **1** analog series.

Substituent	thiophene-isoxazole	furan-isoxazole	methylthiazole-isoxazole	phenyl-isoxazole	pyridyl-isoxazole	fluorophenyl-isoxazole	chlorophenyl-isoxazole	phenyl-pyrazole
morpholinoethyl	**1** 4.25 (1.91)	**47** 1.13 (1.25)	**132** 1.14 (1.01)	**75** 5.38 (2.80)	**187** 1.09 (1.06)	**179** 1.88 (1.08)	**140** 1.71 (0.94)	**124** 1.44 (0.95)
morpholinopropyl	**24** 3.51 (2.71)	**49** 1.18 (1.61)	**134** 0.96 (0.96)	**67** 2.43 (1.29)	N/A	N/A	**145** 1.00 (1.28)	**117** 0.96 (0.73)
methoxyethyl	**9** 4.29 (3.88)	**43** 2.43 (1.58)	**128** 1.01 (1.23)	**162** 3.71 (2.59)	**188** 1.08 (0.94)	**180** 2.92 (2.15)	**139** 1.94 (1.32)	**114** 1.29 (0.87)
isopropoxypropyl	**10** 5.34 (4.43)	**44** 3.55 (3.90)	**129** 1.0 (0.90)	**63** 4.98 (3.37)	**189** 1.17 (1.01)	**181** 3.78 (1.91)	**141** 2.09 (1.01)	**106** 2.10 (1.28)
imidazolylpropyl	**21** 8.11 (4.15)	**46** 2.72 (3.92)	**131** 1.20 (1.13)	**80** 4.66 (5.20)	**191** 1.52 (1.34)	**183** 3.54 (3.01)	**144** 1.97 (1.08)	**115** 1.27 (1.04)
methylfuranylethyl	**6** 6.98 (4.71)	**51** 3.17 (5.89)	**136** 1.06 (0.99)	**68** 4.99 (2.26)	**193** 1.5 (0.80)	**185** 4.2 (1.50)	**147** 1.07 (1.14)	**109** 2.78 (2.09)
pyridin-2-ylethyl	**31** 12.9 (4.49)	**54** 6.53 (3.64)	N/A	**69** 11.99 (3.51)	N/A	N/A	N/A	**111** 1.72 (1.07)
pyridin-3-ylethyl	**17** 4.97 (5.41)	**45** 4.50 (6.46)	**130** 1.23 (1.52)	**65** 4.78 (2.36)	**190** 1.79 (1.14)	**182** 4.48 (3.78)	**142** 2.01 (1.45)	**107** 2.67 (1.61)
pyridin-4-ylethyl	**32** 13.25 (3.50)	**55** 8.60 (3.02)	N/A	**70** 13.69 (1.16)	N/A	N/A	N/A	**112** 3.50 (1.71)
pyridin-4-ylbutyl	**34** 8.91 (4.44)	**57** 9.6 (6.33)	N/A	**72** 9.60 (4.36)	N/A	N/A	N/A	**116** 7.66 (5.68)
hydroxyphenylethyl	**5** 7.30 (5.03)	**50** 4.11 (4.54)	**135** 0.96 (0.95)	**73** 11.25 (5.80)	**192** 1.48 (1.19)	**184** 1.72 (1.43)	**146** 1.31 (1.50)	**122** 6.41 (1.81)
chlorophenylethyl	**36** 8.44 (3.12)	**59** 7.98 (1.37)	N/A	**77** 2.52 (0.91)	N/A	N/A	N/A	**119** 10.27 (1.48)
dimethoxyphenylethyl	**30** 2.71 (4.21)	**53** 2.82 (3.64)	**138** 0.94 (1.12)	**88** 2.32 (2.00)	**194** 1.7 (0.90)	**186** 2.96 (1.76)	**148** 0.90 (0.94)	**110** 4.16 (1.73)

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by Western blot or In-Cell Western (as seen in parentheses) for compound treatment at 10 μM; N/A = not synthesized (TSA = 3.74 Western blot; 2.37 ICW).

Interestingly, we found that phenyl rings with *ortho* substituents, e.g., 2-chlorophenyl derivative (**36**) and 2-hydroxylphenyl derivative (**5**), maintained efficacy. In terms of western heterobiaryl SAR, we found that unsubstituted isoxazoles were devoid of activity (data not shown), but a furyl moiety on the western portion of the molecule was a reasonable alternative for the thiophene ring. A methyl thiazole proved to be inferior to the thiophene in nearly all cases. A phenyl ring was found to be an essentially equipotent replacement for the thiophene, however differences were noted depending on the nature of the amine in the eastern tail. Interestingly, neither a pyridine (**187**), an *ortho*-flouro (**179**) nor a *para*-chloro (**141**) substitution were well tolerated with the morpholine based amine, demonstrating a 2-fold reduction in activity relative to the phenyl substitutent. However, with other eastern moieties, including the ether (**181**), imidazole (**183**), furyl (**185**) and pyrimidinyl (**182**) containing amines, the *ortho*-flouro derivative retains potency. Finally, mixed success was obtained with the introduction of heteroaryl substituents, inspired by the acyl hydrazone hits **2**-**4**, into the western portion. A 4-pyridyl analog was uniformly inactive (not shown) whereas the phenyl pyrazole from **2** retained activity in several examples, the best of which (**116** and **119**) possessed potency as much as 2-fold better than the initial hit **1**. The overall SAR for this matrix library was intriguing, as increased expression of E-cadherin was dependent on the nature of both the eastern and western fragments.
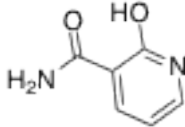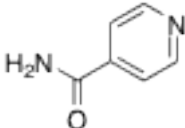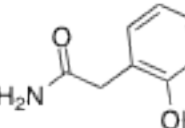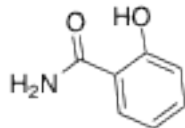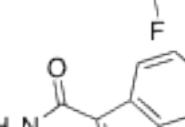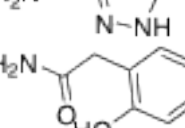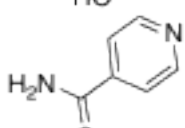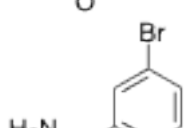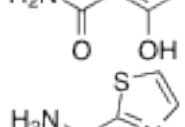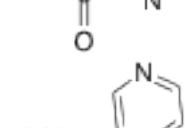
Synthesis of the most resent libraries focused on changes within the linker region to try and expand the SAR through addition of points for further building off the centra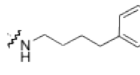l carbon linker (**Table 3**). Originally, the changes made to the linker where in an effort to open up the central carbon linker for potential addition of photoaffinity labels, for future

efforts towards identifying the molecular target. Additionally, hydrogen bond donor and acceptor moieties were moved around the pyridine ring in order to determine which position(s) and moiety types were best tolerated. When these molecules were screened in the ICW assay in both the SW620 and H520 cells many were seen to have improved E-cadherin restoration as compared to the previously profiled compounds. In addition, EC_{50} values of select newly synthesized analogs were much lower, hovering around 1 μM. Further discussion of the EC_{50} values will occur in Chapter 3.

Table 3. General synthesis scheme and representative library of analogs that include an amine tethered carbon linker in the mid-section of the molecules.

X = O, NH
n = 1, 2

Conditions: a) CDI, $NH_2CH_2(CH_2)_nNHBOC$, DCM, DIPEA; b) TFA, DCM
c) EDC, HOBT, or HAYU, DIPEA, DMF, DCM

Compound	X	n	R	ICW[a] SW620	EC_{50} (μM)	ICW[a] H520	EC_{50} (μM)
286	NH	1		1.37		1.24	
287	NH	1		1.45		1.04	
288	NH	1		1.77		1.10	

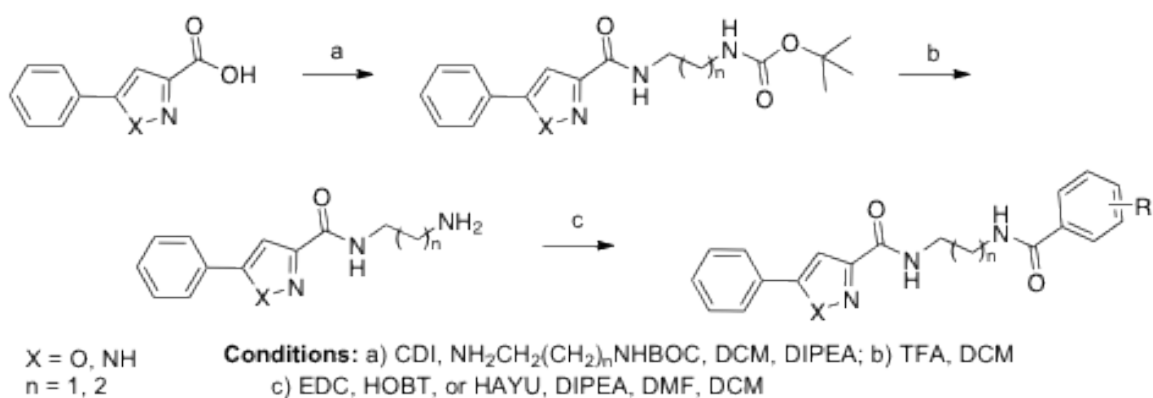|							
289	O	2		5.23	5.0	5.54	1.0
290	NH	1		1.29		1.22	
291	O	2		4.92	4.97	5.92	1.5
292	O	2		2.92	5.0	5.27	1.4
293	O	2		1.62		1.88	
294	O	1		3.40	3.7	3.51	5.6
295	O	1		2.79	8.8	2.87	5.1
296	O	1		1.18		1.14	
297	O	1		8.98	4.8	4.22	
298	O	1		3.05	7.1		

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by In-Cell Western assay for compound treatment at 10 μM. (TSA = 2.37 in SW620 cell line).

Screening of Analogs to Quantify E-cadherin Restoration

The initial HTS screening hits were confirmed using a standard Western blot procedure using enhanced chemiluminescence (ECL) to detect E-cadherin protein on the Western blot membrane (**Figure 4**). The major disadvantage to using ECL was that we were unable to quantify the band intensity and thus were left making decisions on the compounds' abilities to restore E-cadherin expression by comparing band intensity and size by eye.
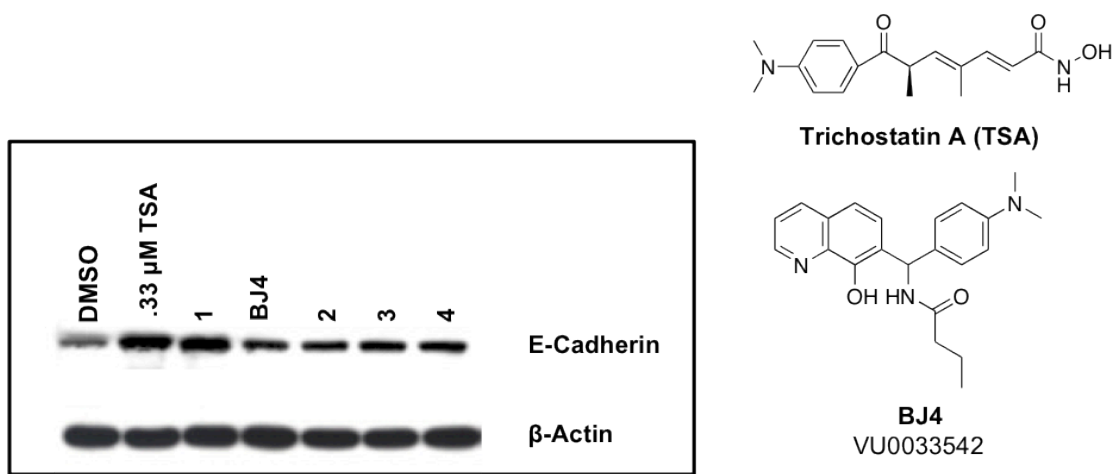
Figure 4. Preliminary Western blot to confirm restoration of E-cadherin protein expression by positive screening hits. SW620 cells were treated with a 10 µM treatment for 24 hours. DMSO is used as a negative control, TSA is used as positive control, and β-actin is used as a loading control. Additionally, the chemical structure for TSA and BJ4 (another positive screening hit) are displayed on the right.

This method of analysis continued through several preliminary libraries of compound **1** analogs until we decided to take advantage of the fluorescent secondary antibodies used with the Odyssey Imaging System allowing for the quantification of

fluorescent band intensity. By quantifying the E-cadherin band intensity for each compound we were able to calculate the fold change by normalizing to the DMSO control treatment. This allowed for a more accurate comparison of compounds to the negative control (DMSO), positive control (TSA), and original screening hit (**1**). Representative blots from the SW620 and H520 cells treated with selected analogs can be seen below. Additionally, the derived graphs below confirm that changes in E-cadherin expression can be both visualized and quantified (**Figure 5**).
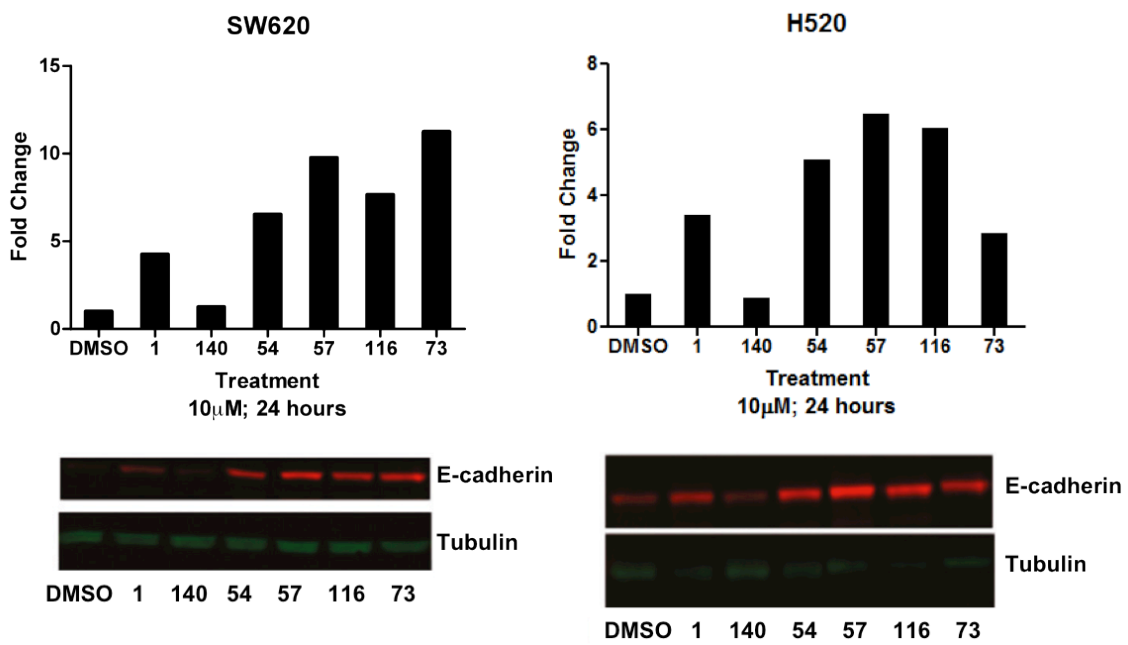
Figure 5. Representative images of Odyssey Imaging System Western blot read-out from SW620 (left) and H520 (right) protein samples and the subsequent calculation of fold change as normalized to the DMSO control. Cells were treated for 24 hours with a single point 10 μM treatment. Whole cell lysates were collected and subjected to Western blot analysis. Samples were probed for E-cadherin; Tubulin was the loading control. Values were quantified as the fold change in E-cadherin expression above the DMSO control.

Ultimately, we wanted to develop a method that would be high throughput and allow for replicate testing of the synthesized analogs. Piggy-backing from the original HTS assay and our use of the Odyssey Imaging System, we developed and utilized an In-Cell Western (ICW) assay. The ICW assay would allow for triplicate or quadruplicate replicates for each concentration and direct quantification of the intensity of fluorescent secondary antibody in each well. In addition, we could screen up to 32 compounds per 96-well plate instead of 15 compounds per Western blot. Use of the ICW assay also allowed for fixation, staining, and analysis of the treatments in a single day, approximately 3 hours, as opposed to an overnight incubation with primary antibody necessary with the Western blot. Preliminarily, an optimized single point, 10 μM concentration, ICW assay was developed and representative wells with calculated data for both the SW620 and H520 cell lines can be seen in Figure 6. While the sensitivity was not as great in the ICW, it did trend with the Western blot data, suggesting that such an assay could be utilized to screen all future synthesized analog libraries; data for which can be seen in Table 2 and Table 3.
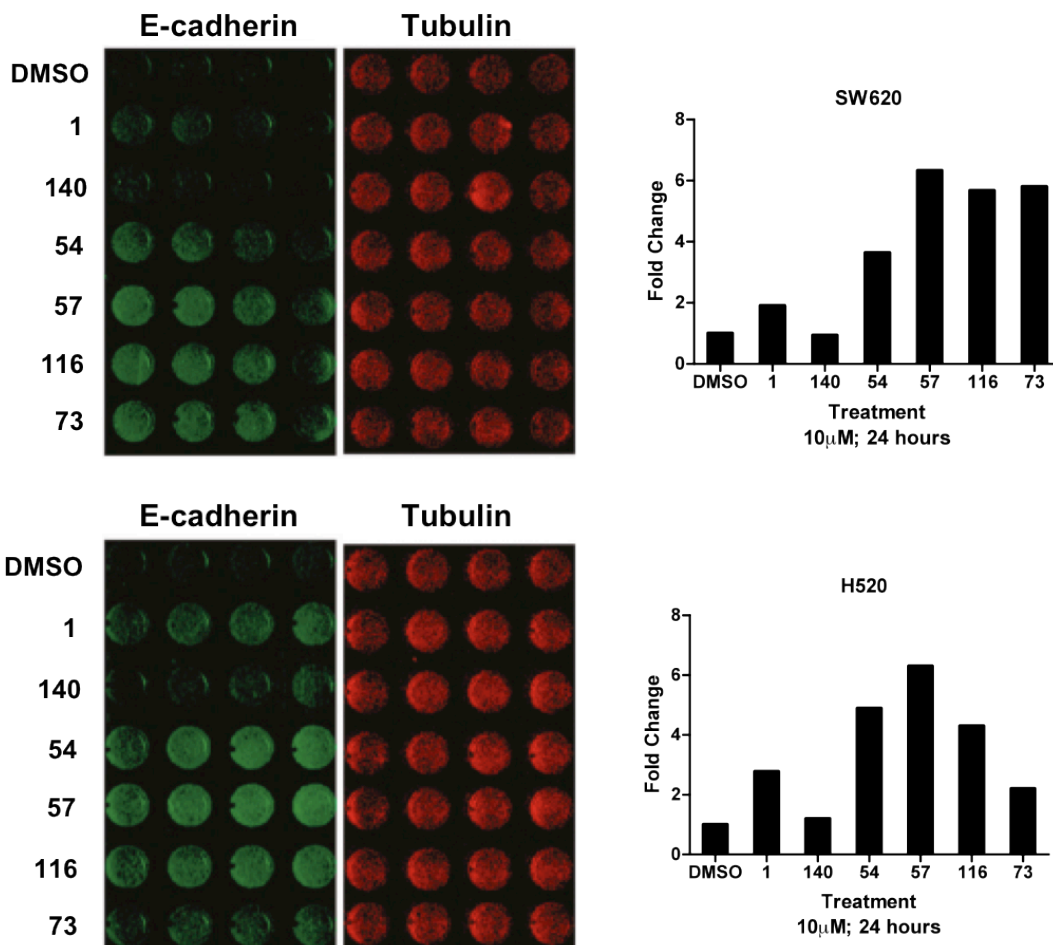
Figure 6. Snap shot images of the Odyssey Imaging System read-out for the ICW assay with the SW620 (top) and H520 (bottom) cells and the subsequent calculation of fold change as normalized to the DMSO control. Cells were treated for 24 hours with a single point 10 µM treatment. After treatment, cells were fixed, permeabilized, and probed for E-cadherin and Tubulin (control). The In-Cell Western assay measured fluorescence intensity of the secondary antibodies directly in each well as seen above using the Odyssey Imaging System. For each treatment the intensity values were averaged and normalized to the DMSO control at 1; this is shown as fold change in the corresponding graph.

Discussion

Here we have shown the development and application of two unique assays. The first is a novel high-throughput screen used to identify compounds capable of restoring E-cadherin expression levels in SW620 cells, which are deficient in E-cadherin. The use of

a phenotypic assay to identify a lead compound is very rare and comes with several hurdles. The major hurdle is that mechanism of action and specifically the molecular target for which the active compounds are binding remains unknown. However, there are two major advantages to phenotypic screening assay. One is that the desired phenotypic response, E-cadherin protein expression, is achieved in the positive screening hits. The other is that since the desired response is being achieved by positive screening hits it is known that molecules are able to cross the cell membrane, a problem seen when *in vitro* enzymatic and biochemical assays developed for a specific protein's activity are used to screen synthesized analogs prior to cell-based experiments.

The second assay developed was an ICW assay that permitted the quantification of E-cadherin restoration in a relatively high-throughput manner using the Odyssey Imaging System. Several iterations of the ICW assay protocol allowed for further reduction in assay length as well as increased sensitivity such that the fold changes calculated for compounds were more similar to those seen by the standard Western blot procedure. Increased sensitivity allowed for clear separation between active and inactive compounds, and thus the standard Western blot was not needed for confirmation. This allowed for screening solely using the ICW assay. Overall, the ICW assay allowed for single point screening of synthesized compound libraries in triplicate or quadruplicate. It also allowed for the establishment of concentration response curves, and quantification of an EC_{50} value, for each positive compound, which will be discussed in the following Chapter.

Further, a matrix library approach, conducted using iterative parallel library synthesis, was undertaken in order to examine a broad spectrum of chemical alterations to

62

both the eastern amine tail and the western heterobiaryl portion of the original lead compound. Such an approach allowed for a relatively rapid synthesis of nearly 100 compounds that were preliminarily screened in both the Western blot and/or ICW assays. This permitted the visualization of chemical structural trends, allowing for more specific library synthesis in order to further optimize the original lead compound, **1**.

In an effort to expand the SAR through the middle section carbon linker, we were able to identify compounds that further improved the restoration of E-cadherin protein expression. Insertion of a carbon linker attached to the western and eastern portions of the molecule by amines will provide additional areas to expand the molecule. This may also allow for the insertion of a photoaffinity moiety, such as an azide, in order to utilize affinity chromatography as a mechanism by which to identify the molecular target.

In total, 300 compounds have been synthesized and screened to date, with the identification of a handful of compounds that restore E-cadherin expression greater then 9- to 10-fold in both the SW620 and H520 cells in the Western blot assay, and greater then 7-fold in the ICW assay. All synthesized compounds to date can be seen in Appendix 1, Tables 1-20 at the back.

References

1. Cavallaro, U. and G. Christofori, *Cell adhesion and signalling by cadherins and Ig-CAMs in cancer.* Nat Rev Cancer, 2004. **4**(2): p. 118-32.

2. Birchmeier, W. and J. Behrens, *Cadherin expression in carcinomas: role in the formation of cell junctions and the prevention of invasiveness.* Biochim Biophys Acta, 1994. **1198**(1): p. 11-26.

3. Hirohashi, S., *Inactivation of the E-cadherin-mediated cell adhesion system in human cancers.* Am J Pathol, 1998. **153**(2): p. 333-9.

4. Yoshiura, K., et al., *Silencing of the E-cadherin invasion-suppressor gene by CpG methylation in human carcinomas.* Proc Natl Acad Sci U S A, 1995. **92**(16): p. 7416-9.

5. Hennig, G., et al., *Progression of carcinoma cells is associated with alterations in chromatin structure and factor binding at the E-cadherin promoter in vivo.* Oncogene, 1995. **11**(3): p. 475-84.

6. Yoo, C.B. and P.A. Jones, *Epigenetic therapy of cancer: past, present and future.* Nat Rev Drug Discov, 2006. **5**(1): p. 37-50.

7. von Burstin, J., et al., *E-cadherin regulates metastasis of pancreatic cancer in vivo and is suppressed by a SNAIL/HDAC1/HDAC2 repressor complex.* Gastroenterology, 2009. **137**(1): p. 361-71, 371 e1-5.

8. Swinney, D.C. and J. Anthony, *How were new medicines discovered?* Nat Rev Drug Discov. **10**(7): p. 507-19.

9. Hewitt, R.E., et al., *Validation of a model of colon cancer progression.* J Pathol, 2000. **192**(4): p. 446-54.

10. Kramer, O.H., et al., *A phosphorylation-acetylation switch regulates STAT1 signaling.* Genes Dev, 2009. **23**(2): p. 223-35.

11.   Dhawan, P., et al., *Claudin-1 regulates cellular transformation and metastatic behavior in colon cancer.* J Clin Invest, 2005. **115**(7): p. 1765-76.

12.   Zhang, J.H., T.D. Chung, and K.R. Oldenburg, *A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays.* J Biomol Screen, 1999. **4**(2): p. 67-73.

13.   van Roy, F. and G. Berx, *The cell-cell adhesion molecule E-cadherin.* Cell Mol Life Sci, 2008. **65**(23): p. 3756-88.

CHAPTER III

BIOLOGICAL EVALUATION OF SMALL MOLECULES THAT RESTORE

E-CADHERIN EXPRESSION

Introduction

The majority of human cancers arise from epithelial cells, which are held together through junction structures: tight junctions, adherens junctions, and desmosomes [1]. There are several classes of cell adhesion molecules, including cadherins, immunoglobulin-like cell adhesion molecules (Ig-CAMs), the hyluronan receptor CD44, and integrins [2] . The development of malignant tumors, in particular the transition from benign growths to more invasive or metastatic cancer, is often characterized by a tumor cell's ability to overcome cell-to-cell adhesion and to invade the surrounding tissue, lymph system, and the circulatory system. During the transition from a normal epithelial cell to a malignant (mesenchymal-like) cell, expression of some of these junction molecules is drastically reduced or terminal. This is often referred to as the epithelial-to-mesenchymal (EMT) transition, and is believed to play a prominent role in invasion, extravasion, and colonization during metastasis [3].

Cell-adhesion molecules are implicated in human carcinogenesis, and recently much attention has been directed towards E-cadherin [2]. E-cadherin is a single-spanning transmembrane domain protein that forms homodimers at the cell surface membrane and interacts with the corresponding E-cadherin homodimers of neighboring cells (**Figure 1**). Aside from cell-to-cell adhesion, E-cadherin is a key component in cell polarity induction and epithelium organization. The loss of E-cadherin function elicits active signals that support tumor-cell migration, invasion, and metastatic dissemination [4, 5].
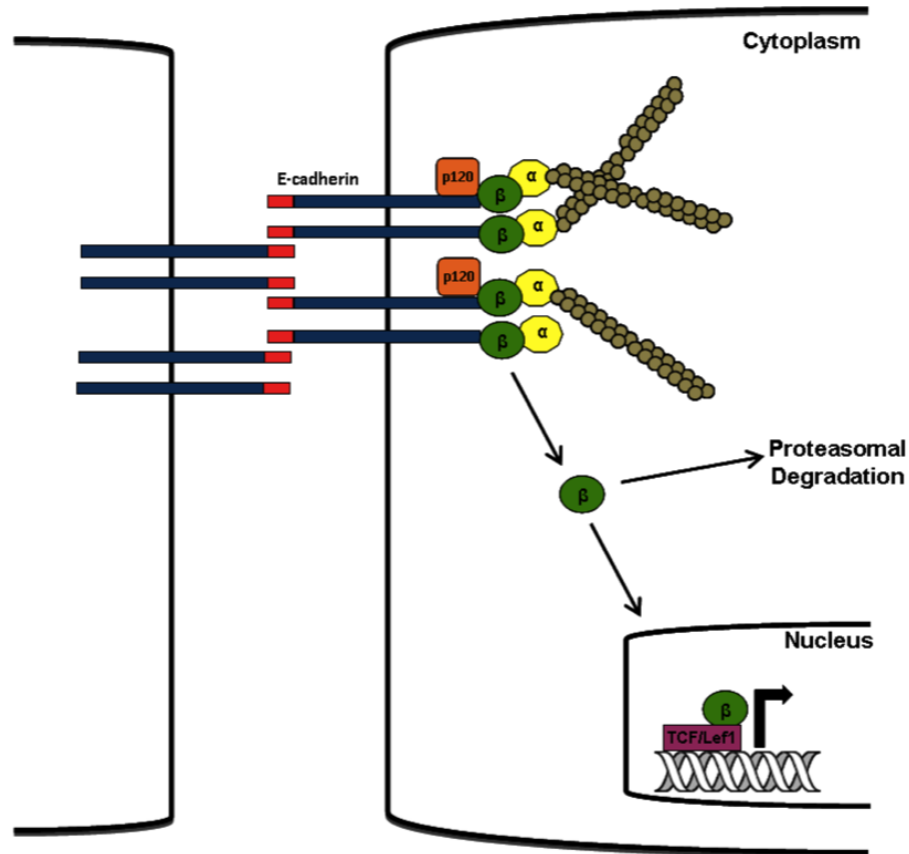
Figure 1. E-cadherin is a single-transmembrane spanning molecule that forms homodimers at the cellular membrane and interacts in a zipper-like manner with homodimers on neighboring cellular membranes. The cytoplasmic cell-adhesion complex of E-cadherin consists of p120, β-catenin, and α-catenin, which links E-cadherin homodimers to the actin cytoskeleton. These interactions aid in cell polarity induction and epithelium organization. Loss of E-cadherin leads to disassembly of the cytoplasmic cell-adhesion complex and release of p120, β-catenin, and α-catenin, into the cytoplasm. There are two pathways for the β-catenin protein to take once released into the cytoplasm. First, β-catenin can be sequestered by the APC/GSK-3β complex and ultimately tagged for proteosomal degradation via ubiquitination. Second, upon activation of the Wnt signaling pathway or mutation in the APC/GSK3β complex, β-catenin can no longer be phosphorylated, and tagged for degradation; and therefore is translocated to the nucleus where together with TCF/Lef1 transcription factor it modulates the expression of target genes. These genes are known to be involved in cell proliferation and tumor progression (Abbreviated from Chapter 1, Figure 4).

The loss of E-cadherin function during tumor progression can be caused by genetic or epigenetic mechanisms, the most common of which is down regulation at the transcriptional level. Repressor transcription factors Snail, Slug, and SIP1, as well as the helix-loop-helix transcription factor E12/E47, have been found to bind to the E2 boxes in

the promoter region of the E-cadherin gene and actively repress its transcription. DNase I hypersensitive site mapping indicated the loss of transcription factor binding, resulting in chromatin rearrangement in the regulatory region of the E-cadherin gene. As a direct consequence of transcriptional inactivation, the E-cadherin locus is epigenetically silenced by hypermethylation and deacetylation [6-10]. It was shown through cloning and sequencing of the E-cadherin gene promoter that CpG methylation around the promoter region of the E-cadherin gene was present in cell lines that lacked E-cadherin expression and that E-cadherin could be restored in these cell lines upon treatment with the DNA methyltransferase inhibitor 5-azacytidine [11, 12]. Deacetylation of histone lysine residues by histone deacetylase (HDAC) enzymes results in chromatin compaction and inactivation of genes. Deacetylation has been shown to occur around the E-cadherin gene promoter region by a repressor complex comprised of Snail, HDAC1, and HDAC2. It has been shown that Snail preferentially binds the E2 box in the promoter region, while binding directly to HDAC2 and indirectly to HDAC1. Treatment of cell lines that have reduced E-cadherin expression with Trichostatin A (TSA), a histone deacetylase inhibitor, leads to restored expression of E-cadherin [13, 14].

As reported in Chapter 2, we developed and used a high-throughput screen to discover lead compounds that restored E-cadherin expression in a metastatic colon adenocarcinoma cell line, SW620, which exhibited low levels of E-cadherin expression. We were able to develop preliminary structure-activity relationships (SAR) around the lead compound **1** using a Western blot-based E-cadherin restoration assay to screen analogs. In addition, we were able to develop an In Cell Western (ICW) assay with the Odyssey Imaging System. Within this chapter we will discuss use of the ICW to quantify

69

EC$_{50}$ values for selected active compounds. Additionally, we were able to confirm E-cadherin restoration via visualization of E-cadherin and β-catenin at the membrane after treatment with profiled compounds. The compounds reduced the invasion of colon cancer cells (SW620) and lung cancer cells (H520) with minimal effects on cellular proliferation. Finally, we show that the compounds increase acetylation of the H4 histone, but do not appear to function via HDAC inhibition, leading to a preliminary high-throughput attempt at identifying the molecular target through submitting selected analogs for outsourced screening.

Methods and Materials

Cell Culture

A colorectal adenocarcinoma cell line, SW620, and a squamous cell lung carcinoma cell line, H520, were obtained from the American Type Culture Collection (Manassas, VA) and maintained in a humidified atmosphere of 5% CO_2 in air at 37 °C. The cells were routinely cultured in RPMI 1640 supplemented medium with 10% fetal bovine serum (Atlanta Biologicals, GA) and 100 μg/mL penicillin-streptomycin.

Protein Expression – Western Blot Analysis

SW620 and H520 cells (7.5 x 10^5/mL) were seeded in 6 cm round plate 24 hours prior to treatment. Cells were treated with a 10 μM concentration of synthesized compound for 24 hours in RPMI 1640 supplemented medium and 100 μg/mL penicillin-streptomycin. Total protein was isolated from cells with the use of RIPA lysis buffer with

protease inhibitors in order to determine E-cadherin protein expression levels. The complete protocol can be referred to in the Methods & Materials section of Chapter 2.

Protein Expression – In Cell Western Analysis

SW620 and H520 cells (10 x 10^4/100 μL) were seeded in a 96-well plate prior to treatment. Cells were treated with a 10 μM concentration of synthesized compound in quadruplicate for 24 hours in RPMI 1640 supplemented medium and 100 μg/mL penicillin-streptomycin. The complete protocol can be referred to in the Methods & Materials section of Chapter 2.

Immunofluorescence Analysis and Deconvolution Microscopy

SW620 and H520 cells (7 x 10^4/well) were seeded on 8-well chamber slides for 24 hours prior to treatment. Cells were treated with 10 μM concentration of synthesized compound for 24 hours in RPMI 1640 supplemented medium and 100 μg/mL penicillin-streptomycin. The cells were rinsed with PBS and fixed with 100% methanol for 15 minutes at 4 °C. The cells were rinsed with PBS and blocked and permeabilized with 2% BSA and 0.2% Triton X-100 in PBS. The cells were then incubated with the following primary antibodies: anti-β-catenin and anti-E-cadherin diluted in 1% BSA in PBS blocking solution overnight at 4 °C. After 3 washes with PBS, the cells were incubated with appropriate secondary antibodies conjugated to Fluorescein (1:200; Sigma, St. Louis, MO), Texas Red (1:500; Invitrogen, Carlsbad, CA), and DAPI (1:2000, Sigma, St. Louis, MO) for 40 minutes at room temperature. The cells were washed 3 times with PBS, and then mounted with Vectashield mounting medium (Vector Laboratories, Burlingame, CA). Deconvolution microscopy analyses were performed using the

DeltaVision® Core (Applied Precision) microscope. All images were taken using a 600x oil immersion objective and converted to a TIFF format and arranged using Photoshop 7.0 (Adobe, Seattle, WA).

Viability Assay

SW620 and H520 cells (2.5 x 10^4/100 μL) were seeded in 96-well microtiter plates prior to treatment. Cells were treated with 10 μM concentration of synthesized compound in quadruplicate for 24 hours and 48 hours in RPMI 1640 supplemented medium and 100 μg/mL penicillin-streptomycin. The Quick Cell Proliferation Assay Kit from BioVision (Mountain View, CA) was used to measure proliferation. The RPMI media was removed and replaced with 100 μL of the WST-1/ECS reagent diluted 1:10 in RPMI supplemented medium. The plates were incubated for 1 hour in an incubator with 5% CO_2 in air at 37 °C. The change in proliferation was quantified by measuring the absorbance of the dye solution at 450 nm on a microtiter plate reader.

Proliferation Analysis

SW620 and H520 cells (2.5 x 10^4/100 μL) were seeded in 96-well microtiter plates prior to treatment. Cells were treated with a 10 μM concentration of selected compounds in triplicate for 48 hours in RPMI 1640 supplemented medium and 100 μg/ml penicillin-streptomycin. The CycLex® Cellular BrdU ELISA Kit from MBL International (Woburn, MA) was used to measure proliferation. The RPMI medium was removed and replaced with 100 μL of 1x BrdU label mix in RPMI medium for 2 hours at 37 °C in 5% CO_2 in the air. The BrdU label mix was removed and 200 μL of the Fix/Denature solution was added and incubated for 30 minutes at room temperature. The

plate was drained, incubated with 50 µL of primary antibody for 1 hour at room temperature, rinsed with wash buffer, and incubated with 50 µL of secondary antibody. The wells were rinsed with wash buffer followed by a single rinse with PBS and drained. 50 µL of substrate solution was added and incubated for 6 minutes followed immediately by 50 µL of Stop solution. The change in proliferation was quantified by measuring the absorbance of the dye solution at 450 nm on a microtiter plate reader.

Invasion Analysis

SW620 or H520 cells (2.5×10^5/mL) were seeded in 6-well plates prior to treatment. Cells were treated with a 10 µM concentration of synthesized compound for 24 hours in RPMI 1640 supplemented medium and 100 µg/mL penicillin-streptomycin. 40 µL (2.5 mg/mL) of BD Matrigel Basement Membrane Matrix (BD Biosciences, Bedford, MA) was added to the top of the insert of a 24-well Transwell Permeable Support plate with a polycarbonate membrane with 8 µm pore size (Corning Inc, Corning, NY). Then the cells were trypsinized and 3×10^5/250 µL cells were added to the top of the chamber in serum-free RPMI medium, and 1 mL of RPMI medium with 10% FBS was added to the bottom of the well. Then the plates were incubated for 72 hours at 37 °C in a humidified incubator with 5% CO_2 in the air. Then the wells were stained with 1% crystal violet in 50% methanol for 1 hour and washed in PBS. The membrane was cut off, adhered to a slide with glycerol, and analyzed in 20x field via microscopy. 3 − 20x fields were quantified per membrane.

Cell viability of the remaining cells plated in the top chamber was observed using Calcein AM. The media was carefully removed from the top chamber, 1 µM

concentration of Calcein AM in RPMI medium (100 µL) was added to each well, and the plate was incubated for 1 hour at 37 °C in a humidified incubator with 5% CO_2 in the air. Calcein staining of viable cells was observed and captured at 200x using a fluorescence microscope.

RNA-Seq Experiment and Analysis

The RNA-Seq experiments and analyses were carried out at HudsonAlpha Institute for Biotechnology (Huntsville, AL). SW620 cells (1.5×10^6/1 mL) were seeded in 3.5 cm dish; 9 total. 3 plates were treated with a 10 µM concentration of DMSO, compound **1**, or compound **57** for 24 hours in RPMI 1640 supplemented medium and 100 µg/mL penicillin-streptomycin. RNA was extracted from each of the samples and quantified prior to shipment to HudsonAlpha. Upon arrival, the 3 samples per treatment were pooled and prepared for sequencing. Raw data was aligned using Tophat and then analyzed through the Cufflinks pipeline on site.

Histone Acetylation Anaylsis

SW620 and H520 cells (7.5×10^5/mL) were seeded in 6 cm round plates prior to treatment. Cells were treated with a 10 µM concentration of synthesized compound for 24 hours in RPMI 1640 supplemented medium and 100 µg/mL penicillin-streptomycin. The Nuclear Extract Kit from Active Motif (Carlsbad, CA) was used to collect the nuclear fraction from each sample. 20 µg of proteins were loaded per sample onto 10% SDS-polyacrylamide gels that were run at 100V for approximately 1.5 hours. The Western blot protocol is same as previous described in the 'Protein Expression – Western Blot Analysis' section. To determine Histone H4 acetylation, the membrane was incubated

with the Acetyl-Histone H4 antibody (1:1000; Millipore, Temecula, CA) in 10 mL LI-COR blocking buffer (1:1 dilution in PBS) overnight at 4 °C. The next day, the antiserum was removed and the membrane was washed in PBS-T before addition of secondary antibody: IRDye 800-conjugated goat anti-rabbit IgG (1:10,000) in 10 mL LI-COR blocking buffer (1:1 dilution in PBS) for 1 hour at room temperature. The membrane was washed in PBS-T and analyzed on the Odyssey IR imaging system. Membranes were stripped with 10 mL of LI-COR stripping buffer (1:5 dilution in ddH$_2$0) for 30 minutes at room temperature with gentle agitation followed by 3 ten-minute washes with PBS. Membranes were re-probed with anti-Histone H4 Pan-acetylation (1:1000; Upstate, Lake Placid, NY), anti-RhoGDIα (1:1000; Santa Cruz Biotechnologies, Santa Cruz, CA), and anti-PARP 1/2 (1:1000; Santa Cruz Biotechnologies, Santa Cruz, CA) in 10mL of LI-COR blocking buffer (1:1 dilution in PBS) for 1 hour at room temperature. The antiserum was removed and the membrane was washed 3 times in PBS-T before addition of secondary antibody: IRDye 700-conjugated goat anti-rabbit IgG (1:10,000) in 10 mL LI-COR blocking buffer (1:1 dilution in PBS) for 1 hour at room temperature. At the end of the incubation period, membranes were washed 3 times with PBS-T and analyzed on the Odyssey IR Imaging System.

Statistical Analysis.

GraphPad Prism Software (GraphPad Software, San Diego, CA, USA) was used to analyze all data.

Visualization of E-cadherin Restoration by Selected Analogs

      As mentioned previously, all compounds were screened by Western blot to measure their ability to restore E-cadherin expression in the SW620 cells. Compounds that restored E-cadherin to a similar or greater level than compound **1** were further screened in the same E-cadherin assay format in the H520 cell line. Four compounds, **54**, **57**, **73**, and **116** were chosen for further profiling based on their performance in the E-cadherin restoration assays and their structural variability. Additionally, compound **140** was utilized as a negative control since it does not restore E-cadherin expression in either the SW620 or H520 cell lines (**Figure 2**).
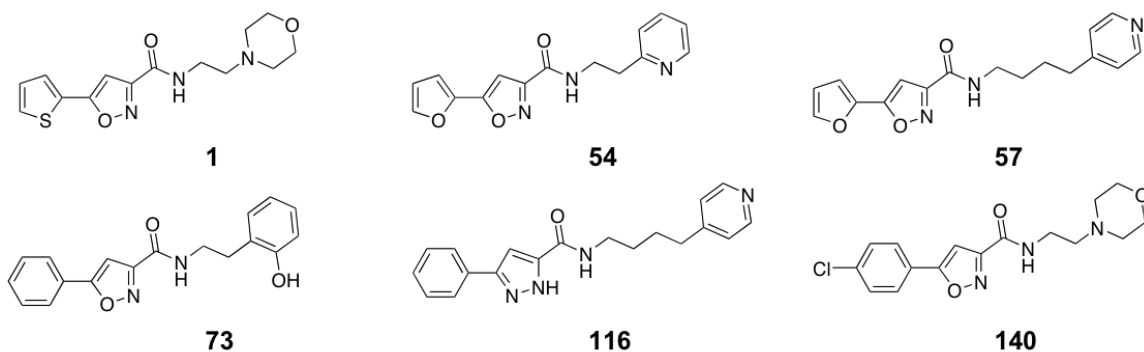
Figure 2. Compounds selected for further profiling.

Figure 3A shows the ability of these compounds to restore E-cadherin in both the SW620 and H520 cell lines at a 10 μM concentration. It can be seen that the compounds display improved E-cadherin restoration compared to **1** and show slight variability between the two cell lines.

As mentioned in Chapter 2, we wanted to develop a method that would facilitate the generation of concentration response curves and quantify EC_{50} values for restored E-cadherin expression. For this, we developed and utilized an ICW assay using the Odyssey Imaging System, which would allow for quadruplicate replicates for each concentration and direct quantification of the intensity of fluorescent secondary antibody within each well (**Figure 3B**). Use of the ICW assay also allowed for fixation, staining, and analysis of the compound treatments in a single day, approximately 3 hours, as opposed to an overnight incubation with primary antibody necessary with the Western blot. Preliminarily, an optimized single point, 10 μM concentration, ICW assay was developed and the data for the profiled compounds in both the SW620 and H520 cell lines can be seen in Figure 3B.
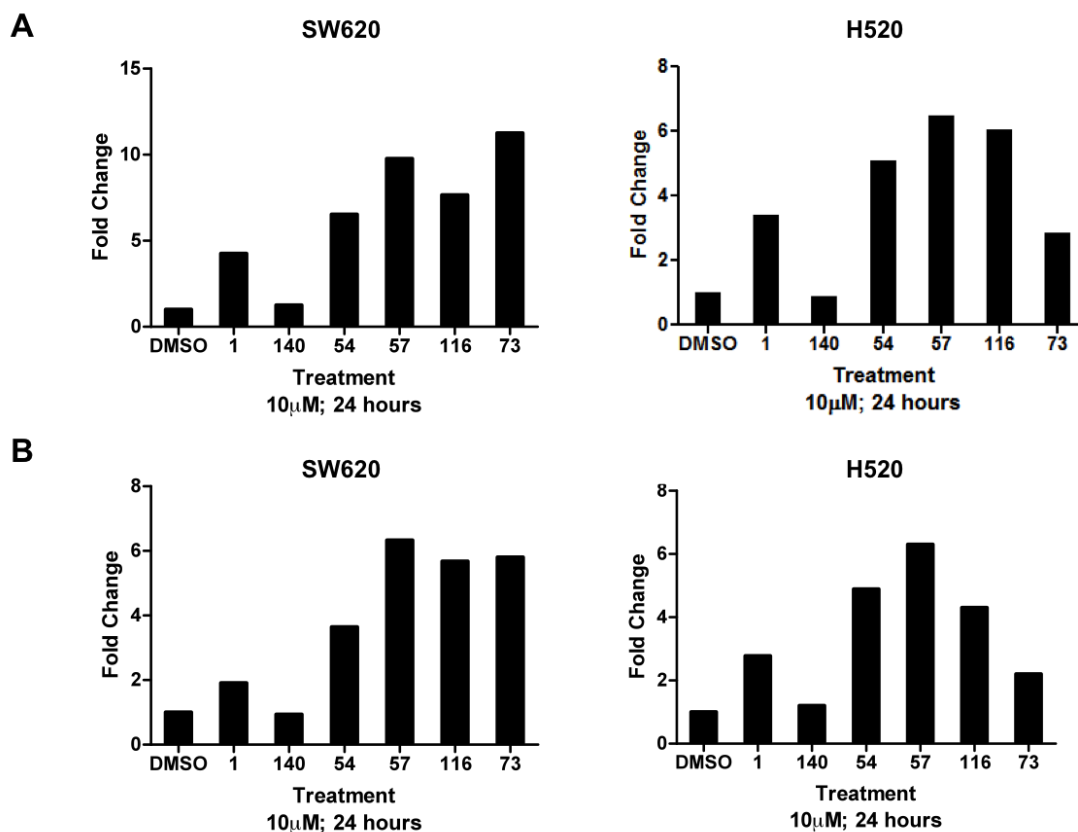
Figure 3. Western blot (A) and ICW (B) data for profiled compounds. Cells were treated for 24 hours with a single point 10 µM treatment prior to the assay. Samples were probed for E-cadherin and Tubulin (control) and secondary antibody fluorescence intensity was measured using the Odyssey Imaging System. Fold change was calculated by normalizing all samples to the DMSO control at 1.

The ICW assay appeared to trend with the western blot assay data, confirming that it could be utilized for the development of concentration response curves (CRC) based on E-cadherin restoration (**Figure 3**). Representative curves can be seen in Figure 4 for compound **1**, compound **57**, and compound **289** in both the SW620 and H520 cells. All EC_{50} values for selected active analogs screened can be seen in Appendix 1, Table 21. The EC_{50} values for **1**, **57**, and **293** in the SW620 cell line were 10.6 µM, 2.13 µM, and 1.1 µM, and 5.3 µM, 1.25 µM, and 1.0 µM in the H520 cell line, respectively (**Figure 4**).
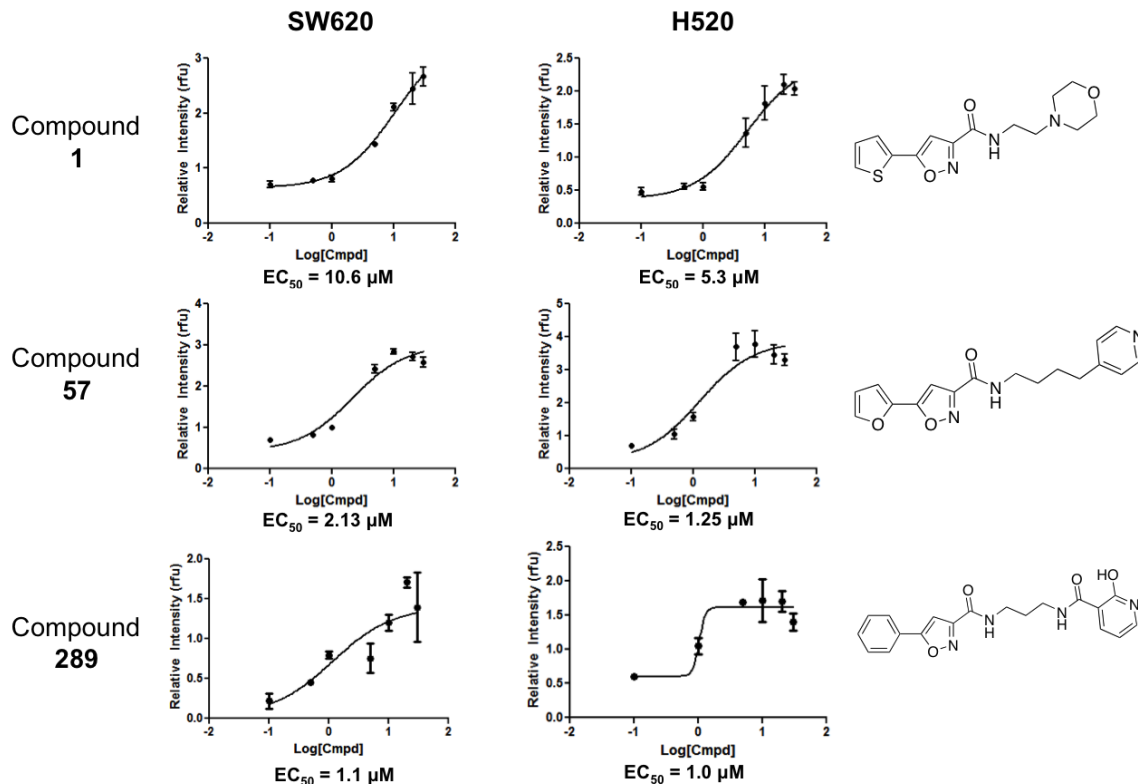
Figure 4. Concentration response curves for selected analogs. SW620 and H520 cells were treated for 24 hours with compound. Concentration response curves were developed from 7 treatment concentrations from 30 µM to 0.1 µM using the In-Cell Western assay. Each concentration was tested in triplicate and plotted as the relative intensity of fluorescence; EC_{50} values were determined and are shown below each graph. The chemical structure is also shown to the right side of the figure for each profiled compound to display structural variability.

While we were able to quantify that a subset of selected compound **1** analogs were restoring E-cadherin, we also wanted to determine the localization of the E-cadherin protein within the cell. For this, we used immunofluorescent microscopy to visualize the localization of E-cadherin and β-catenin in SW620 cells after treatment with the profiled compounds or DMSO (**Figure 5**). E-cadherin is not present and β-catenin appears dispersed throughout the cytoplasm in the DMSO control and compound **140**, an inactive analog, treated cells. β-catenin dispersal throughout the cytoplasm was suspected in such treatments since E-cadherin was not present to sequester β-catenin to the member via

79

CCC formation. In Addition, the cells treated with active analogs displayed localization of both E-cadherin and β-catenin to the membrane, especially where cell-to-cell contact was made. If E-cadherin was being restored to the membrane, as was observed via immunofluorescent microscopy, it is logical to hypothesize that β-catentin would be sequestered back to the membrane. Co-localization of these proteins can be seen when the images are merged. This suggests that our compounds are not only capable of restoring E-cadherin protein expression, but E-cadherin is a also successfully being transported to the membrane. Additionally, the localization of E-cadherin and β-catenin to the membrane suggests that the E-cadherin cytoplasmic cell-adhesion complex is being restored at the membrane on the intracellular portion of the E-cadherin protein.
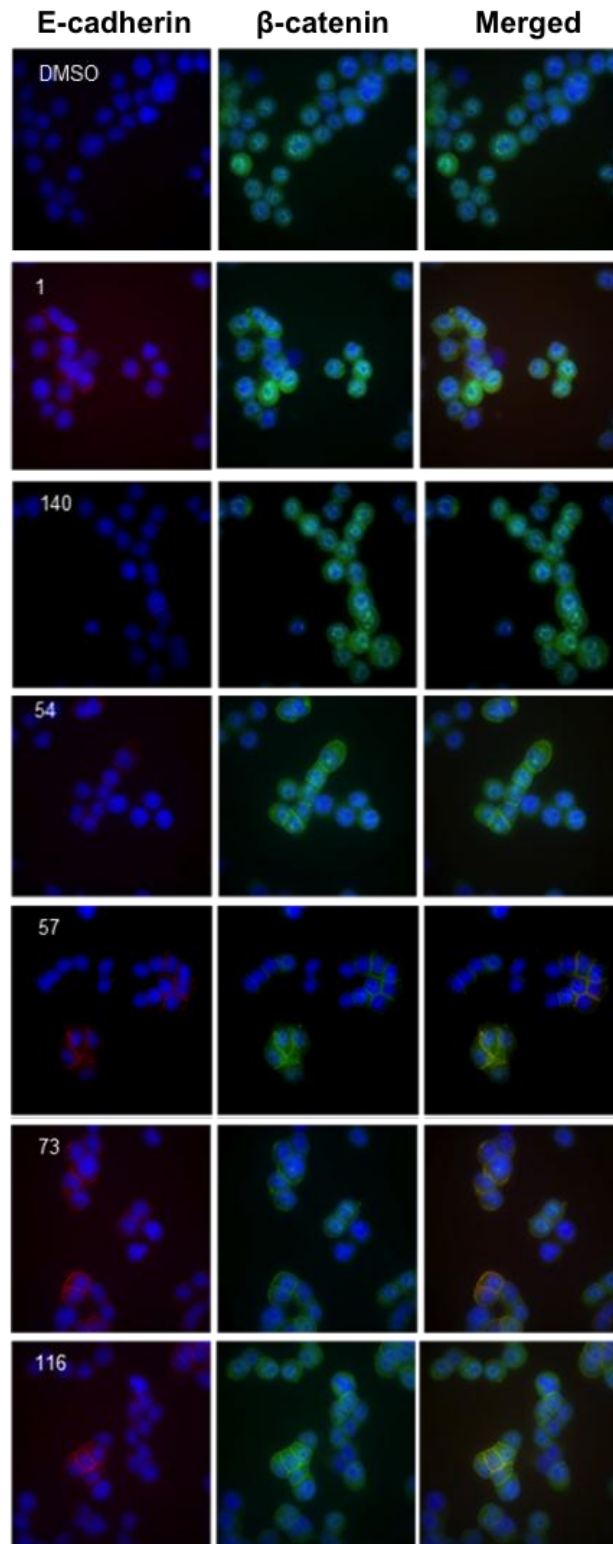
Figure 5. Visualization of E-Cadherin localization. SW620 cells were treated for 24 hours with a 10 µM concentration of compound. Each panel of three images represents a single treatment with immunofluorescent localization of E-cadherin or β-catenin as well as the merged image. Cells were viewed with 600x total magnification.

Having shown that the E-cadherin protein is being restored (**Figure 3**) and that the protein is localizing to the membrane in cells treated with active compounds (**Figure 5**), we wanted to identify if the restoration was occurring at the transcriptional level. For this we used primers for the E-cadherin mRNA transcript and PCR to determine if there were changes in mRNA levels after treatment of SW620 cells with DMSO, compound **1**, compound **57**, or compound **140**. We saw that treatment with compound **1** and **57** resulted in a modest increase in E-cadherin mRNA transcripts as compared to the DMSO and **140** treated samples (data not shown). In extension of this finding we decided to send 3 samples to HudsonAlpha Institute for Biotechnology for RNA-Seq analysis. RNA-Seq analysis, much like PCR, involves reverse transcription of mRNA samples into cDNA, which is then sequenced. This provides a very detailed view of both the exact sequence of the genes expressed and the magnitude of the expression. An example can be seen in Figure 6 in which the *CDH1* gene reads are shown for each of the three treatments. After analysis, cells treated with compound **1** had a 1.46 fold increase in E-cadherin expression and cells treated with compound **57** had a 10.22 fold increase in E-cadherin expression when compared to DMSO treated sample. This data further validates that the compounds are restoring E-cadherin protein by enhancing the transcription of *CHD1* gene.
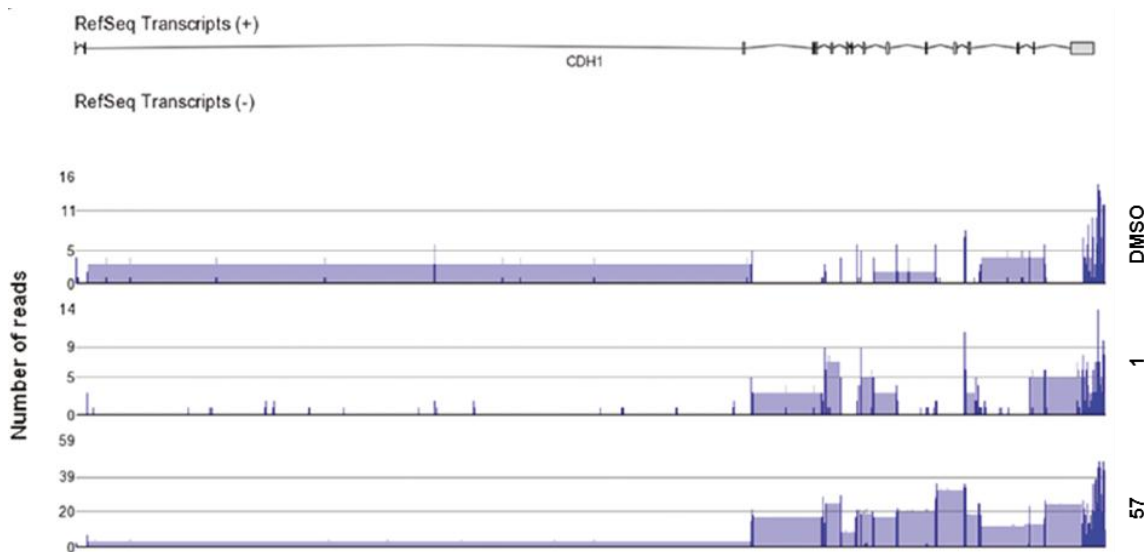
Figure 6. Quantification of mRNA transcript levels after treatment of SW620 cells with selected analogs. RNA-Seq experiments and data analyses were conducted at HudsonAlpha Institute for Biotechnology. The panel shows the *CDH1* (E-cadherin gene) expression for DMSO, compound **1**, and compound **57** treated samples. Reads are in bright blue and the grayish blue boxes between the reads show junctions between exons.

It became apparent within the first initial libraries that a distinct morphological change was occurring in the SW620 cells that were treated with active compounds (**Figure 7A**). Thus, we could identify prior to the Western blot or ICW assay if the compounds were likely to be active or inactive. It became apparent that the compounds that elicited a greater fold increase in E-cadherin expression also had a larger percentage of cells within the well that appeared to have undergone this distinct morphology change. The SW620 cells typically have a spherical shape, as seen in the DMSO and compound **140** treated cells in Figure 7A. The cells treated with active compounds, as seen in the compound **1** and compound **57** treated cells, appear to have spread out in all directions. Curious about this morphology change, we plated SW620 cells in an E-plate, which allowed us to measure changes in cell index using the xCelligence System by Roche

Diagnostics. Cell index is a reflection of cell attachment, number, and morphology of cells within the well. After a 10 μM treatment we observed the cells for 50 hours; the cells were treated a second time after 24 hours. It can be seen that the cells treated with the active compounds have an increased cell index as compared to cells treated with DMSO or inactive analogs (**Figure 7B**). This suggests that the compounds have improved adherence to the bottom of the E-plate after treatment with the active compounds, which may explain the morphology changes seen in Figure 7A. One could hypothesize that the increase in cell index observed after treatment with active analogs was a result of E-cadherin restoration to the membrane leading to functional assembly of additional adhesion complexes [15, 16]. In addition, this would result in CCC formation on the intracellular portion that mediates the association and stability with the actin cytoskeleton [17-19].
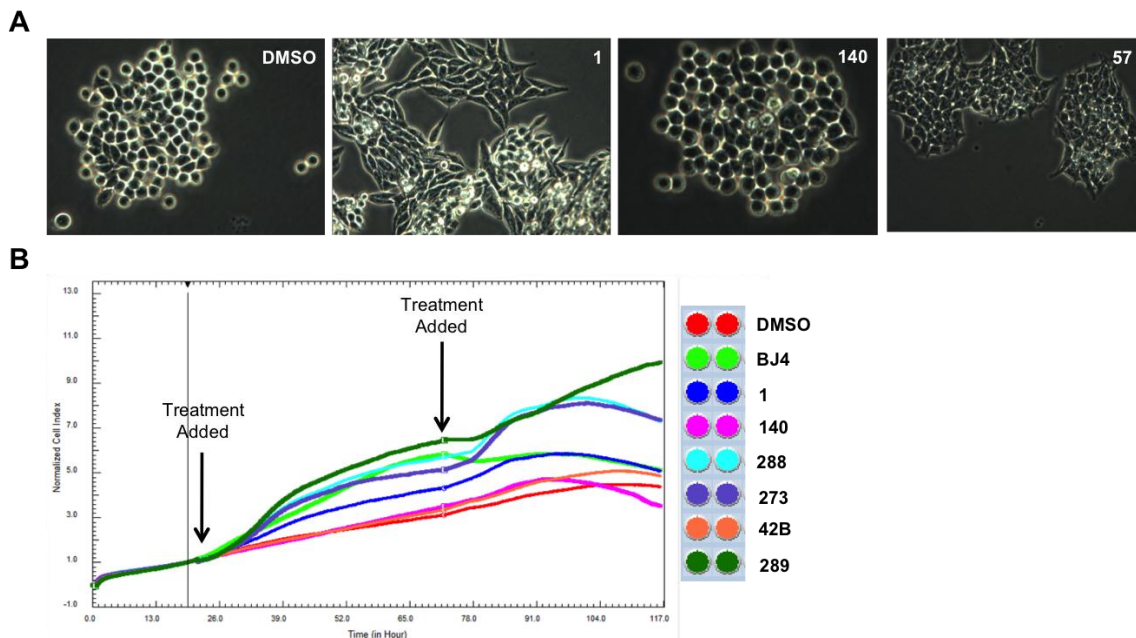
Figure 7. Distinct changes in cell morphology in the SW620 cells were observed after treatment with selected analogs. (A) Pictures acquired at 20x using a basic light microscope (B) Cell Index data collected from the xCelligence machine over 50 hours and legend indicating the 10 µM treatment of selected analogs per colored line.

Cells Remain Viable and Invasion is Inhibited after Treatment by Selected Analogs

Next we wanted to look at viability of the SW620 and H520 cells after 48 hours of treatment with profiled analogs. This was conducted using the WST-1 reagent, which is cleaved in metabolically active cells to produce a soluble substrate whose absorbance can be measured. Overall, the active analogs had no effect on cell viability in either cell line with exception to compound **57**. Compound **57** displayed a moderate decrease of cell viability in both the SW620 and H520 cell lines (**Figure 8A**).

Additionally, we wanted to determine if the compounds were capable of inhibiting cell proliferation, as uncontrolled proliferation is common in carcinogenesis.

For this we utilized a standard bromodeoxyuridine (BrdU) incorporation assay to determine if the compound treatments were inhibiting DNA synthesis. Compound **57** displayed significant inhibition of proliferation, while compounds **116** and **73** displayed moderate inhibition in the SW620 cells. However, only **57** and **73** showed moderate inhibition in the H520 cells (**Figure 8B**). Overall, compound **57** is the only profiled analog that displays significant effects by decreasing both cell viability and proliferation in both the SW620 and H520 cell lines.
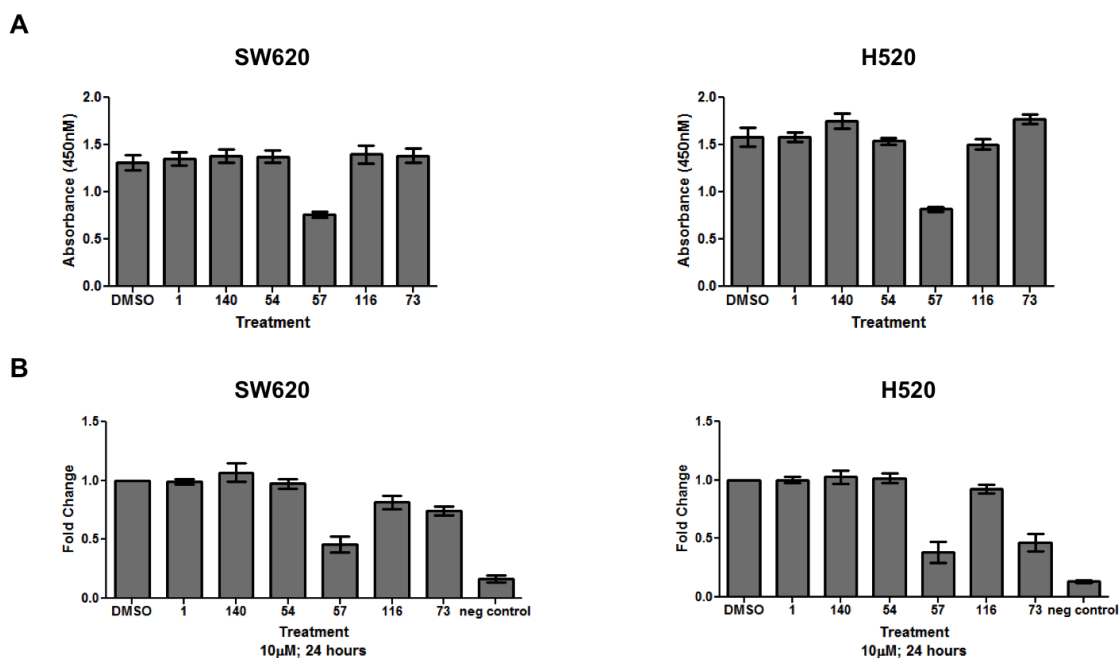
Figure 8. Compounds have minimal effect on viability (A) and proliferation (B). (A) In the viability assay, SW620 or H520 cells were treated for 48 hours with a 10 µM concentration of compound then analyzed by the WST-1 assay. (B) In the proliferation assay, SW620 or H520 cells were treated for 48 hours with a 10 µM concentration of compound then analyzed by the BrdU incorporation assay. For both experiments each treatment was performed in triplicate, and absorbance was measured at 450 nm.

In addition, we wanted to confirm that the compounds were not cytotoxic to a normal epithelial cell line. We screened the MCF10A cell line, a normal-like mammary epithelial cell line, in both the viability (WST-1) and proliferation (Brd-U) assays after treatment with the profiled compounds for 48 hours (**Figure 9**). There were no significant decreases in viability or proliferation after treatment with active compounds, suggesting that these compounds are not cytotoxic to normal epithelial cells.
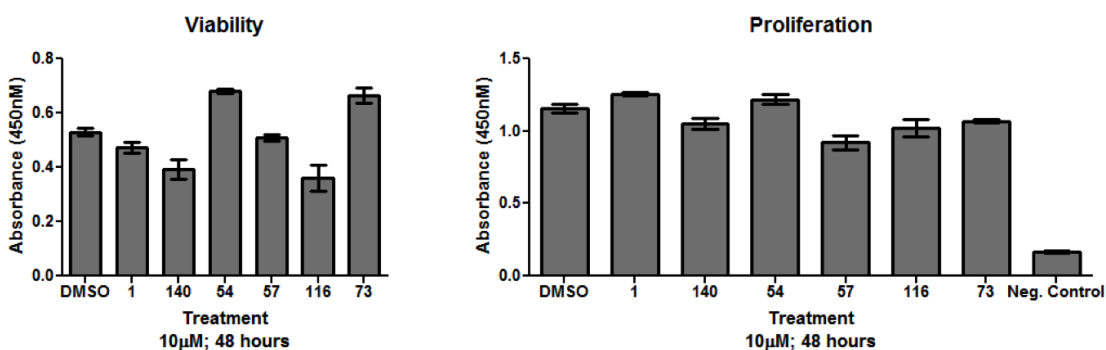
Figure 9. Compounds are not cytotoxic to MCF10A cells – a normal-like human mammary epithelial cell line. (A) In the viability assay, cells were treated for 48 hours with a 10 µM concentration of compound then analyzed by the WST-1 assay. (B) In the proliferation assay, cells were treated for 48 hours with a 10 µM concentration of compound then analyzed by the BrdU Incorporation Assay. For both experiments each treatment was performed in triplicate, and absorbance was measured at 450 nm.

Further, we sent compounds **1** and **57** to be screened in a 60 cancer cell line panel at the National Cancer Institute (NCI). The panel of cancer cell lines represented leukemia, melanoma, lung, colon, brain, ovary, breast, prostate, and kidney cancers. The compounds were screened initially at a single 10 µM concentration in all 60 cancer cell lines to look at growth inhibition. There were only a few cell lines that displayed a

moderate amount of growth inhibition after either treatment. Compound **1** showed moderate growth inhibition against the PC-3, A498, and UACC-257 cell lines (**Table 1**). Compound **57** showed moderate growth inhibition against the RPMI-8226, NCI-H522, HCC-2996, MALME-3M, SK-MEL-28, UACC-257, OVCAR-4, OVCAR-8, NCI/ADR-RES, A498, PC-3, MCF-7, T-47D, and MDB-MB-468 cell lines (**Table 2**). While there was moderate growth inhibition in selected cell lines there wasn't any significant growth inhibition, thus the compounds were not screened in a dose response manner. However, it did provide additional cell lines that could be pursued in the future if necessary in order to validate the mechanism of action.

Table 1. NCI Cytotoxicity Screen for Compound **1**.

Developmental Therapeutics Program	NSC: D-758127 / 1	Conc: 1.00E-5 Molar	Test Date: May 09, 2011
One Dose Mean Graph	Experiment ID: 1105OS49		Report Date: Oct 17, 2011

Panel/Cell Line	Growth Percent	Mean Growth Percent - Growth Percent
Leukemia		
HL-60(TB)	81.37	
K-562	84.07	
MOLT-4	77.64	
RPMI-8226	49.37	
SR	72.18	
Non-Small Cell Lung Cancer		
A549/ATCC	78.45	
EKVX	75.71	
HOP-62	69.84	
HOP-92	69.69	
NCI-H226	70.66	
NCI-H23	79.37	
NCI-H322M	78.62	
NCI-H460	77.60	
NCI-H522	48.02	
Colon Cancer		
COLO 205	69.78	
HCC-2998	55.69	
HCT-116	70.50	
HCT-15	84.21	
HT29	68.71	
KM12	83.88	
SW-620	102.69	
CNS Cancer		
SF-268	85.77	
SF-295	85.36	
SF-539	71.33	
SNB-19	69.38	
U251	77.09	
Melanoma		
LOX IMVI	83.11	
MALME-3M	41.92	
M14	67.64	
MDA-MB-435	68.88	
SK-MEL-28	57.46	
SK-MEL-5	64.20	
UACC-257	34.41	
UACC-62	62.98	
Ovarian Cancer		
IGROV1	90.24	
OVCAR-3	75.00	
OVCAR-4	55.58	
OVCAR-5	64.89	
OVCAR-8	49.23	
NCI/ADR-RES	40.65	
SK-OV-3	86.78	
Renal Cancer		
786-0	94.09	
A498	16.81	
ACHN	84.51	
CAKI-1	84.92	
RXF 393	92.15	
SN12C	74.33	
TK-10	90.30	
UO-31	81.58	
Prostate Cancer		
PC-3	62.95	
DU-145	73.09	
Breast Cancer		
MCF7	56.77	
MDA-MB-231/ATCC	107.08	
HS 578T	75.42	
BT-549	96.37	
T-47D	54.30	
MDA-MB-468	32.74	
Mean	71.18	
Delta	54.37	
Range	90.27	

150   100   50   0   -50   -100   -150

Table 2. NCI Cytotoxicity Screen for Compound **57**.

Developmental Therapeutics Program				
One Dose Mean Graph	NSC: D-758127 / 1	Conc: 1.00E-5 Molar	Test Date: May 09, 2011	
	Experiment ID: 1105OS49		Report Date: Oct 17, 2011	

Panel/Cell Line	Growth Percent	Mean Growth Percent - Growth Percent
Leukemia		
HL-60(TB)	81.37	
K-562	84.07	
MOLT-4	77.64	
RPMI-8226	49.37	
SR	72.18	
Non-Small Cell Lung Cancer		
A549/ATCC	78.45	
EKVX	75.71	
HOP-62	69.84	
HOP-92	69.69	
NCI-H226	70.66	
NCI-H23	79.37	
NCI-H322M	78.62	
NCI-H460	77.60	
NCI-H522	48.02	
Colon Cancer		
COLO 205	69.78	
HCC-2998	55.69	
HCT-116	70.50	
HCT-15	84.21	
HT29	68.71	
KM12	83.88	
SW-620	102.69	
CNS Cancer		
SF-268	85.77	
SF-295	85.36	
SF-539	71.33	
SNB-19	69.38	
U251	77.09	
Melanoma		
LOX IMVI	83.11	
MALME-3M	41.92	
M14	67.64	
MDA-MB-435	68.88	
SK-MEL-28	57.46	
SK-MEL-5	64.20	
UACC-257	34.41	
UACC-62	62.98	
Ovarian Cancer		
IGROV1	90.24	
OVCAR-3	75.00	
OVCAR-4	55.58	
OVCAR-5	64.89	
OVCAR-8	49.23	
NCI/ADR-RES	40.65	
SK-OV-3	86.78	
Renal Cancer		
786-0	94.09	
A498	16.81	
ACHN	84.51	
CAKI-1	84.92	
RXF 393	92.15	
SN12C	74.33	
TK-10	90.30	
UO-31	81.58	
Prostate Cancer		
PC-3	62.95	
DU-145	73.09	
Breast Cancer		
MCF7	56.77	
MDA-MB-231/ATCC	107.08	
HS 578T	75.42	
BT-549	96.37	
T-47D	54.30	
MDA-MB-468	32.74	
Mean	71.18	
Delta	54.37	
Range	90.27	

150   100   50   0   -50   -100   -150

Since E-cadherin expression is down regulated at the invasive front in a variety of cancers, it was important to determine if compound **1** and selected analogs reduced invasion in the SW620 and H520 cell lines [18, 20, 21]. The cells were treated with a 10 μM concentration of profiled compounds for 24 hours and then allowed 72 hours to invade through a matrigel-covered chamber. In our preliminary analyses of invasion we fixed and stained the membranes with 1% crystal violet and counted the cells within a 20x field under the microscope. We saw that compound **1** and the profiled compounds all reduced invasion in both the SW620 and H520 cells (**Figure 10A**).   Specifically, compounds **57** and **73** had the most significant effect on reducing invasion of the SW620 and H520 cells through the matrigel. However, there was some concern as both of these compounds were seen to reduce proliferation in both the SW620 and H520 cells. Therefore, we had to question if the reduction in invasion was a result of reduced cell proliferation or viability, since the cells were allowed to invade for 72 hours, or because the compounds do reduce invasion separate from decreasing cellular proliferation.

To test this, Calcein AM was used to stain the cells on both the top and bottom of the chamber.   A fluoroblok membrane within the chamber inserts allowed for fluorescence to be read on either side of the chamber without interference.   This will account for any inhibition of proliferation, as seen in Figure 8A, as the experiments can be normalized to the number of cells remaining in the top chamber. This also confirms that treatment with the compounds is not cytotoxic as Calcein AM is converted to its fluorescent form (Calcein) by living cells. As seen in Figure 10B the compounds were able to reduce the number of invading cells compared to the DMSO control in the SW620 cells as well as the H520 cells (data not shown). Additionally, the inactive analog, **140**,

did not reduce invasion, as the fold change was similar to the DMSO negative control. This suggests that the decrease in invasion observed with compound treatment alterations is not likely to be due to cytotoxicity or inhibition of proliferation, but a result of molecular effects of the compound on the cells.
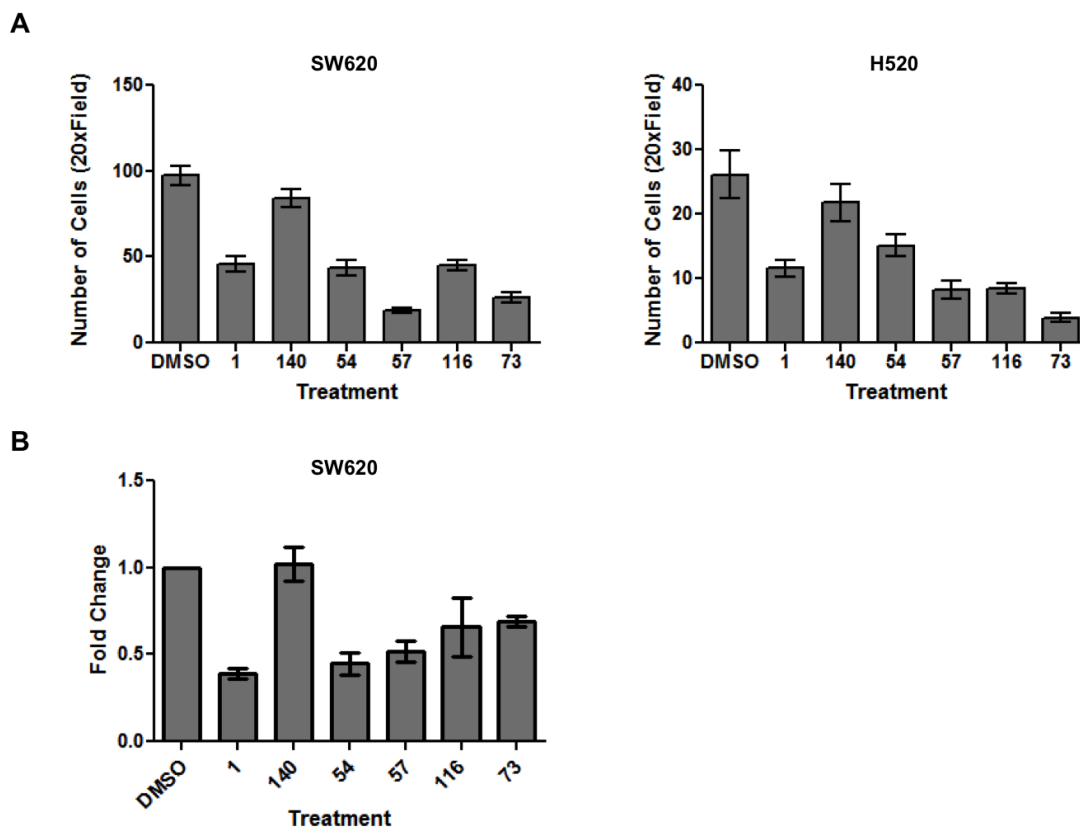

Figure 10. Profiled compounds reduced invasion in both the SW620 and H520 cells. (A) SW620 or H520 cells were treated for 24 hours with a 10 μM concentration of compound and then plated in a matrigel covered Boyden invasion chamber. Cells were given 72 hours to invade, were strained with crystal violet, and then counted. 3 - 20x fields per membrane, 3 membranes per treatment. (B) SW620 cells were treated for 24 hours with a 10 μM concentration of compound and then plated in the matrigel-covered invasion chamber. Cells were given 72 hours to invade, were strained with Calcein AM, and then counted on a fluorometer. A fold change of invading cells was calculated as normalized to the DMSO control.

Increase in Histone Acetylation After Treatment with Active Compounds

We used our knowledge of TSA, a known HDAC inhibitor and our high-throughput screen positive control, to hypothesize that the compounds may be altering histone acetylation and therefore restoring E-cadherin expression. It has been shown that histone deacetylation may lead to transcriptional repression of a gene. More specifically the HDAC1/2-Snail complex has been shown in several cancer models to repress E-cadherin expression directly [14, 22, 23]. Thus, we treated SW620 and H520 cells with selected compounds and used a Western blot assay to probe for Histone H4 pan-acetylation in addition to total H4 histone present in each sample. Figure 11 shows the data expressed as the percent change in histone acetylation above a DMSO control. The cells treated with active compounds shown to restore E-cadherin had a marked increase in histone acetylation compared to the DMSO control and **140**, the negative analog.
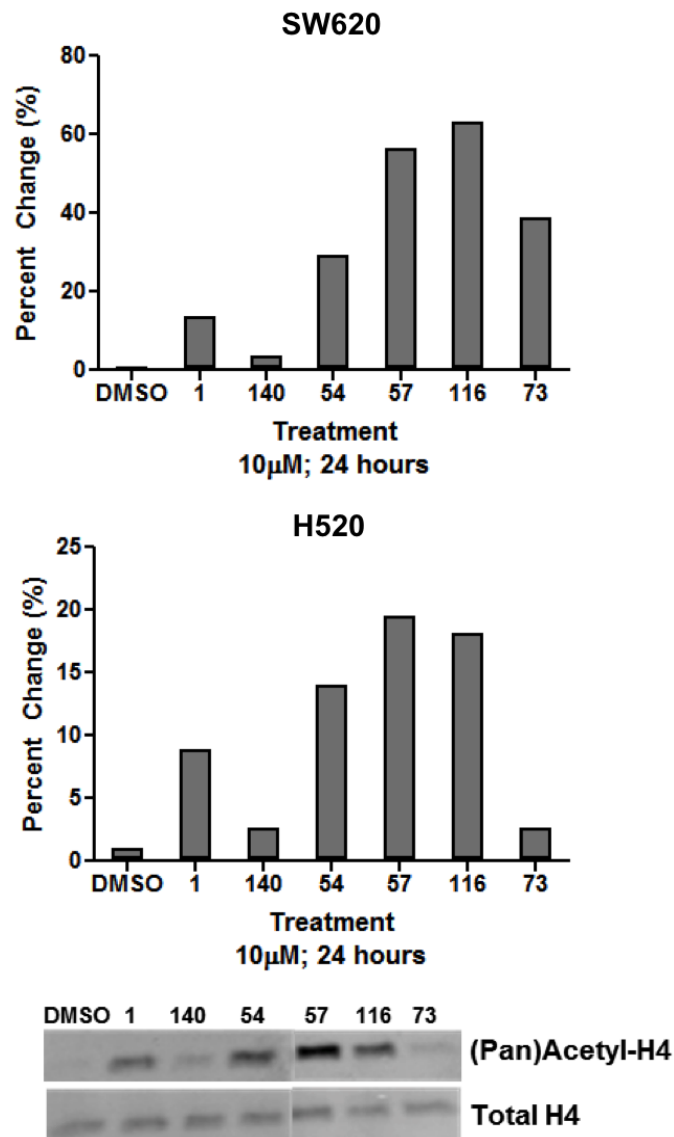
Figure 11. Compounds increased Histone H4 pan-acetylation. SW620 or H520 cells were treated for 24 hours with a 10 μM concentration of compound. Nuclear fractions were isolated and subjected to western blot analysis. Samples were probed for Histone H4 acetylation and total Histone H4. Values were quantified as the fold change in histone acetylation above the DMSO control. The blot was also probed with Parp1/2 and RhoGDI to confirm clean fractionation for each sample (not shown).

Preliminary Efforts to Identify the Molecular Target via Outsourced Screens

Several analogs of compound **1** were profiled in screening assays, both general and specific protein family screens, in an attempt to identify the molecular target. A preliminary screen of one analog, compound **80**, was evaluated in the MDS Pharma Lead Profiling radioligand binding screen, which consists of a panel of 68 GPCRs, ion channels, transporter and nuclear hormone receptors. Compound **80** displayed no significant activity (<50% inhibition at 10 μM), with the lone exception of activity at the imidazoline I_2 central receptor (**Table 3**). The data indicated that the compounds discovered in this project are not broadly promiscuous. A more specific screen was conducted at MDS Pharma Services against a set of potential targets that might be of importance for epigenetic regulation, such as the sirtuins and matrix metalloproteases (MMP). However, no significant activity was found for the compounds at 10 μM against the sirtuins, nine MMP isoforms, catechol-O-methyl transferase, or histamine N-methyltransferase.

Table 3. MDS Pharma Lead Profiling results for Compound **80**.

EXPERIMENTAL RESULTS - BIOCHEMICAL ASSAYS

Cat. #	TARGET	BATCH*	SPP.	n=	CONC.	† % INHIBITION %	IC_{50}	K_I	n_H	R
200510	Adenosine A_1	258235	hum	2	10 µM	16				
200610	Adenosine A_{2A}	258263	hum	2	10 µM	7				
200720	Adenosine A_3	258334	hum	2	10 µM	4				
203100	Adrenergic α_{1A}	258298	rat	2	10 µM	10				
203200	Adrenergic α_{1B}	258558	rat	2	10 µM	3				
203400	Adrenergic α_{1D}	258559	hum	2	10 µM	6				
203620	Adrenergic α_{2A}	258300	hum	2	10 µM	7				
204010	Adrenergic β_1	258301	hum	2	10 µM	4				
204110	Adrenergic β_2	258302	hum	2	10 µM	-3				
285010	Androgen (Testosterone) AR	258508	rat	2	10 µM	3				
212510	Bradykinin B_1	258392	hum	2	10 µM	-15				
212610	Bradykinin B_2	258267	hum	2	10 µM	-5				
214510	Calcium Channel L-Type, Benzothiazepine	258240	rat	2	10 µM	5				
214600	Calcium Channel L-Type, Dihydropyridine	258239	rat	2	10 µM	-7				
216000	Calcium Channel N-Type	258433	rat	2	10 µM	4				
219500	Dopamine D_1	258244	hum	2	10 µM	-1				
219700	Dopamine D_{2S}	258246	hum	2	10 µM	-6				
219800	Dopamine D_3	258247	hum	2	10 µM	12				
219900	Dopamine $D_{4.2}$	258248	hum	2	10 µM	11				
224010	Endothelin ET_A	258250	hum	2	10 µM	2				
224110	Endothelin ET_B	258251	hum	2	10 µM	11				
225510	Epidermal Growth Factor (EGF)	258253	hum	2	10 µM	2				
226010	Estrogen ERα	258254	hum	2	10 µM	2				
226300	G Protein-Coupled Receptor GPR103	258479	hum	2	10 µM	-9				
226600	$GABA_A$, Flunitrazepam, Central	258304	rat	2	10 µM	14				
226500	$GABA_A$, Muscimol, Central	258303	rat	2	10 µM	-1				
228610	$GABA_{B1A}$	258255	hum	2	10 µM	-4				
232020	Glucocorticoid	258256	hum	2	10 µM	-6				
232700	Glutamate, Kainate	258451	rat	2	10 µM	21				
232810	Glutamate, NMDA, Agonism	258453	rat	2	10 µM	14				
232910	Glutamate, NMDA, Glycine	258455	rat	2	10 µM	-9				

* 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=Additional Comments
ham=hamster; hum=human

96

EXPERIMENTAL RESULTS - BIOCHEMICAL ASSAYS

Cat. #	TARGET	BATCH*	SPP.	n=	CONC.	† % INHIBITION %	IC$_{50}$	K$_I$	n$_H$	R
233000	Glutamate, NMDA, Phencyclidine	258306	rat	2	10 µM	0				
239610	Histamine H$_1$	258307	hum	2	10 µM	-6				
239710	Histamine H$_2$	258389	hum	2	10 µM	-4				
239810	Histamine H$_3$	258296	hum	2	10 µM	-18				
♦ 241000	Imidazoline I$_2$, Central	258297	rat	2	10 µM	54				
243520	Interleukin IL-1	258260	mouse	2	10 µM	4				
250460	Leukotriene, Cysteinyl CysLT$_1$	258261	hum	2	10 µM	-3				
251600	Melatonin MT$_1$	258387	hum	2	10 µM	-12				
252610	Muscarinic M$_1$	258547	hum	2	10 µM	8				
252710	Muscarinic M$_2$	258308	hum	2	10 µM	0				
252810	Muscarinic M$_3$	258309	hum	2	10 µM	3				
257010	Neuropeptide Y Y$_1$	258369	hum	2	10 µM	1				
257110	Neuropeptide Y Y$_2$	258370	hum	2	10 µM	1				
258590	Nicotinic Acetylcholine	258310	hum	2	10 µM	-10				
258700	Nicotinic Acetylcholine D1, Bungarotoxin	258311	hum	2	10 µM	6				
260110	Opiate δ (OP1, DOP)	258585	hum	2	10 µM	30				
260210	Opiate κ (OP2, KOP)	258266	hum	2	10 µM	-11				
260410	Opiate µ (OP3, MOP)	258312	hum	2	10 µM	-2				
264500	Phorbol Ester	258313	mouse	2	10 µM	0				
265010	Platelet Activating Factor (PAF)	258314	hum	2	10 µM	-4				
265600	Potassium Channel [K$_{ATP}$]	258315	ham	2	10 µM	-2				
265900	Potassium Channel hERG	258316	hum	2	10 µM	27				
268420	Prostanoid EP$_4$	258317	hum	2	10 µM	1				
268700	Purinergic P$_{2X}$	258372	rabbit	2	10 µM	-24				
268810	Purinergic P$_{2Y}$	258482	rat	2	10 µM	7				
270000	Rolipram	258318	rat	2	10 µM	-9				
271110	Serotonin (5-Hydroxytryptamine) 5-HT$_{1A}$	258320	hum	2	10 µM	5				
271910	Serotonin (5-Hydroxytryptamine) 5-HT$_3$	258321	hum	2	10 µM	6				
278110	Sigma σ$_1$	258324	hum	2	10 µM	46				
278200	Sigma σ$_2$	258325	rat	2	10 µM	23				

* 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=Additional Comments
ham=hamster; hum=human

97

EXPERIMENTAL RESULTS - BIOCHEMICAL ASSAYS

Cat. #	TARGET	BATCH*	SPP.	n=	CONC.	† % INHIBITION %	IC_{50}	K_i	n_H	R
279510	Sodium Channel, Site 2	258326	rat	2	10 µM	29				
255510	Tachykinin NK₁	258384	hum	2	10 µM	5				
285900	Thyroid Hormone	258480	rat	2	10 µM	5				
220320	Transporter, Dopamine (DAT)	258243	hum	2	10 µM	6				
226400	Transporter, GABA	258445	rat	2	10 µM	6				
204410	Transporter, Norepinephrine (NET)	258238	hum	2	10 µM	-4				
274030	Transporter, Serotonin (5-Hydroxytryptamine) (SERT)	258383	hum	2	10 µM	2				

* 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=Additional Comments
ham=hamster; hum=human

Based on published work concerning the involvement of HDAC's in E-cadherin repression and cancer and our visualization of increased histone acetylation by treatment with our compounds, we elected to test two compounds, **32** and **70**, for direct inhibition of HDAC isoforms, using commercially available assays measuring direct inhibition of HDACs 1-11 at Research Biology Corporation. Neither analog showed direct inhibition against any of the HDAC proteins in the screen. From this we were able to conclude that while our compounds alter histone acetylation, which may explain the restoration of E-cadherin expression, unlikely to be a result of the direct inhibition of the HDAC proteins. However, this does leave a variety of potential molecular targets such as histone acetyltransferase (HAT) as well as repressor transcription factors, such as Snail, which are known to complex with HDACs.

98

Discussion

Here we report that the restoration of E-cadherin protein within the cells as a consequence of cell treatment with active analogs (**54**, **57**, **73**, **116**) is being localized to the membrane. Co-localization with β-catenin at the membrane, as seen in the immunofluorescent microscopy analysis, would suggest that the E-cadherin complex is being restored at the membrane (although complex function was not tested). These data further validate the success of the phenotypic high-throughput screen that was utilized to identify compound **1**. In addition, a distinct morphological change was observed in the SW620 cells after treatment with active compounds during the western blot and ICW screening assays. The cells appear to be changing from a small spherical morphology to a more flattened and spread out morphology. This information, coupled with the increase in cell index observed in cells treated with active compounds, suggests that the restoration in E-cadherin may be playing a role in the cells flattening out and adhering better to the bottom of the plate or well.

Additionally, we undertook the biological evaluation of compound **1** and selected analogs in a variety of assays in an effort to better understand how these compounds are affecting the function of the cell aside from restoring E-cadherin expression. The profiled compounds had minimal effect on viability and proliferation, with exception to compound **57**, in both the SW620 and H520 cells. Unfortunately, further analysis of two compounds in a 60 cancer cell line panel at NCI did not identify any specific cancer types that were sensitive to a single point 10 μM treatment. While this is disappointing, it still leaves the possibility for the compounds being used in combination with more aggressive

treatments to induce a synergistic effect. This would ideally allow for a sub-therapeutic dose of a more cytotoxic therapy (with severe side effects), allowing for a more enhanced targeted response at the tumor, but fewer debilitating side effects.

Finally, in regards to our preliminary efforts to identify the molecular target, we received a lot of negative data. Due to the small size and low molecular weight of these molecules our first thought was that they would be extremely promiscuous and bind to a wide array of intracellular proteins and membrane channels and receptors rendering the molecular target impossible to elucidate. However, from the MDS Pharma general radioligand panel screen we learned that these small molecules are very clean. Many of the targets within the MDS Pharma general panel screen elicit adverse events and are often avoided during drug development efforts. Therefore, it is excellent that our molecules are clean both from target identification and drug development stand points.

Additionally, it was surprising to learn that the compounds were not HDAC inhibitors. Due to known information of HDAC inhibitors restoring E-cadherin expression (ie. TSA, our positive control) and similarities between our analogs and known HDAC inhibitors there was high hope that a screen of HDAC isoforms would identify at least one target. Further, HDAC inhibitors have been an area of focus in cancer therapeutic development, and specifically around developing isoform specific inhibitors. Identifying a selective HDAC inhibitor would have quickly moved the project forward; although now we believe we may have a potentially novel target and mechanism of action. Therefore, this data led us to believe that there may be a single molecular target, and that understanding the mechanism of action and ultimately identifying the molecular target were the next questions to address.

References

1. Christofori, G. and H. Semb, *The role of the cell-adhesion molecule E-cadherin as a tumour-suppressor gene.* Trends Biochem Sci, 1999. **24**(2): p. 73-6.

2. Cavallaro, U. and G. Christofori, *Cell adhesion and signalling by cadherins and Ig-CAMs in cancer.* Nat Rev Cancer, 2004. **4**(2): p. 118-32.

3. Steeg, P.S., *Metastasis suppressors alter the signal transduction of cancer cells.* Nat Rev Cancer, 2003. **3**(1): p. 55-63.

4. Miyoshi, J. and Y. Takai, *Structural and functional associations of apical junctions with cytoskeleton.* Biochim Biophys Acta, 2008. **1778**(3): p. 670-91.

5. Voulgari, A. and A. Pintzas, *Epithelial-mesenchymal transition in cancer metastasis: mechanisms, markers and strategies to overcome drug resistance in the clinic.* Biochim Biophys Acta, 2009. **1796**(2): p. 75-90.

6. Batlle, E., et al., *The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells.* Nat Cell Biol, 2000. **2**(2): p. 84-9.

7. Cano, A., et al., *The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression.* Nat Cell Biol, 2000. **2**(2): p. 76-83.

8. Comijn, J., et al., *The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion.* Mol Cell, 2001. **7**(6): p. 1267-78.

9. Hajra, K.M., D.Y. Chen, and E.R. Fearon, *The SLUG zinc-finger protein represses E-cadherin in breast cancer.* Cancer Res, 2002. **62**(6): p. 1613-8.

10. Perez-Moreno, M.A., et al., *A new role for E12/E47 in the repression of E-cadherin expression and epithelial-mesenchymal transitions.* J Biol Chem, 2001. **276**(29): p. 27424-31.

11. Hennig, G., et al., *Progression of carcinoma cells is associated with alterations in chromatin structure and factor binding at the E-cadherin promoter in vivo.* Oncogene, 1995. **11**(3): p. 475-84.

12. Yoshiura, K., et al., *Silencing of the E-cadherin invasion-suppressor gene by CpG methylation in human carcinomas.* Proc Natl Acad Sci U S A, 1995. **92**(16): p. 7416-9.

13. Yoo, C.B. and P.A. Jones, *Epigenetic therapy of cancer: past, present and future.* Nat Rev Drug Discov, 2006. **5**(1): p. 37-50.

14. von Burstin, J., et al., *E-cadherin regulates metastasis of pancreatic cancer in vivo and is suppressed by a SNAIL/HDAC1/HDAC2 repressor complex.* Gastroenterology, 2009. **137**(1): p. 361-71, 371 e1-5.

15. Tunggal, J.A., et al., *E-cadherin is essential for in vivo epidermal barrier function by regulating tight junctions.* EMBO J, 2005. **24**(6): p. 1146-56.

16. Lewis, J.E., et al., *Cross-talk between adherens junctions and desmosomes depends on plakoglobin.* J Cell Biol, 1997. **136**(4): p. 919-34.

17. Kemler, R., *From cadherins to catenins: cytoplasmic protein interactions and regulation of cell adhesion.* Trends Genet, 1993. **9**(9): p. 317-21.

18. van Roy, F. and G. Berx, *The cell-cell adhesion molecule E-cadherin.* Cell Mol Life Sci, 2008. **65**(23): p. 3756-88.

19. Gumbiner, B.M., *Cell adhesion: the molecular basis of tissue architecture and morphogenesis.* Cell, 1996. **84**(3): p. 345-57.

20. Birchmeier, W. and J. Behrens, *Cadherin expression in carcinomas: role in the formation of cell junctions and the prevention of invasiveness.* Biochim Biophys Acta, 1994. **1198**(1): p. 11-26.

21. Hirohashi, S., *Inactivation of the E-cadherin-mediated cell adhesion system in human cancers.* Am J Pathol, 1998. **153**(2): p. 333-9.

22. Onder, T.T., et al., *Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways.* Cancer Res, 2008. **68**(10): p. 3645-54.

23. Bolos, V., et al., *The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors.* J Cell Sci, 2003. **116**(Pt 3): p. 499-511.

CHAPTER IV

UNDERSTANDING THE MECHANISM OF ACTION OF SMALL MOLECULES THAT RESTORE E-CADHERIN EXPRESSION

Introduction

The major disadvantage to running a high-throughput screen using a phenotypic assay to identify lead compounds, is that the mechanism of action, and more specifically the molecular target, remain unknown. As noted in Chapter 3, we were able to synthesize small molecules that restored E-cadherin expression in both a colorectal and lung cancer cell lines, which had a repressed but functional *CDH1* gene. In Chapter 2, it was noted that the SAR surrounding the original screening hit, compound **1**, was shallow with activity lost from subtle changes made to the molecule. While in more recently synthesized libraries, we were able to install a carbon linker tethered by amines on either side allowing for the molecule to be homologated with good activity; we are hindered by not knowing the molecule target in more aggressively pursing SAR.

Loss of E-cadherin protein expression is frequently found during tumor progression and has been identified as a clinical marker for poor prognosis in some cancers [1-3]. E-cadherin expression can be regulated at the transcriptional level as well as the post-translational level. Based on preliminary data, histone acetylation, and RNASeq data provided in Chapter 3 we focused our attention on regulation of E-cadherin expression at the transcriptional level.

Silencing of E-cadherin at the transcriptional level is a result of repressor transcription factor binding complexes that bind to the promoter region of the *CDH1* gene. These repressor transcription factors include Snail1 (Snail), Snail2 (Slug), ZEB1 (δEF1), ZEB2 (Sip1), E47, and Twist [4]. Binding of these repressor transcription factors in complexes leads to epigenetic silencing of the *CDH1* gene by histone modifications and DNA hypermethylation [5-7]. The process of silencing the E-cadherin gene promoter is a complex cascade of events that is still being identified; although there is a dynamic range of repression both reversible and irreversible. Snail is first recruited and forms a complex with HDACs, thereby inducing histone deacetylation, while a second repressor complex is recruited to the site to promote histone methylation. This preliminary repression can induce expression of additional repressor proteins, such as ZEB1, which bind to the *CDH1* promoter further promoting E-cadherin gene silencing [1].

Methods and Materials

Cell Culture

A colorectal adenocarcinoma cell line, SW620, and a squamous cell lung carcinoma cell line, H520, were obtained from the American Type Culture Collection (Manassas, VA) and maintained in a humidified atmosphere of 5% CO_2 in air at 37 °C. The cells were routinely cultured in RPMI 1640 supplemented medium with 10% fetal bovine serum (Atlanta Biologicals, GA) and 100 μg/mL penicillin-streptomycin.

Quantitative (q)PCR Analysis

RNA extraction was performed using the RNEasy Mini Kit (Qiagen Corporation), following the manufacturer's protocol. The Transcriptor Universal cDNA Master kit (Roche Diagnostics) was used for cDNA synthesis. 1.0 μg of RNA per sample was used for each reaction. Samples were incubated at room temperature for 5 minutes, 55 °C for 10 minutes, 85 °C for 5 minutes, on ice for 5 minutes, and then diluted to a total volume of 100 μL. Real-time analysis was performed on the Light Cycler 480 (Roche Diagnostics) with Universal Probe Master system; primers and probes for *CDH1, NFATc1, NFATc2*, and *PMM-1* genes were selected according to the Software Probe Finder (Roche Diagnostics). The protocol consisted of 45 cycles: 10 seconds at 95 °C, 30 seconds at 60 °C, and 1 second at 72 °C.

Table 1. PCR Primers and UPL Probes for each gene analyzed.

Gene	Forward Primer	Reverse Primer	UPL Probe
CDH1	5'- TTG ACG CCG AGA GCT ACA C -3'	5'- GTC GAC CGG TGC AAT CTT -3'	80
NFATc1	5' - CCA AGG TCA TTT TCG TGG AG - 3'	5'- GGT CAG TTT TCG CTT CCA TC -3'	45
NFATc2	5'- CAT CTA ACC CCA TCG AGT GC -3'	5'- GCT GTC TGT GTC TTG TCT TTC AA -3'	44
PMM-1	5'- TTC TCC GAA CTG GAC AAG AAA -3'	5'- CTC TGT TTT CAG GGG TTC CA -3'	7

Plasmid Constructs

The plasmids constructs were graciously provided from two separate labs. The 1.4 kbp E-cadherin promoter plasmid (E1) was synthesized in Eric R Fearon's Laboratory at The University of Michigan [8]. The E-cadherin promoter plasmids E2-E8 which ranged in size from 1.2 kbp to 100bp of the E-cadherin promoter were synthesized in Ju Hshiung Chen Laboratory at Tzu Chi University in Hualien, Taiwan [9]. All plasmid constructs were sequenced at the Vanderbilt University Sequencing Core to confirm the E-cadherin promoter fragments using the PCR primers listed below.

Table 2. PCR primers for Plasmid Construct E2 – E8 Sequencing

E2	-995/135 (forward)	5'-GGTACCGCCGCTCGAGCGAGAGTGCAGTGG-3'
E3	-833/135 (forward)	5'-GGTACCGCCTGTAGTCCCAGCTACTC-3'
E4	-677/135 (forward)	5'-GGTACCAAAAATTAGGCTGCTAGCTCAGTGG-3'
E5	-517/135 (forward)	5'-GGTACCTCTCTCTACAAAAAGGCAAAAGAAAA-3'
E6	-357/135 (forward)	5'-GGTACCGAAAGAGTGAGCCCCATCTCCAAAA-3'
E7	-195/135 (forward)	5'-GGTACCCACCTAGACCCTAGCAACTCCAGGCT-3'
E8	-38/135 (forward)	5'-GGTACCTCCGGGGCTCACCTGGCTGCAGCC-3'
	(Reverse)	5'-AAGCTTCTGCGGCTCCAAGGGCCCATGGCTG-3'

Transient Transfection Experiments

Cells were plated 12–24 hours before transfection at $3\text{-}4\text{x}10^5$ cells per well in a twenty-four well plate. In all, 0.15 µg of various DNA constructs and 0.10 µg of β-galactosidase construct (Promega) were mixed with 75 µL of Effectene (Qiagen). The mixture was incubated at room temperature for 10 min. After washing the cells with 1x PBS, the DNA/Effectene was added to serum free RPMI, transferred into the wells, and then incubated at 37 °C in a 5% CO_2 humidified incubator for 4 hours. After the transfection incubation, the cells were washed with 1x PBS and then treated with 10 µM treatment of DMSO, compound **1**, or selected analogs in RPMI with 10% FBS for 24 hours. Each treatment was conducted in triplicate per plasmid construct. At the end of the 24 hour incubation, the transfected cells were lysed with reporter lysis buffer (Promega).

Luciferase and β-Galactosidase Reporter Assays

The enzymatic activity was measured for firefly luciferase using the Luciferase Assay System (Promega) with a luminometer. All luciferase assays were carried out in triplicate and experiments were carried out at least twice.

The enzymatic activity was measure for β-galactosidase using the β-Galactosidase Enzyme Assay System (Promega) with a spectrophotometer. All β-galactosidase assays were carried out in triplicate and experiments were carried out at least twice.

Evaluation of Increase in E-Cadherin mRNA Transcript Levels

Our preliminary findings using standard PCR (data not shown) as well as the RNASeq analysis data (Chapter 3, **Figure 6**) showed that compound **1** and active analogs increased mRNA transcript levels in the SW620 cells. We wanted to confirm our preliminary findings using RT-qPCR (**Figure 1**). Therefore, SW620 cells were treated with a 10 µM treatment of previously profiled compounds that had been screened in the biological assays presented in Chapter 2. Compounds **1** and **73** had slight increases in mRNA transcription levels relative to DMSO; however, the remaining positive analogs screened showed at least a 10-fold increase in E-cadherin mRNA transcript levels (**Figure 1A**).

A

B
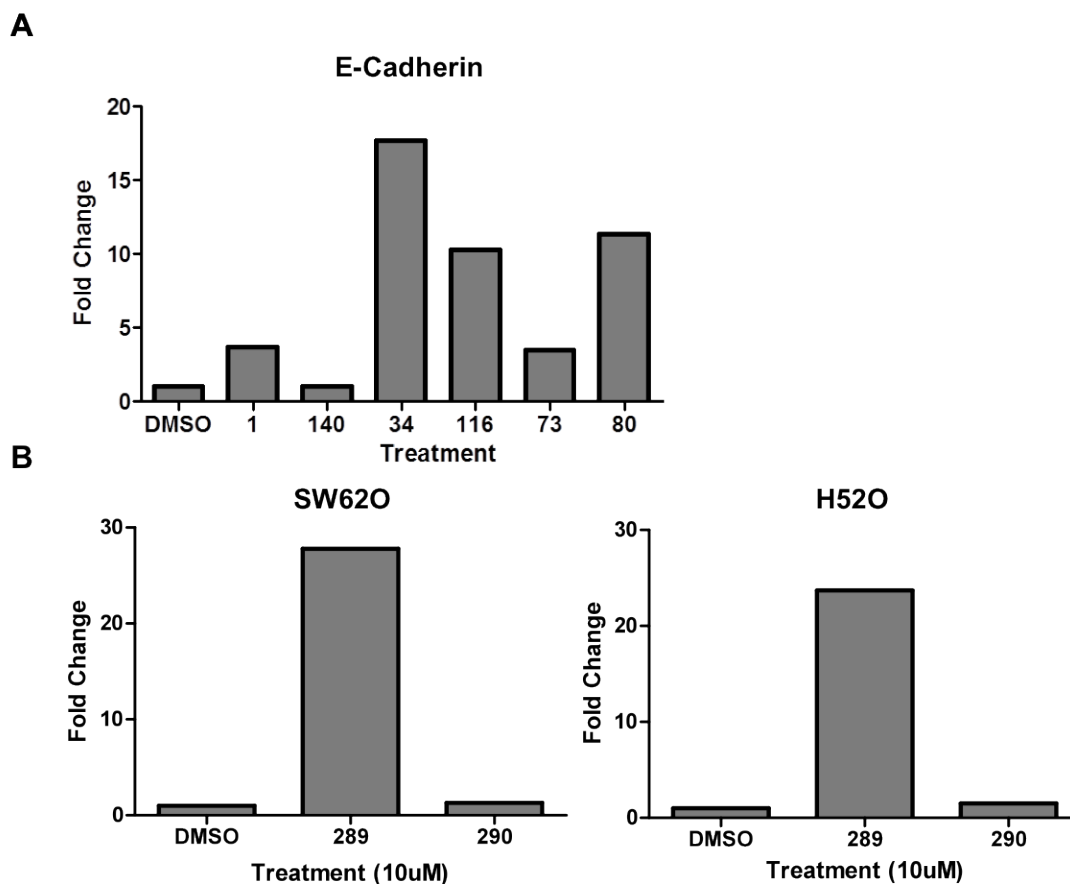
Figure 1. qPCR data for E-cadherin after a 10 µM treatment with selected compounds for 24 hours. (A) SW620 (B) E-cadherin mRNA transcript levels in both cell lines treated with more recently synthesized analogs. All values normalized to the DMSO control and compounds **140** and **290** are inactive analogs.

Additionally, two of the newly synthesized analogs, represented in Figure 2, were analyzed as well in both the SW620 and H520 cells (**Figure 1B**). In both cell lines, the active compound, **289**, had a significant increase in E-cadherin mRNA transcript levels as compared to DMSO and the negative control, compound **290**.

289 **290**
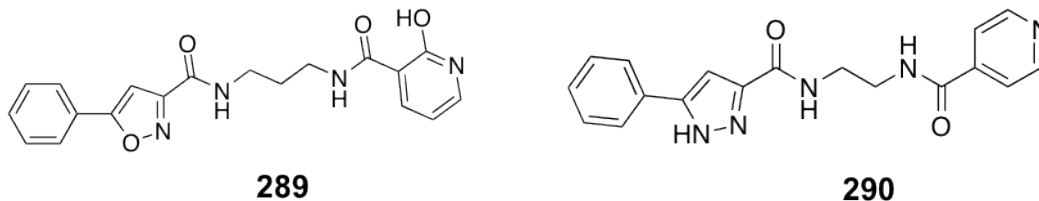
Figure 2. Selected newly synthesized analog (**289**) with improved E-cadherin restoration and an inactive analog (**290**)

We were curious to identify a more specific time point in which activity of the active analogs occurs within the cell. Until this point, we had been using a 24 hour treatment incubation time, which is a standard incubation time. For the time course experiment, RNA was extracted from treated cells at 6 time points: 30 minutes, 1 hour, 3 hours, 6 hours, 12 hours, and 16 hours. A 24 hour time point was not done due to it being the standard incubation time used in all prior experiments and had already been conducted (**Figure 1**). Result for RT-qPCR analysis of E-cadherin mRNA transcript levels at all 6 times points can be seen in Figure 3, in which the data is normalized to the DMSO control samples for each time point.
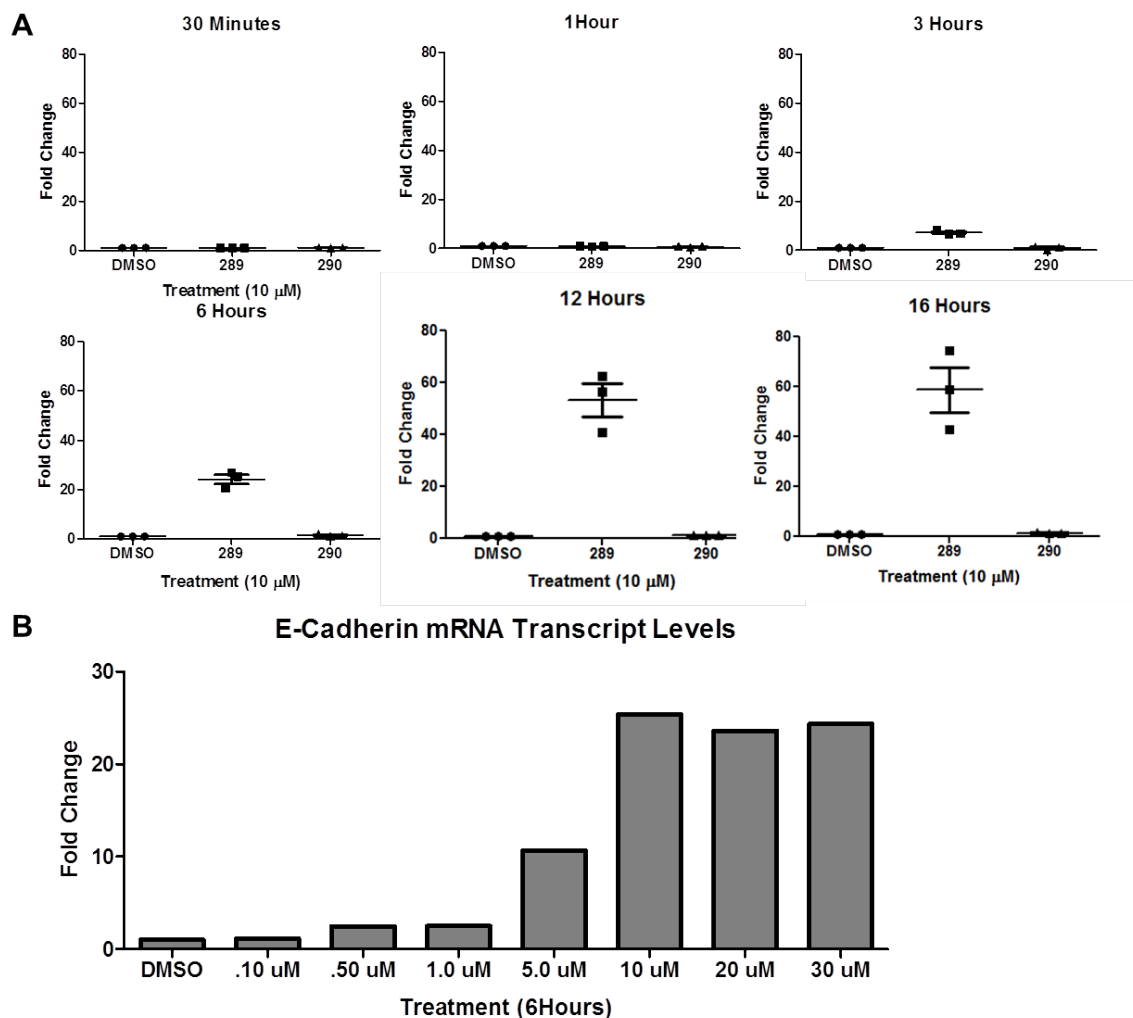
Figure 3. Time course experiment and concentration response curve looking at E-cadherin mRNA transcript levels via qPCR. (A) SW620 cells were treated with a 10 μM concentration of selected analogs or DMSO. RNA was extracted at 6 time points indicated above, and qPCR was run to analyze transcription levels of E-cadherin (N=3). (B) SW620 cells were treated with 7 concentrations between 100 nM and 30 μM for 6 hours, RNA was extracted, and qPCR was turn to analyze transcription levels of E-cadherin. Samples were normalized to DMSO control and compound **290** is an inactive analog.

As can be seen in Figure 3 there is a 7.26-fold increase in E-cadherin mRNA transcript levels after 3 hours of treatment with compound **289**. However, at 6 hours, we see a drastic increase in E-cadherin mRNA transcript levels with treatment of compound **289**, a 24.22-fold increase as compared to DMSO control. We were surprised to see that

111

compound **289** had an effect after only 6 hours, and that levels continued to increase through 16 hours of treatment; 53.40-fold increase after 12 hours and 58.68-fold increase after 16 hours of treatment. I think the most notable bit of information from the time course experiment is the 7.26-fold increase in E-cadherin transcription levels at 3 hours. It would be interesting to further analyze the changes in transcription levels of E-cadherin between 1 – 3 hours to narrow down a point when E-cadherin mRNA transcript levels begin to increase.

Additionally, we developed a concentration response curve for E-cadherin mRNA transcript levels with various concentrations of compound **289** after 6 hours of incubation. 7 concentrations of compound **289** from 100 nM to 30 μM were used to treat the cells for 3 hours, RNA was extracted, and mRNA levels were quantified using qPCR. We saw that E-cadherin mRNA transcript levels begin to increase at 500 nM and plateau at 10 μM (~25 fold increase) as can been seen in Figure 3B. Analyzing E-cadherin mRNA transcript levels at key time points (3 or 6 hours) or across various concentrations may provide additional information when screening newly synthesized libraries.

Increase in Transcript Levels Observed with Additional Transcription Factors

While we were planning to further analyze the RNASeq data in an effort to shed some light on the mechanism of action or identify potential candidates as the molecular target of these small molecules, we were able to use some accidental revelations to direct our analysis towards repressive transcription factors. The family of nuclear factor of activated T cells (NFAT) transcription factors, specifically NFATc1 and NFATc2, are of interest to the research conducted in our lab in regards to invasion and metastasis in colorectal cancer.

In general, research is uncovering that various isoforms of the NFAT transcription factors are functional in tumor cells as well as in the tumor microenvironment, including fibroblasts, endothelial cells, and immune cells and understanding the role for each isoform will be key. For example, NFATc1 and NFATc2 have been shown to have distinct and opposing roles in tumorigenesis. NFATc1 is thought to be an oncogene and constitutive activity studies in fibroblasts have shown that NFATc1 increases proliferation. NFATc2 is thought to be a tumor suppressor protein and constitutive activity studies in fibroblasts have shown that NFATc2 induces cell cycle arrest and apoptosis [10, 11].

Curiously, RNA samples from SW620 and H520 cells treated with DMSO along with a preliminary set of active and inactive analogs were analyzed for effects on NFATc1 and NFATc2 mRNA transcript levels while looking at E-cadherin mRNA transcript levels (data already collected). Originally, the data was looked at separately and it was noticed that an increase in NFATc1 and NFATc2 transcript levels was observed. Although when we compared all three analyses – E-Cadherin, NFATc1, and NFATc2 – we saw a similar trend of increased mRNA transcript levels (**Figure 4**).

We further analyzed this finding by collecting RNA from SW620 and H520 samples treated with DMSO, compound **289**, or compound **290** for 24 hours. The results can be seen in Figure 5, in which a similar increase in all three genes is observed with the active compound, **289**. In the SW620 cells, we see a much smaller increase in NFATc2 transcription levels as compared to E-cadherin and NFATc1, which are relatively similar. However, in the H520 cells the increase in transcription levels appears relatively constant between the three genes analyzed.
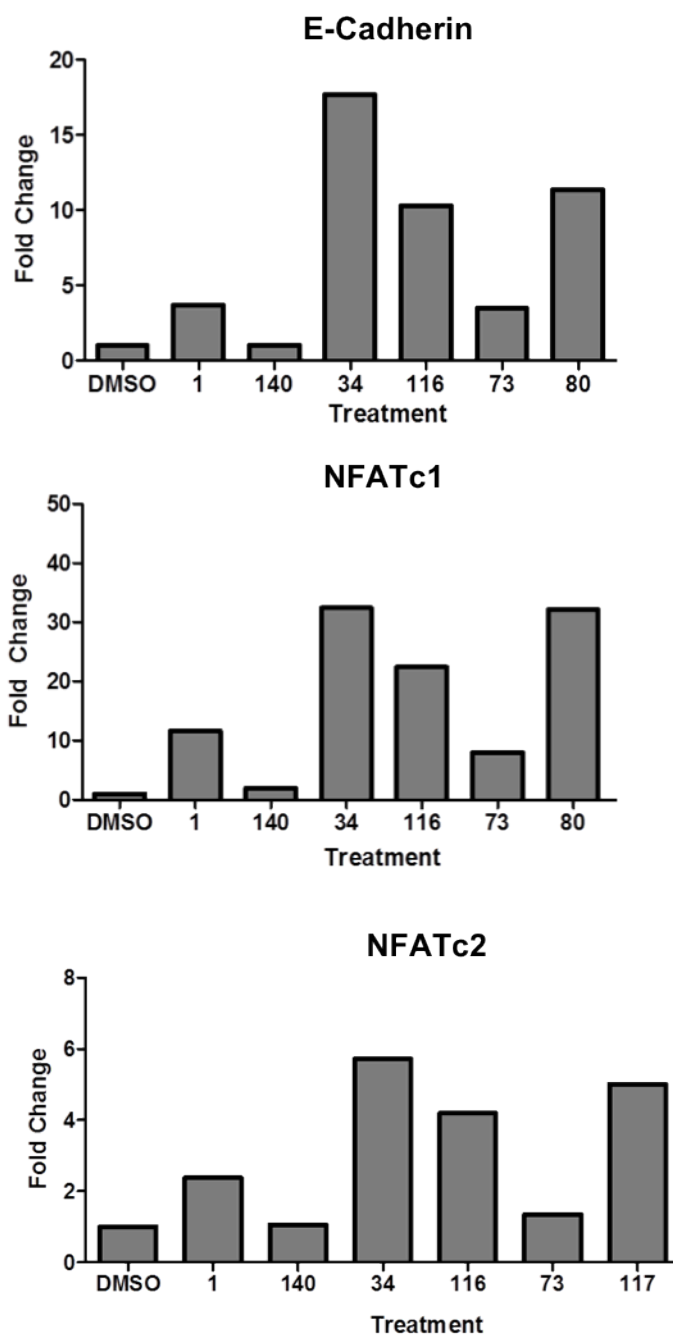
Figure 4. qPCR analysis comparison of E-cadherin, NFATc1, and NFATc2 after SW620 cells were treated with a 10 µM concentration of selected compounds for 24 hours. All values normalized to DMSO control and compound **140** is an inactive analog.
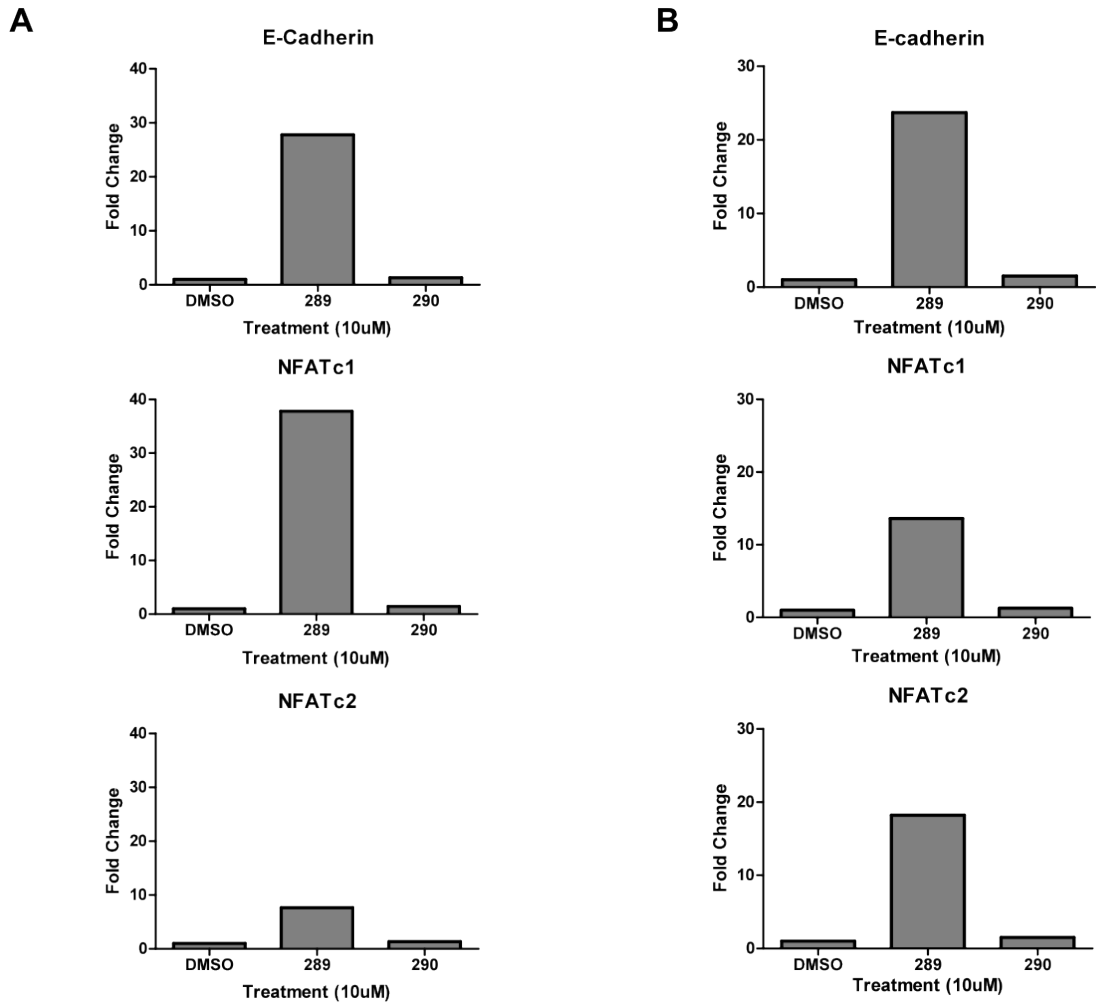
Figure 5. qPCR analysis of E-cadherin, NFATc1, NFATc2 mRNA transcript levels after a 10 µM treatment of selected analogs for 24 hours in the (A) SW620 and (B) H520 cells.

Having observed a 24.22-fold increase in E-cadherin mRNA transcript levels after 6 hours of treatment, we were curious to see the mRNA transcript levels of NFATc1 and NFATc2 after 6 hours of treatment. If the mRNA transcript levels were low then it may suggest that the increase in NFATc1 and NFATc2 transcript levels observed at 24 hours was a downstream effect resulting from E-cadherin transcription. However, if the transcription levels are elevated after 6 hours of treatment, one could hypothesize that the

active analogs are targeting the same mechanism of regulation for all three genes. Thus, NFATc1 and NFATc2 were analyzed in both the SW620 and H520 cells after 6 hours of treatment with DMSO, compound **289**, or compound **290**.

As seen in Figure 6, RNA was extracted from SW620 or H520 cells treated for 6 hours with DMSO, compound **289**, or compound **290** and used to measure E-cadherin, NFATc1, and NFATc2 mRNA transcript levels. In the SW620 cells all 3 genes have increased transcript levels after treatment with compound **289** as compared to DMSO and the inactive analog, compound **290**. When comparing transcription levels with the SW620 cells from Figure 5a and Figure 6a it appears that E-cadherin levels are consistent, NFATc1 levels continue to increase from 6 hours to 24 hours, and NFATc2 levels are much higher after 6 hours of treatment as compared to 24 hours. When comparing mRNA transcript levels from Figure 5b and Figure 6b it appears that the transcription levels remain relatively constant for E-cadherin, NFATc1, and NFATc2 with only a minor decrease occurring at 24 hours. However, what is most important from Figure 6, is that transcription is increased as well for NFATc1 and NFATc2 suggested that the changes in gene transcription must be consistent and a result of treatment by compound **289**.

We used this knowledge to hypothesize that the small molecules are promoting transcription via the same mechanism of action for E-cadherin and the NFAT family of transcription factors. Preliminary analysis from a colleague, Bing Zhang, identified two E-box binding sites within the promoter regions of NFATc1 and NFATc2. Similarly, there are several known E-box binding sites within the promoter region of the *CDH1* gene (**Figure 9**). Thus, further suggesting there must a be constant regulatory

transcription factor protein complex that bindings to conserved binding sites within the promoter regions of these genes.
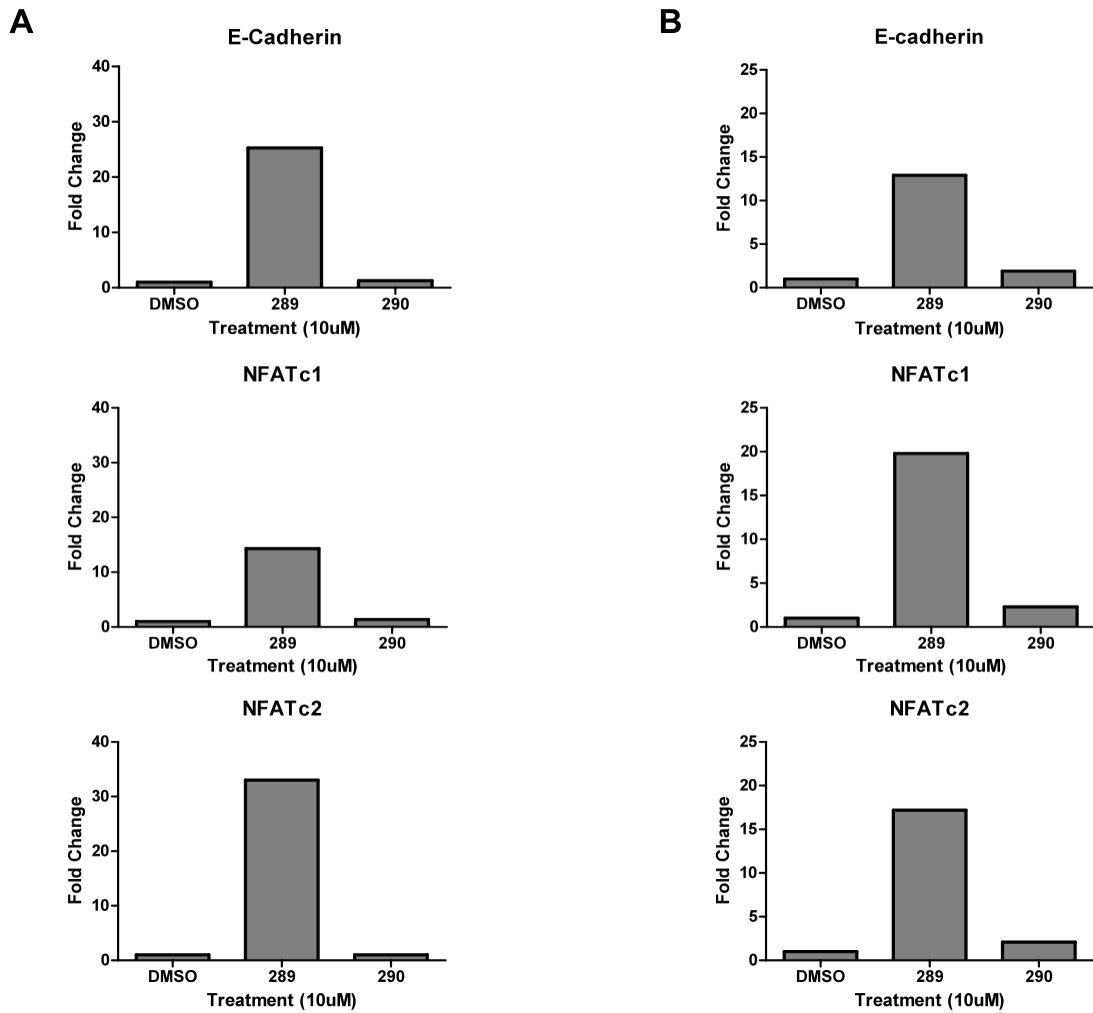
Figure 6. qPCR analysis of E-cadherin, NFATc1, NFATc2 mRNA transcript levels after a 10 µM treatment of selected analogs for 6 hours in the (A) SW620 and (B) H520 cells.

Active Analogs are Believed to be Promoting Transcription by Altering Transcription Factor Binding to E-cadherin Promoter Region

We have been working backwards from the original HTS phenotypic response assay, restoration of E-cadherin protein expression, in an effort to better understand the mechanism of action of these small molecules as well as to move closer towards identifying the molecular target. Thus, observing an increase in E-cadherin mRNA transcript levels after treatment with analogs of compound **1**, we wanted to look specifically at the E-cadherin promoter activity. A reporter plasmid construct obtained from Eric R. Fearon's laboratory, which consisted of a 1.4 kbp fragment of the E-cadherin promoter upstream of the firefly *Luc* gene, was used in preliminary experiments to determine if the small molecules acted within this sequence [8]. A representation of the E-cadherin promoter fragment located within the reporter plasmid construct (E1) can be seen in Figure 7A.
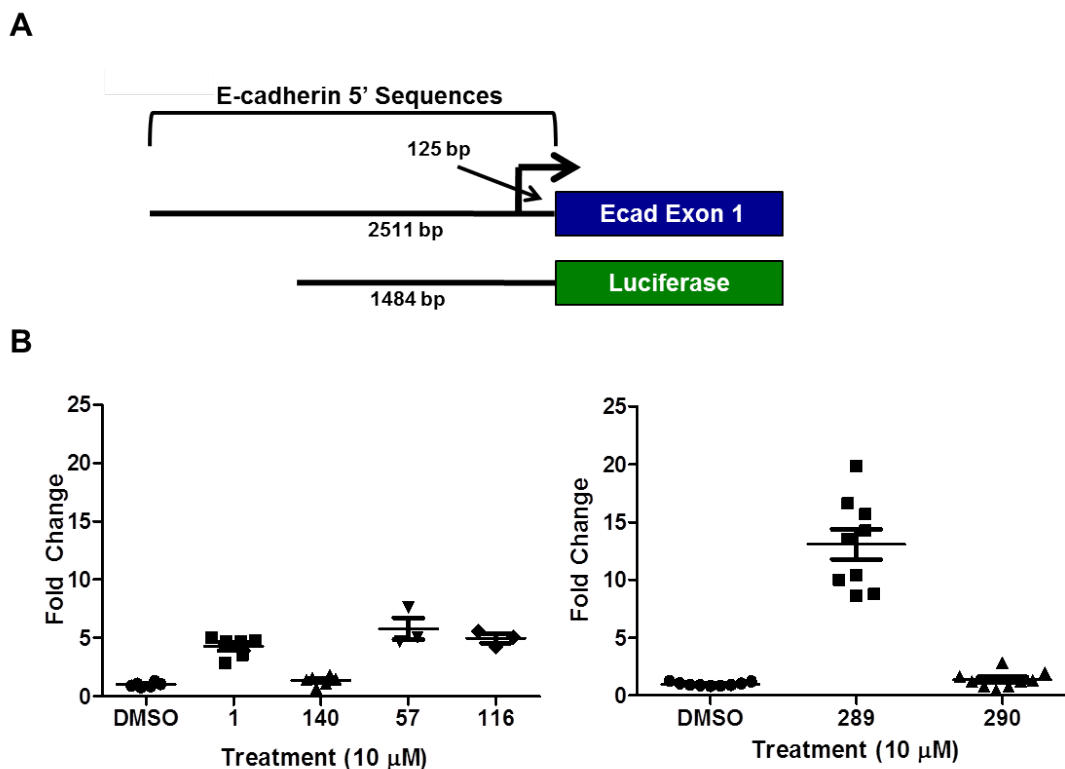
Figure 7. (A) E1 E-cadherin promoter reporter plasmid construct (B) Luciferase activity in SW620 cells transfected with plasmid E1 and treated immediately with a 10 µM concentration of selected compounds for 24 hours. All samples are normalized to DMSO control and compounds **140** and **290** are inactive analogs.

The plasmid was transfected into SW620 and H520 cells and then the cells were treated with a 10 µM concentration of selected analogs for 24 hours. A β-galactosidase reporter plasmid was used as a transfection control and was co-transfected with the E-cadherin promoter reporter plasmid in each well. The luciferase assay, which was a measure of E-cadherin promoter activity, was normalized to the β-galactosidase assay results for each sample and then the treatment groups were normalized to the DMSO control samples. The results for the SW620 and H520 cells can be seen in Figure 7B and Figure 8 respectively.
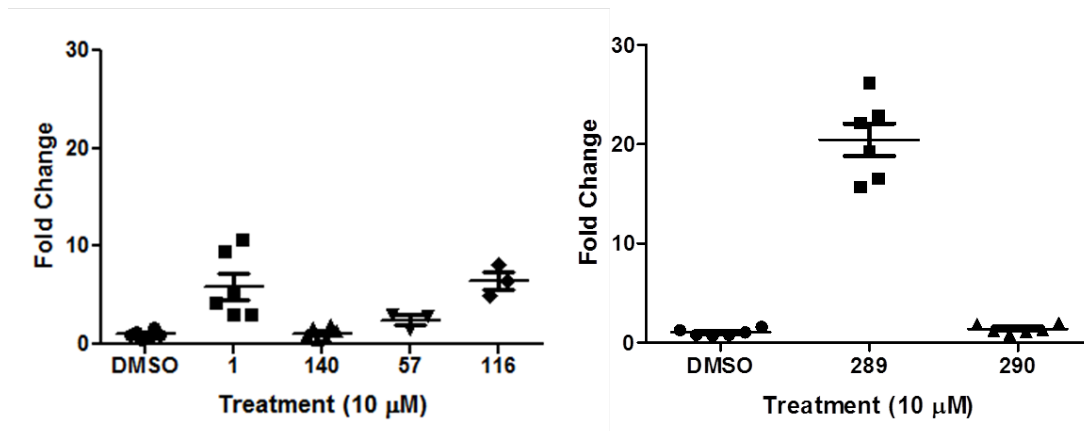
Figure 8. Luciferase activity in H520 cells transfected with plasmid E1 and treated immediately with a 10 μM concentration of selected compounds for 24 hours. All samples are normalized to DMSO control and compound **140** is an inactive analog.

Much to our surprise, as this was merely a fragment of the E-cadherin promoter and there are known upstream and downstream regulatory elements, we observed significant (p value < .005) increases in luciferase activity between the active selected analogs and the DMSO. Additionally, we saw no luciferase activity in the inactive analogs, compounds **140** and **290**, further confirming that the induction observed was a result of the mechanism of action of the active analogs.

With this information, we then directed our attention to narrowing down the region in which the active analogs were interacting and thus promoting transcription. In order to do this, we obtained 7 E-cadherin promoter reporter plasmid constructs from Ju Hshiung Chen's laboratory. Representation of these 7 E-cadherin promoter plasmid constructs (E2-E8) can be seen in Figure 9, and varied from 1.2 kbp of the E-cadherin promoter to 200 bp of the E-cadherin promoter upstream of the start codon [9]. Similar to plasmid E1, the E-cadherin promoter fragments in plasmids E2-E8 were directly

upstream from the firefly *Luc* gene; thus luciferase activity was utilized as a readout for

E-cadherin promoter induction after treatment with selected analogs.
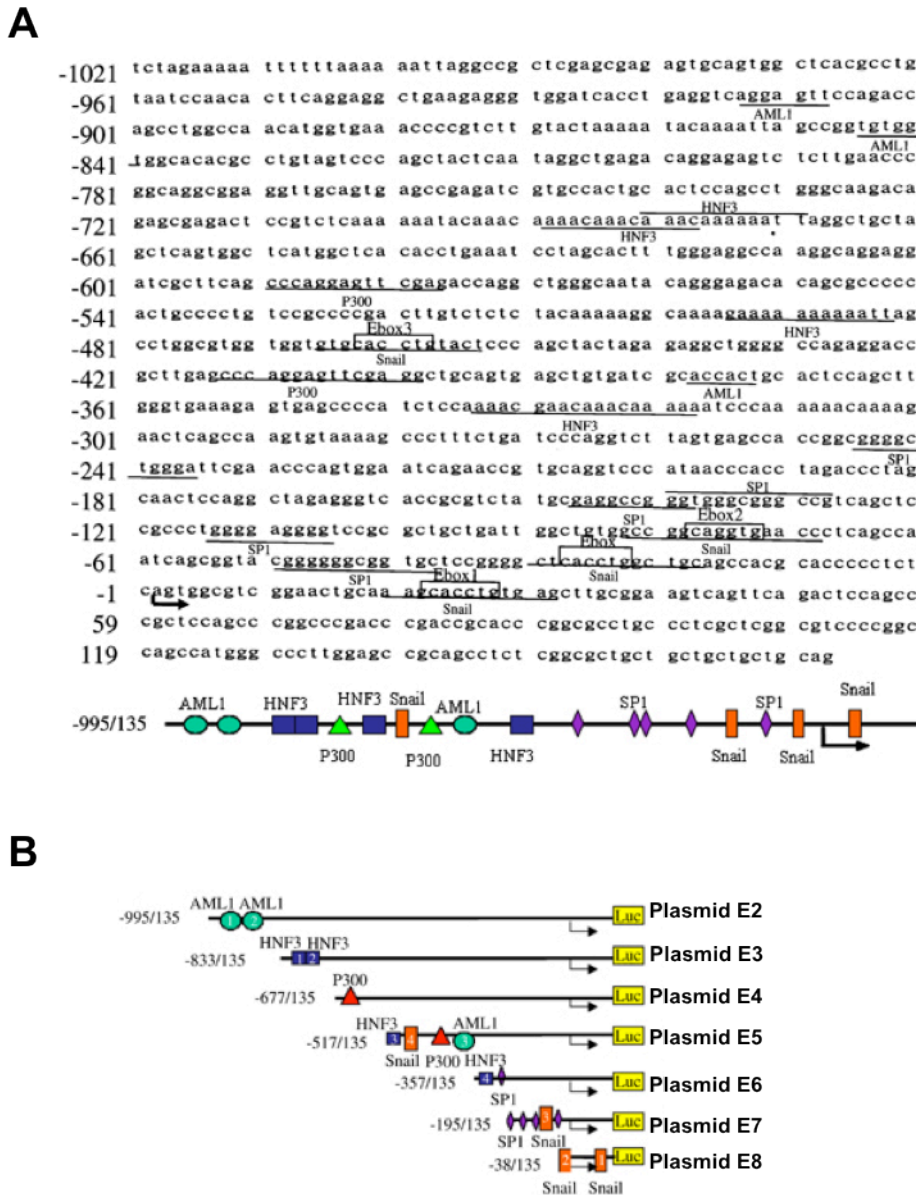
Figure 9. (A) E-Cadherin promoter sequence and highlighted transcription factor binding sites within the promoter sequence (B) E-cadherin promoter plasmid constructs that were synthesized (Liu, et al 2005).

I began by transfecting the SW620 cells with the smallest E-cadherin promoter plasmid (Plasmid E8) and then transfecting sequentially from smallest to largest E-cadherin promoter plasmid (E8 → E2). Based on the data published from Chen's laboratory using the plasmids, I thought that there would be no activity in at least the smallest plasmid, E8, as seen in the MCF7 positive cell line [9]. As can be seen in Figure 10, there was a 14.5-fold induction in luciferase activity after the treatment with the active analog, compound **289**, following transient transfection of plasmid E8 in the SW620 cells. This induction in luciferase activity remained elevated with transfection of all the plasmids, E2-E8, followed by treatment with the compound **289** for 24 hours. However, there was a noticeable increase from 10-15-fold induction of luciferase activity in Plasmids E8 and E7 and the 26.9-fold induction in luciferase activity from Plasmid E6. There was a dip in induction of luciferase activity with the transfection of Plasmid E4, 16.9-fold, but it returned with the transfection of Plasmid E3. While luciferase induction remains elevated throughout each plasmid transfection, the fluctuation may be due to additional activating and repressing binding sites being available on the E-cadherin promoter region within the plasmid. However, the important note is the significant induction in the smallest plasmid, E8.
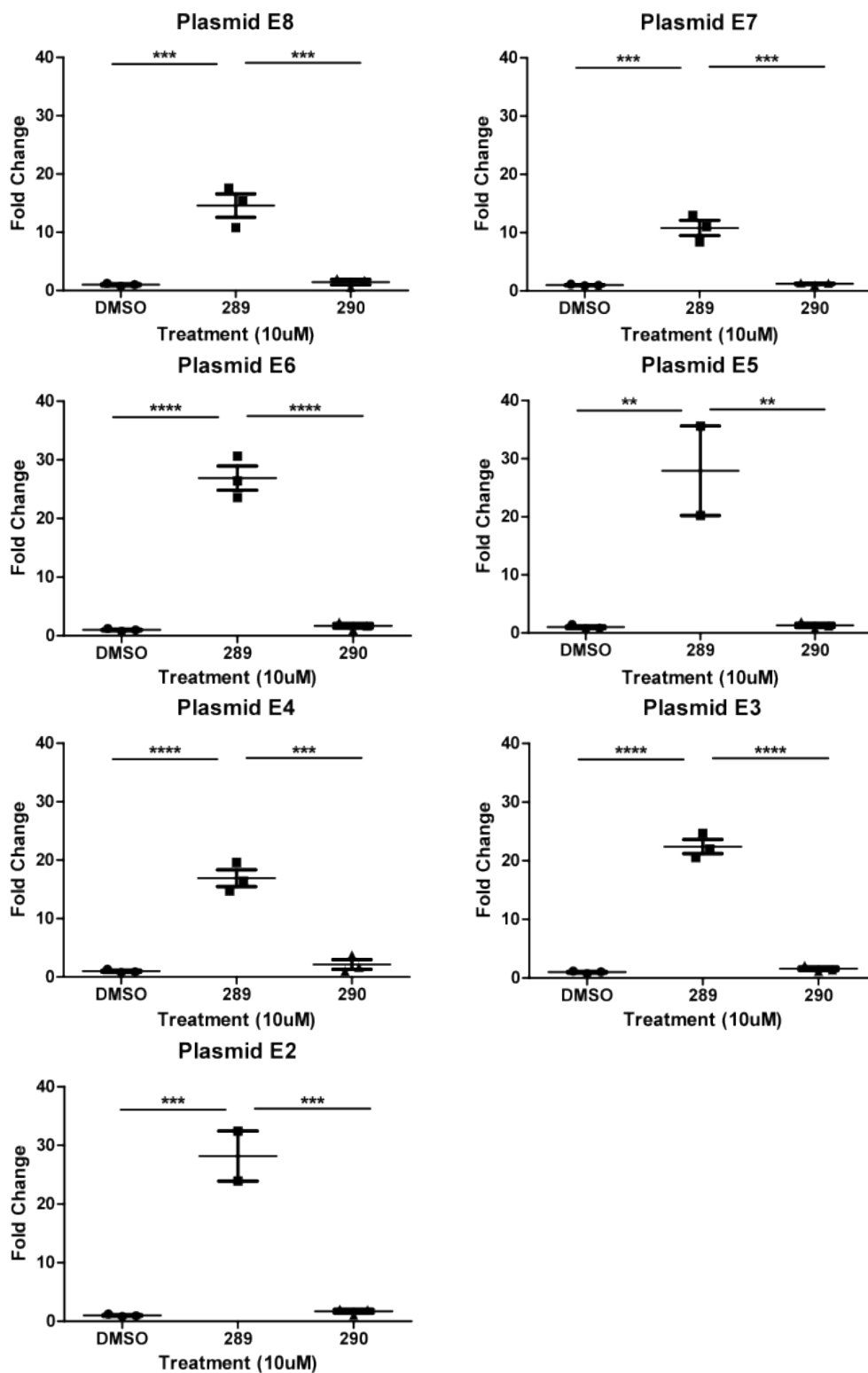
Figure 10. Luciferase activity in SW620 cells transfected with plasmid E8 – E2 and treated immediately with a 10 µM concentration of selected compounds for 24 hours. All samples are normalized to DMSO control and compound **290** an inactive analog.

It is known that there are 2 Snail transcription factor binding sites, or E-box sites, within the first 200 base pairs of the E-cadherin promoter. Additionally, one of these Snail transcription factor binding sites is located after the start codon [9]. Thus, the active compounds may be relieving the repression of Snail and ultimately the entire Snail repression binding complex, which includes HDAC 1/2. Further research will need to be conducted in order to determine if the Snail – HDAC 1/2 complex is being disrupted by our active compounds.

Discussion

Here we report the preliminary understanding of the mechanism of action for identified small molecules that restore E-cadherin expression. Due to steep SAR surrounding the original screening hit, compound **1**, and results from our preliminary screens we opted to take a biological approach to uncover the mechanism of action. We began by confirming that E-cadherin mRNA levels were upregulated upon treatment with compound **1** and active analogs. In addition, we looked specifically at when transcription was being "turned on" by the active compounds by conducting a time course experiment over 16 hours. Since all of our protein expression assays had been conducted after a 24 hour treatment, we were surprised to see that transcription of the *CDH1* gene was being "turned-on" after 3 hours, with a 7.26-fold increase in E-cadherin mRNA transcript levels with treatment of compound **289**. Additionally, it appears that transcription levels plateau at a 10 μM treatment with a slight induction occurring at 500 nM.

However, our thoughts of pursuing transcriptional regulation were further supported when we observed that NFATc1 and NFATc2 transcription levels increased in parallel with E-cadherin expression after both 6 hour and 24 hour treatments with compounds. This was observed in both the SW620 and H520 cells. At this point it appeared the next step would be to take advantage of a 1.4 kbps E-cadherin promoter reporter plasmid that was present in the lab. As this is only a section of the E-cadherin promoter, which is known to be 2.5 kbps, and with known enhancer elements up and downstream of the promoter region, we were unsure if the compounds would elicit an induction in luciferase activity after transfection of plasmid E1 into the SW620 and H520 cells. However, we were excited to observe a 5-fold induction of luciferase activity with treatment of compound **1** and preliminarily profiled active analogs. Additionally, a roughly 10-fold induction was seen with compound **289** in the SW620 cells, suggesting that the compounds are acting somewhere along the 1.4 kbp portion of the E-cadherin promoter to alleviate transcription repression or promote transcription activation. Through literature searches to identify the regulatory element binding sites within this fragment of the E-cadherin promoter we found Chen's laboratory's published research, which included the synthesis of 7 truncated E-cadherin promoter reporter plasmids. Compounds **289** and **290** were used to screen the 7 plasmids (Plasmids E2 –E8) in order to identify a narrow region of the E-cadherin promoter affected by the treatment of active compounds to relieve repression of *CDH1* transcription. To our surprise, a 14-fold induction of luciferase activity was observed in the smallest plasmid, E8, which includes only a 38 bp promoter fragment prior to the start codon. There are two known Snail binding sites within this region of the E-cadherin promoter. One is within the 38 bp prior

to the start codon, and one is found within the 135 bp following the start codon but before the first exon (or in the plasmid the luciferase gene). In addition, the induction of luciferase activity remained elevated (25-30 fold induction in some plasmids) within all of the plasmids suggesting that this small fragment contains the region of the promoter effected by our molecules.

Having narrowed the search to a small portion of the E-cadherin promoter we will be able to focus our attention on the transcription factor binding sites within this ~200 bp region of the promoter in an effort to identify the molecular target and further elucidate the mechanism of action. Additionally, as seen in Figure 9, it is shown that there are two Snail binding sites. Due to the similarity between our compounds and HDAC inhibitors as well as HDAC1/2 forming a known complex with Snail, this seems like a logical place to begin our efforts to determine if the molecules disrupt complex formation.

References

1. Yil maz, M. and G. Christofori, *EMT, the cytoskeleton, and cancer cell invasion.* Cancer Metastasis Rev, 2009. **28**(1-2): p. 15-33.

2. Cavallaro, U. and G. Christofori, *Cell adhesion and signalling by cadherins and Ig-CAMs in cancer.* Nat Rev Cancer, 2004. **4**(2): p. 118-32.

3. Perl, A.K., et al., *A causal role for E-cadherin in the transition from adenoma to carcinoma.* Nature, 1998. **392**(6672): p. 190-3.

4. Peinado, H., D. Olmeda, and A. Cano, *Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype?* Nat Rev Cancer, 2007. **7**(6): p. 415-28.

5. Kouzarides, T., *Chromatin modifications and their function.* Cell, 2007. **128**(4): p. 693-705.

6. Jenuwein, T. and C.D. Allis, *Translating the histone code.* Science, 2001. **293**(5532): p. 1074-80.

7. Zhang, Y. and D. Reinberg, *Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails.* Genes Dev, 2001. **15**(18): p. 2343-60.

8. Ji, X., et al., *Transcriptional defects underlie loss of E-cadherin expression in breast cancer.* Cell Growth Differ, 1997. **8**(7): p. 773-8.

9. Liu, Y.N., et al., *Regulatory mechanisms controlling human E-cadherin gene expression.* Oncogene, 2005. **24**(56): p. 8277-90.

10. Mancini, M. and A. Toker, *NFAT proteins: emerging roles in cancer progression.* Nat Rev Cancer, 2009. **9**(11): p. 810-20.

11. Robbs, B.K., et al., *Dual roles for NFAT transcription factor genes as oncogenes and tumor suppressors.* Mol Cell Biol, 2008. **28**(23): p. 7168-81.

ADDITIONAL BIOLOGICAL ANALYSIS OF NATURAL PRODUCTS

AND UNNATURAL ANALOGS

Introduction

In addition to my thesis research, I also conducted experiments to biologically evaluate synthesized natural compounds and their unnatural analogs. There were two main projects that I provided biological data for: the total synthesis of Tambjamine K and unnatural analogs synthesized by Leslie N. Aldrich and the total synthesis of (+)-7-Bromotrypargine and unnatural analogs synthesized by John T. Brogan. Both projects will be discussed in brief within this chapter.

The tambjamines A-J are a 2,2'-bipyrrolic class of cytotoxic alkaloids with diverse aliphatic termini isolated from bacteria and marine invertebrates. Members of this class have demonstrated a wide range of biological activities, including antitumor, antimicrobial, and immunosuppressive properties [1-8]. Gavagnin and co-workers described the isolation and characterization of a new member of the tambjamine family, tambjamine K, from the Azorean nudibranch *Tambja ceutae*, which displayed

antiproliferative and cytotoxic effects against tumor cell lines [9]. Based on these data, we initiated an effort for the total synthesis and biological evaluation of tambjamine K, along with a library of unnatural analogs with unprecedented diversity in the eastern C7 position (**Figure 1**).
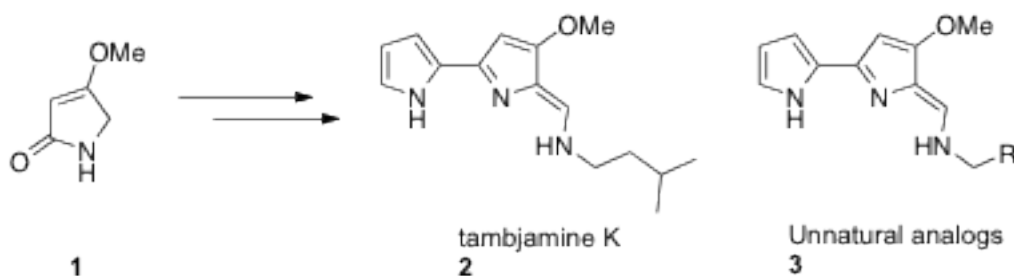
Figure 1. Total synthesis of tambjamine K and unnatural analogs

β-Carboline alkaloids are a prevalent class of biologically active natural products from marine organisms. They exhibit diverse structural features and distinct neuropharmacological profiles [10, 11]. Our lab has worked extensively in this arena and we were recently attracted to a class of β-carboline alkaloids represented by the trypargines **4-6** (**Figure 2**), as these alkaloids mapped well onto the H_3 pharmacophore model and offered a synthetic challenge [12-15]. The H_3 receptor (H_3R) is a presynaptic autoreceptor within the Class A GPCR family, but also functions as a heteroreceptor modulating levels of neurotransmitters such as dopamine, acetylcholine, norepinephrine, serotonin, GABA, and glutamate. Thus, H_3R has garnered a great deal of interest from the pharmaceutical industry for the possible treatment of obesity, epilepsy, sleep/wake disorders, schizophrenia, Alzheimer's disease, neuropathic pain, and ADHD [16-18].

Both trypargine and 6-hydroxytrypargine are highly toxic alkaloids. (+)-7-Bromotrypargine, was only recently isolated by Quinn from the Australian marine sponge *Ancornia sp.*, and found to possess antimalarial activity. 6-Bromotyramine was isolated along with (+)-7-Bromotrypargine in similar quantities, and is believed to be a key biosynthetic precursor [19]. Based on the neuropharmacological profiles of β-carboline alkaloids, and the electron-deficient nature of (+)-7-Bromotrypargine, relative to the electron-rich congeners trypargine and 6-hydroxytrypargine, which might diminish the cytotoxicity, the total synthesis of (+)-7-Bromotrypargine and biological evaluation seemed warranted.
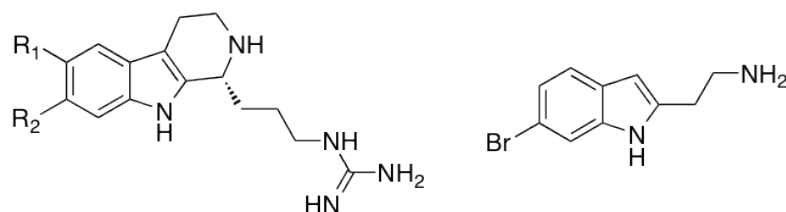
4, R_1=H, R_2=H, (+)-trypargine
5, R_1=OH, R_2=H, (+)-6-OH-trypargine
6, R_1=H, R_2=Br, (+)-7-Br-trypargine

7, 6-bromotyramine

Figure 2. Structures of the (+)-trypargines **4-6** and 6-bromotyramine **7**.

Methods and Materials

130

Cell Culture

A colorectal adenocarcinoma cell line, SW620, and a squamous cell lung carcinoma cell line, H520, were obtained from the American Type Culture Collection (Manassas, VA) and maintained in a humidified atmosphere of 5% CO_2 in air at 37 °C. The cells were routinely cultured in RPMI 1640 supplemented medium with 10% fetal bovine serum (Atlanta Biologicals, GA) and 100 μg/mL penicillin-streptomycin.

Viability Assay

SW620 and H520 cells (2.5×10^4/100 μL) were seeded in 96-well microtiter plates prior to treatment. Cells were treated with a 10 μM concentration of synthesized compound in quadruplicate for 24 hours and 48 hours in RPMI 1640 supplemented medium and 100 μg/mL penicillin-streptomycin. The Quick Cell Proliferation Assay Kit from BioVision (Mountain View, CA) was used to measure viability. The complete protocol can be referred to in the Methods & Materials section of Chapter 3.

Invasion Analysis

SW620 cells (1.0×10^6/mL) were seeded in 6 cm round dishes prior to treatment. Cells were treated with a 10 μM concentration of synthesized compound for 24 hours in RPMI 1640 supplemented medium and 100 μg/mL penicillin–streptomycin. 40 μL (2.5 mg/mL) of BD Matrigel Basement Membrane Matrix (BD Biosciences, Bedford, MA) was added to the top of the insert of a 24-well Transwell Permeable Support plates with polycarbonate membrane with 8 μm pore size (Corning Inc, Corning, NY). Then the cells were trypsinized and 3×10^5/250 μL cells were added to the top of the chamber in serum free RPMI medium, and 1 mL of RPMI medium with 10% FBS was added to the bottom

of the well. Then the plates were incubated for 72 hours at 37 °C in a humidified incubator with 5% CO_2 in the air. Then the wells were stained with 1% crystal violet in 50% methanol for 1 hour and washed in PBS. The membrane was cut off, adhered to a slide with glycerol, and analyzed in 20x□ field via microscopy. 3 – 20x□ fields were quantified per membrane.

Proliferation Analysis

SW620 and H520 cells (2.5 x 10^4/100 μL) were seeded in 96-well microtiter plates prior to treatment. Cells were treated with a 10 μM concentration of selected compounds in triplicate for 48 hours in RPMI 1640 supplemented medium and 100 μg/mL penicillin-streptomycin. The CycLex® Cellular BrdU ELISA Kit from MBL International (Woburn, MA) was used to measure proliferation. The complete protocol can be referred to in the Methods & Materials section of Chapter 3.

Results

Biological Evaluation of Tambjamine K and Synthesized Unnatural Analogs

As mentioned, the total synthesis of tambjamine K was reported along with the synthesis of a library of unnatural analogs. The analogs were designed to incorporate functionalized benzyl, heteroaryl moieties, and other previously undescribed analogs with varying degrees of lipophilicity and basicity to further develop SAR (**Figure 3**).
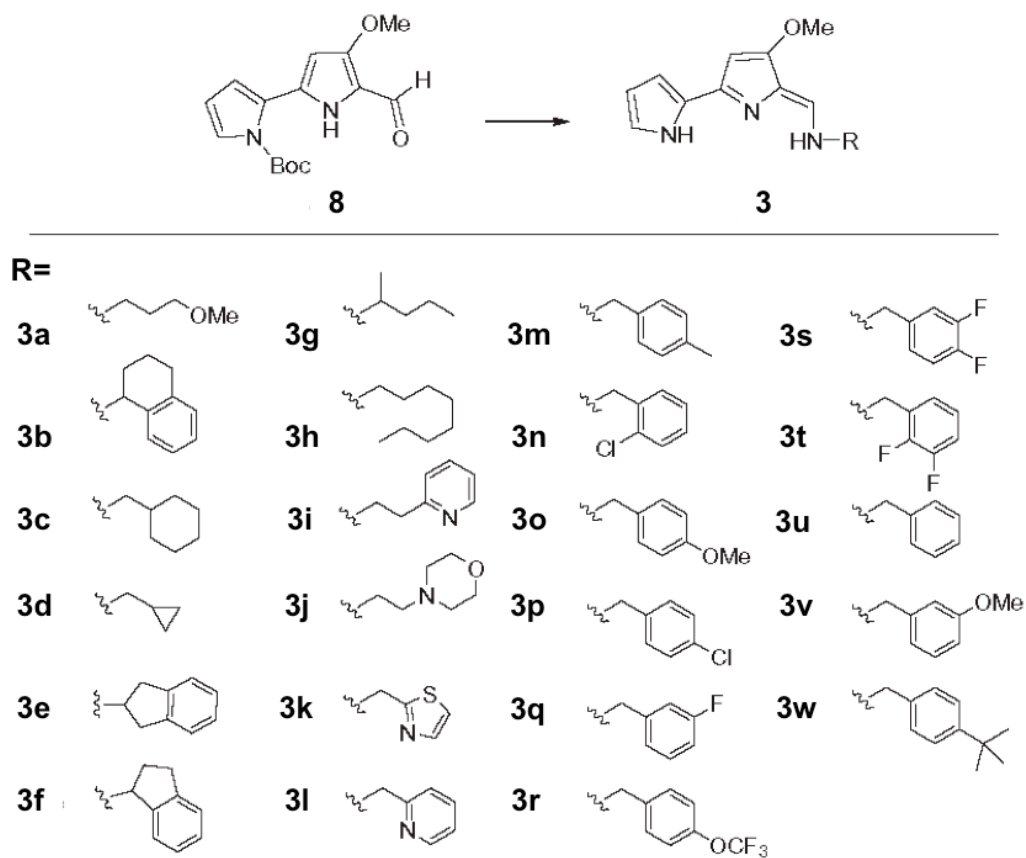
Figure 3. Library of unnatural analogs synthesized to characterize SAR further surrounding this 2,2'-bipyrrolic class of cytotoxic alkyloids. Reagents and conditions: RNH_2, 0.87 M HCl, MeOH, room temperature to 50 °C, 24–48 hours, 35–88%.

Based on viability data collected in the HTC116 and MB231 cell assays, we then

evaluated select unnatural analogs in a standard 48 hour cell viability assay (**Figure 4**) using another colorectal line (SW620) and a lung cancer cell line (H520). Interestingly, **9**, the most potent tambjamine analog in both the HCT116 (IC_{50} = 146 nM) and MB231 (IC_{50} = 362 nM) viability assays, had no effect on viability in either the SW620 or the H520 cell lines. However, unnatural analog **3b**, displayed a significant effect on inhibiting viability in both cancer cell lines, while other analogs showed varying effects. Further analysis would need to be conducted to determine if the decrease in viability was a result of apoptosis (or cell death) or inhibition of proliferation.
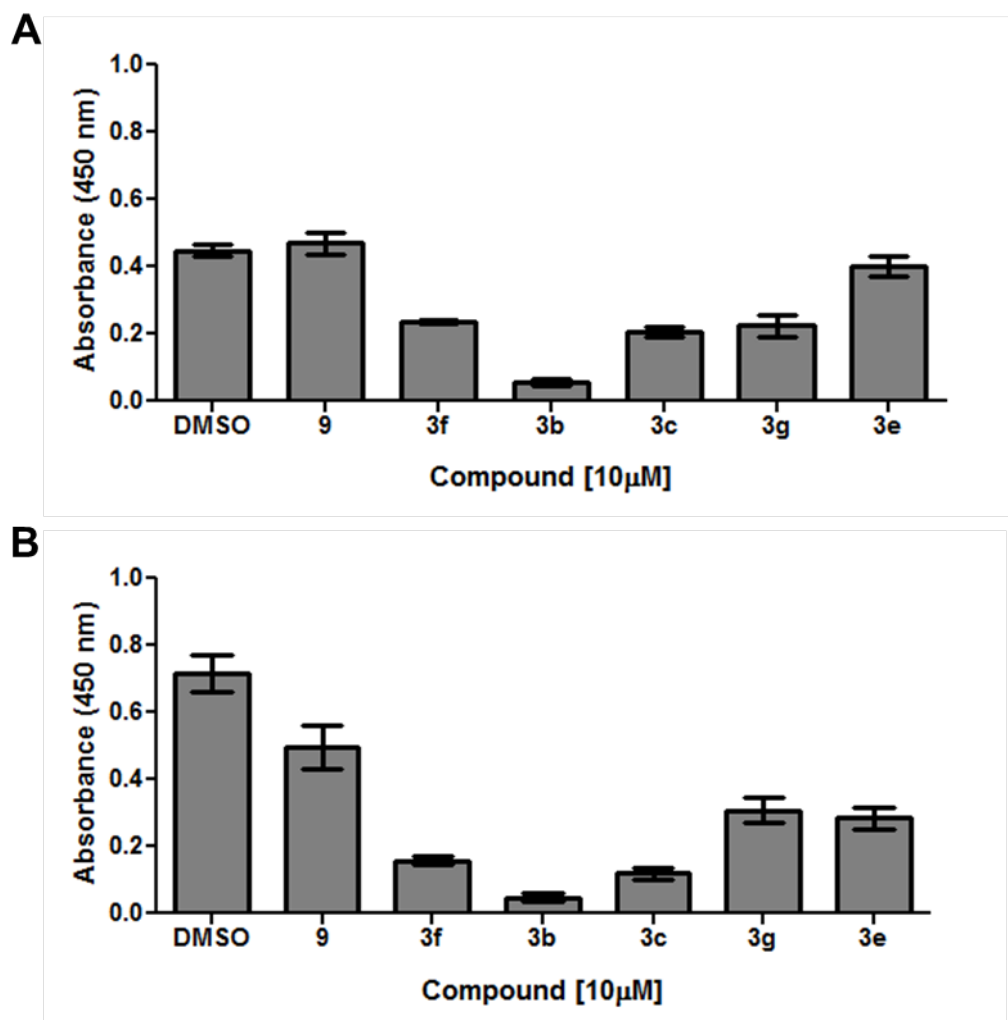
Figure 4. Single point (10 μM) screen of select unnatural tambjamine analogs in a 48 hour cell proliferation assays. (A) Proliferation assay with SW620 cell line; (B) proliferation assay with H520 cell line.

Although we were more interested in the compounds abilities to block invasion. As both unnatural tambjamine analogs **9** and **3e** displayed minimal to no effect on viability in SW620 cells, we examined their ability to block invasion in this cancer cell line, as the ability of tumor cells to invade into the surrounding microenvironment is a defining step in tumor progression. As shown in Figure 5, both **9** and **3e** significantly blocked

invasion, with compound **9** completely inhibiting invasion. The data suggests that while the compounds do not effect SW620 cell viability they do significantly block invasion.
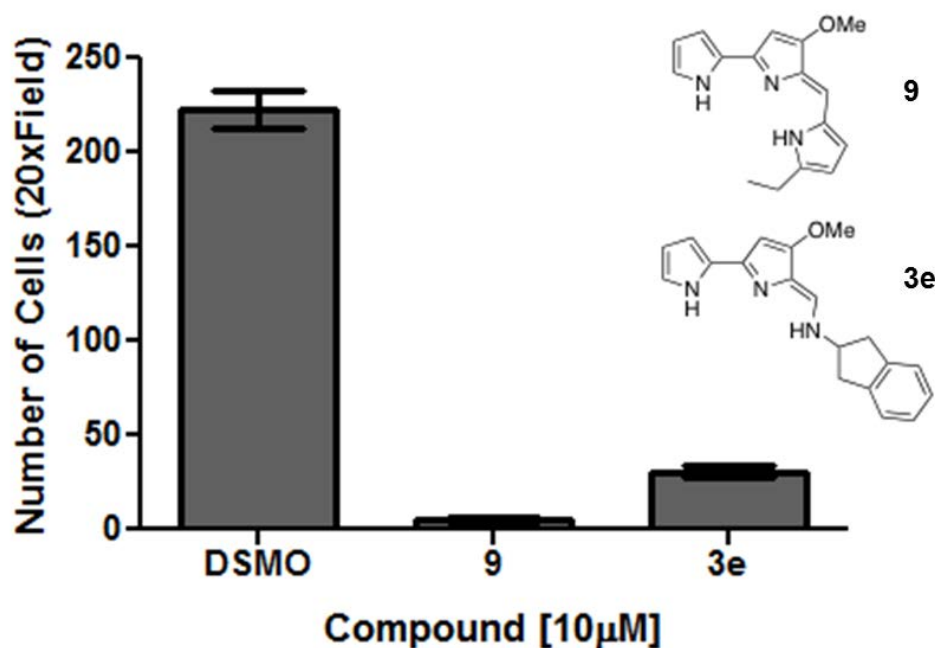
Figure 5. Single point (10 µM) screen of select unnatural tambjamine analogs in 72 hour cell invasion assay in the SW620 cell line.

Biological Evaluation of (+)-Bromotrypargine and Synthesized Unnatural Analogs

As **4** and **5** are highly toxic alkaloids, we first evaluated **6** in a standard cytotoxicity assay and found **6** to be non-toxic up to 20 µM, suggesting the pharmacological profile might diverge from **4** and **5**. This surprising result led us to study **6** in our standard HCT116 colon carcinoma cell viability assay [20, 21]. Here as well, **6** had no affect on HCT116 cell viability after 48 hours. In contrast, advanced

intermediate **8** (**Figure 6**), an unnatural analog of **7**, displayed an IC_{50} of 3 µM in this assay, completely killing the cells after 48 hours.
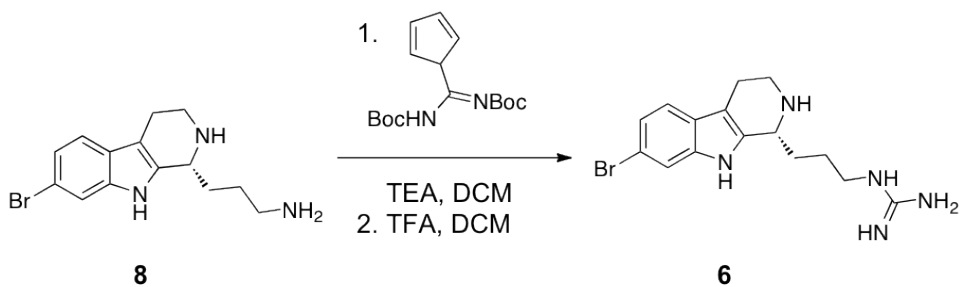
Figure 6. The final step towards the synthesis of (+)-7-bromotrypargine which displays the unnatural precursor analog (**8**) screened in the proliferation and viability assays.

These data prompted examination of unnatural analog **8** in an additional colorectal cancer (SW620) and lung cancer (H520) cell lines [22]. Interestingly, **8** had dramatic effect inhibiting cell viability in the WST-1 assay at various concentrations. This lead us to further elucidate a possibly mechanism for the decrease in viability. For this we screened compound **6** and **8** in the BrdU assay to look at the compounds' effects on proliferation. Again, we saw that compound **8** significantly reduced proliferation of both the SW620 and H520 cells at various concentrations (**Figure 7**). Overall, compound 8 (10 µM) had similar effects on viability and proliferation as the positive control, Sodium Butyrate. Collectively, these data informed us of two important points: 1) the pharmacology of the more electron deficient (+)-7-bromotrypargine (**6**) is distinct from **4** and **5** and warrants further biological evaluation as it lacks toxicity, and 2) unnatural

analog **8** possesses an intriguing pharmacological profile warranting the synthesis and characterization of additional unnatural analogs of **6**.
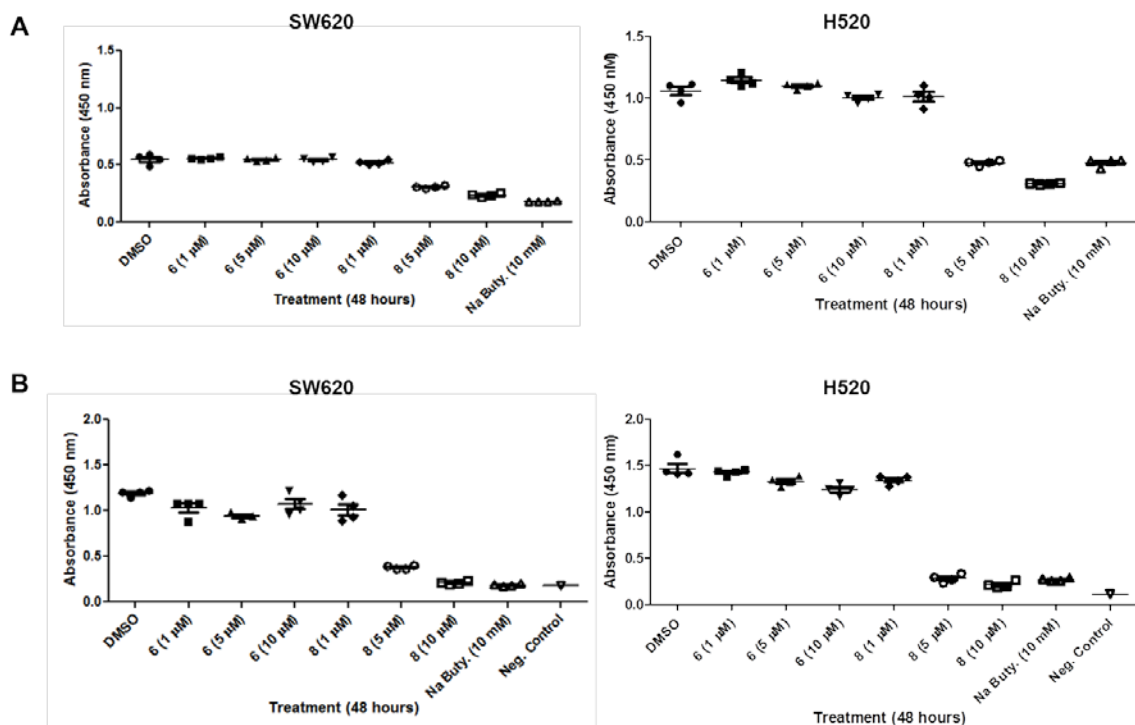
Figure 7. Multiple treatment points for (+)-7-bromotrypargine and the unnatural intermediate analog in cell proliferation and viability assays in both the SW620 and H520 cell lines. Sodium butyrate was used as a positive control. (A) WST-1 viability assay (B) BrdU proliferation assay.

Discussion

In summary, we completed the first total synthesis of tambjamine K in 18% overall yield coupled with biological evaluation in viability assays in both colon (HCT116) and breast cancer (MB231) cell lines. We also prepared a library of unnatural tambjamine analogs with unprecedented diversity and improved biological activity

against a number of tumor cell lines in viability and invasion assays. This effort demonstrated that subtle changes to the tambjamine core afford varying degrees of selectivity against different tumor cell lines. These data argue for further exploration of the tambjamine scaffold coupled with evaluation (viability, proliferation, and invasion) in additional human tumor cell lines. Current efforts are focused on synthesizing a second generation library including the discrete enantiomers of **3b** and **3e**, chiral a-methyl congers of the benzylic analogs **3m–3w**, and more focused analogs based on **9**. In parallel, we are working to identify the molecular target(s) for these unnatural analogs by evaluating **9**, **3b**, and **3e** against large screening panels of kinases, growth factor receptors, and phosphatases as a primary approach.

In summary, we have completed the first total synthesis of (+)-7-bromotrypargine (**6**) in nine steps (8 steps longest linear sequence) in 36.9% overall yield from commercial materials. Biological evaluation of **6** and an advanced intermediate **8** proved very exciting, with **6** displaying divergent pharmacology from related β-carbolines **4** and **5**, while **8** was extremely cytotoxic in mutiple non-transformed and colon cancer cell lines. Importantly, receptor-profiling efforts identified **6** as a moderately potent, dual NET/DAT inhibitor, and only the second known compound, and chemotype, to display such a pharmalogical profile devoid of SERT activity. This finding was in addition to the anticipated activity as an H_3 antagonist, based on the H_3 pharmacophore model. The intriguing pharmacological profile led us to then explore chemistry to access unnatural analogs. These reactions will serve as the groundwork for a larger effort aimed at unnatural analog synthesis to develop SAR around **6** and to optimize dual NET/DAT activity.

References

1. Wrede, F. and O. Hettche, *Über das Prodigiosin, den roten Farbstoff des Bacillus Prodigiosus (I. Mitteil.).* Ber. Dtsch. Chem. Ges. B., 1929. **62**(9): p. 2678-2685.

2. Wasserman, H.H., D.J. Friedland, and D.A. Morrison, *A novel dipyrrolyldipyrromethene prodigiosin analog from Serratia marcescens.* Tetrahedron Lett, 1968. **6**: p. 641-4.

3. Pinkerton, D.M., M.G. Banwell, and A.C. Willis, *Total syntheses of tambjamines C, E, F, G, H, I and J, BE-18591, and a related alkaloid from the marine bacterium Pseudoalteromonas tunicata.* Org Lett, 2007. **9**(24): p. 5127-30.

4. Lindquist, N. and W. Fenical, *New tamjamine class alkaloids from the marine ascidian Atapozoa sp. and its nudibranch predators. Origin of the tambjamines in Atapozoa.* Experientia, 1991. **47**(5): p. 504-506.

5. Kazlauskas, R., et al., *A blue pigment from a compound ascidian.* Australian Journal of Chemistry, 1982. **35**(1): p. 215-217.

6. Carte, B. and D.J. Faulkner, *Role of secondary metabolites in feeding associations between a predatory nudibranch, two grazing nudibranchs, and a bryozoan* Journal of Chemical Ecology, 1986. **12**(3): p. 795-804.

7. Carte, B. and D.J. Faulkner, *Defensive metabolites from three nembrothid nudibranchs.* J. Org. Chem, 1983. **48**(14): p. 2314-2318.

8. Blackman, A.J. and C. Li, *New Tambjamine Alkaloids From the Marine Bryozoan Bugula dentata.* Australian Journal of Chemistry, 1997. **47**(8): p. 1625-1629.

9. Carbone, M., et al., *A new cytotoxic tambjamine alkaloid from the Azorean nudibranch Tambja ceutae.* Bioorg Med Chem Lett. **20**(8): p. 2668-70.

10. Phuong, N.M., et al., *Two new beta-carboline alkaloids from Hedyotis capitellata var. mollis.* Planta Med, 1999. **65**(8): p. 761-2.

11. Cain, M., et al., *Beta-carbolines: synthesis and neurochemical and pharmacological actions on brain benzodiazepine receptors.* J Med Chem, 1982. **25**(9): p. 1081-91.

12. Lindsley, C.W., et al., *A 'one pot' microwave-mediated synthesis of the basic canthine skeleton: expedient acces to unnatural beta-carboline alkaloids.* Tetrahedron Lett, 2003. **44**: p. 4495-4498.

13.   Kennedy, J.P., M.L. Breininger, and C.W. Lindsley, *Total synthesis of Eudistomins Y1-Y6.* Tetrahedron Lett, 2009. **50**: p. 7076-7069.

14.   Cesar, L.M., et al., *Isolation and chemical characterization of PwTx-II: a novel alkaloid toxin from the venom of the spider Parawixia bistriata (Araneidae, Araneae).* Toxicon, 2005. **46**(7): p. 786-96.

15.   Akizawa, T., et al., *Trypargine, a new tetrahydro-beta-carboline of animal origin: isolation and chemical characterization from the skin of the African rhacophorid frog, Kassina senegalensis.* Biomed Res, 1982. **3**: p. 232-234.

16.   Lebois, E.P., C.K. Jones, and C.W. Lindsley, *The evolution of histamine H antagonists/inverse agonists.* Curr Top Med Chem. **11**(6): p. 648-60.

17.   Leurs, R., et al., *The histamine H3 receptor: from gene cloning to H3 receptor drugs.* Nat Rev Drug Discov, 2005. **4**(2): p. 107-20.

18.   Hudkins, R.L. and R. Raddatz, *Recent advances in drug discovery of histamine H3 antagonist.* Annu Rep Med Chem, 2008. **42**: p. 49-63.

19.   Davis, R.A., et al., *(+)--7-Bromotrypargine: an antimalarial beta-carboline from the Australian marine sponge Ancorina sp. .* Tetrahedron Lett, 2010. **51**: p. 583-585.

20.   Aldrich, L.N., et al., *Total synthesis and biological evaluation of tambjamine K and a library of unnatural analogs.* Bioorg Med Chem Lett. **20**(17): p. 5207-11.

21.   Daniels, R.N., et al., *Progress toward the total synthesis of lucentamycin A: total synthesis and biological evaluation of 8-epi-lucentamycin A.* J Org Chem, 2009. **74**(22): p. 8852-5.

22.   Stoops, S.L., et al., *Identification and optimization of small molecules that restore E-cadherin expression and reduce invasion in colorectal carcinoma cells.* ACS Chem Biol. **6**(5): p. 452-65.

CHAPTER VI

CONCLUDING REMARKS

Summary

In summary, a novel high-throughput screen was developed to identify small molecules that restored E-cadherin expression in the SW620 cell line followed by medicinal chemistry employing iterative analog library synthesis to develop SAR. Preliminary optimization of the screening hit has shown it is possible to synthesize small molecules that have an improved ability to restore E-cadherin expression compared to the initial screening hits. This restoration of protein has been confirmed by visualization of E-cadherin at the membrane via immunofluorescent microscopy. Further biological evaluation of profiled analogs has shown a minimal effect on cell proliferation, decrease in cellular invasion, and no cytoxicity in a normal-like epithelial cell line (MCF10A).

In addition we began to focus our attention on understanding the mechanism of action of these small molecules to restore E-cadherin expression. Quantitative PCR analysis has shown that treatment with selected small molecules increased mRNA expression 24.22-fold after 6 hours, which increased further to 58.68-fold after 16 hours of treatment, suggesting the small molecules are altering transcription of the *CDH1* gene. This was supported by experiments conducted using a plasmid construct containing a 1.4 kbp fragment of the E-cadherin promoter and luciferase reporter. After transfection, the cells were treated with selected compounds, lysed, and luciferase activity was measured.

It was shown that selected active analogs had a significant increase in luciferase activity as compared to DMSO or an inactive analog, which were used as controls.

More specifically, this suggests that the small molecules are specifically effecting transcription within the 1.4 kbp promoter region of the *CDH1* gene. Use of 7 truncated E-cadherin promoter reporter plasmids in the same experiment identified the smallest plasmid (E8) consisting of ~200 base pairs as the target for these compounds. Future work will include further pursuing this narrowed region of the E-cadherin promoter in an effort to identify the mechanism of action and molecular target for these small molecules.

On-Going Efforts

On-going efforts within the project are two-fold: analyzing and vetting the RNA-Seq data and EMT studies. The main focus is extracting data from the RNA-Seq analysis that was conducted and used to confirm that E-cadherin was upregulated on a transcriptional level. Knowing that NFATc1 and NFATc2 appear to be under the same repressed transcriptional regulation that is being relieved by treatment with the compounds aids in narrowing down potential transcription factors or associated proteins in complex that are being targeted by the compounds. The number of possible targets can be further narrowed down since we have identified a 200 bp region of the E-cadherin promoter in which we believe the compounds are acting. Additionally, with the synthesis of more efficacious and potent compounds, we will send new RNA samples from DMSO, compound **290**, and compound **289** treated SW620 cells for RNA-Seq analysis. Our rational is that since we see a 7.26-fold increase in E-cadherin mRNA levels with

compound **289** at 3 hours; we may find more distinct changes in mRNA levels genome wide. Additionally, this may aid in identifying a key binding site region that is conversed within the promoter region of all upregulated genes identified in the RNA-Seq analysis. Such information would allow for a more narrowed targeted approach towards identifying the molecular target.

The second on-going effort includes studying compound **289** in two distinct EMT cell models: the LIM1863 and NMuMg cells. Both cell lines can be induced to undergo EMT through the addition of TNF-α and/or TGF-β. The LIM1863 cell line is a unique colon carcinoma cell line that consists of organoids, which are morphologically and functionally organized and maintain tight junctional complexes and epithelial polarity in suspension culture [1, 2]. The LIM1863 organoids will undergo an EMT conversion from this well-differentiated organoid structure to a migratory monolayer phenotype (adhering to the plate bottom) in response to TGF-β [3, 4]. Moreover, it has been found that the EMT process was accelerated dramatically by co-stimulation with the inflammatory cytokine TNF-α [5]. Thus, we plan to take advantage of this unique and inducible EMT model to determine if our active analogs are capable of reverting and/or blocking EMT. The data will be collected using immunofluorescent microscopy and qPCR to identify changes in expression of hallmark EMT markers. In addition the distinct morphological change of the LIM1863 cells from organoids to a monolayer will provide initial confirmation to the effectiveness of compound **289**. Identical experiments will be conducted in the NMuMg cell line, which is an immortalized mouse mammary gland epithelial cell line. Similarly, the NMuMg cell line has been well established as a cell line that can undergo EMT via growth factor stimulation, specifically TGF-β [6, 7].

Overall, we hope to show that our active compounds are capable of reverting or blocking EMT, thus providing proof of concept for the molecules. At a minimum this would allow for the molecules to be used as research tools to further understand EMT during tumor progression.

Future Directions

There are two immediate directions of the project that will be addressed in the upcoming months. The first, and most important, is to identify the molecular target of compound **1** and synthesized analogs. We believe the mechanism of action of these small molecules is to release the repression of the *CDH1* gene thus allowing for transcription of the gene and ultimately restoration of the E-cadherin protein at the membrane. However, there are several repressor complexes as well as multiple binding sites for several repressor transcription factors (e.g. Snail, Snail, Twist, Zeb1). Many of these complexes also coordinate the methylation of DNA and deacetylation of histones within a gene promoter regions. Once the molecular target is identified, we will be able to better understand the specific mechanism of action and series of events within the cell that lead to E-cadherin restoration.

Determining the molecular target will occur through a dual approach: biological and chemical. As seen in Chapter 4, there were 7 plasmids with varying lengths of the E-cadherin promoter region from 200 base pairs to 1.2 kbp, which upon transfection into the SW620 cells and treatment with DMSO or selected analogs showed that the compounds were inducing transcription observed as an induction of luciferase activity. It

is believed the active compounds are releasing repression by interacting with proteins bound to the first 200 base pairs of the E-cadherin promoter (plasmid E8). Within the first 200 base pairs of the promoter region there are several binding sites for the Snail protein; E-box binding sites. Several approaches can be taken to confirm that the small molecules are disrupting the Snail repressor complex. A gel shift experiment, EMSA, using oligos containing the same E-cadherin promoter sequence that is in Plasmid 8 could confirm if Snail is bound after treatment with the selected active analogs or if there is disruption in complex formation on the promoter. Such an experiment can be supported through Co-IP experiments directed at Snail or key repressor complex proteins.

Once the repressor complex has been identified, additional biological research can be conducted in an effort to tease out if the compounds are preventing complex binding directing to the DNA binding site (ChIP), complex formation (Co-IP), or direct inhibition (Co-IP) of one or more specific proteins in the complex. As mentioned the second approach that will be utilized is a chemical approach via affinity chromatography. The most recent series of analogs synthesized opened up the middle carbon linker providing additional points where photoaffinity labels may be tolerated; retaining E-cadherin restoration activity. If such molecules are synthesized, an affinity chromatography protocol can be utilized to try and pull out the molecular target.

Identifying the molecular target will be key in both understanding the mechanism of action as well as further developing SAR around compound **1**. Knowing the molecular target may identify additional pathways that may be targeted in a multi-therapy approach to prevent invasion or induce apoptosis, which will be discussed further below. In addition, it will allow for *in vitro* direct biochemical assays or *in vivo* cellular assays to be

developed for measuring the function of the identified target. This will allow for SAR to be established specifically around the molecular target as opposed to measuring a desired phenotypic response. SAR directed at the molecular target will allow for the increase in potency and efficacy of the small molecules for the molecular target as well as allow us to build small molecule libraries outside the current patent landscape.

The second area of future research will be to identify potential synergistic effects between the most effective compound **1** analogs and the standard of care chemotherapeutics used for the treatment of colorectal cancer. The first chemotherapeutic that will be analyzed in combination with our small molecules will be 5-Fluorouracil (5-FU). 5-FU is a pyrimidine analog and acts as an irreversible thymidylate synthase inhibitor preventing the synthesis of thymidine, a necessary nucleotide required for DNA replication, thus resulting in cell death. In colorectal cancer, 5-FU based chemotherapy treatment regimens improve overall and disease-free survival of patients with resected stage III colorectal cancer. However, developed tumor cell resistance to 5-FU is a major limitation and response rate for advanced stage colorectal cancer is only 10-15% [8-10]. Combinations with new chemotherapies such as oxaliplatin have improved the response rate to 40-50%, but there is still a need to better improve the current therapeutic strategies especially for late stage disease [11].

Therefore, viability and proliferation will be measured in the SW620 cells after the treatment with a subtherapeutic dose of 5-FU in combination with the best compound **1** analogs. Variables to consider during these experiments will include specific doses of 5-FU as well as selected analogs, length of incubation, and treatment time (e.g. whether

co-dosed or a staggered treatment). Additionally, if synergy between 5-FU and the active analogs is evident, induction of apoptosis will also be analyzed.

Upon identification of the molecular target, signaling pathways may be identified that are affected by the signaling cascade triggered in response to our compounds. This may allow for pathway specific inhibitors to be tested in combination with our compounds to elicit a synergistic and tumor specific effect.   Additionally, it may be found that through the molecular target and mechanism of action that tumor cells may become 're-sensitized' to common chemotherapeutics that lead to resistance.   An example of such is restoration of EGFR protein with treatment of an HDAC inhibitor in lung cancer, thus restoring tumor sensitivity to EGFR inhibitors [12].   Since our compounds elicit a transcriptional effect, similar to HDAC inhibitors, this may be another avenue to pursue in specific lung cancer cell lines, such as the H520 cells.

Identifying the molecular target as well as potential synergistic effects with chemotherapeutics currently in the clinic will identify the therapeutic relevance of these molecules. If therapeutic relevance is identified there are several necessary avenues of research that will need to be pursued to push the project forward in a drug discovery manner. First, iterative parallel libraries of small molecules surrounding compound **1** as well as active analogs, such as compound **289**, will need to be synthesized to further optimize the SAR as well as to identify proprietary compounds. Selected optimized compounds will need to undergo traditional *in vitro* and *in vivo* DMPK studies to understand the metabolism of the small molecules. Concurrently, preliminary cancer studies in mouse models suited to replicate the specific cancer type should be conducted to determine tumor response to compound treatment in order to provide preclinical data.

I can only hope that the small molecules synthesized in my research aid in the advancement of cancer therapeutic development either as a targeted therapy or as a research tool to further understand the signaling pathways involved in EMT and invasion.

References

1.	Whitehead, R.H., et al., *A new colon carcinoma cell line (LIM1863) that grows as organoids with spontaneous differentiation into crypt-like structures in vitro.* Cancer Res, 1987. **47**(10): p. 2683-9.

2.	Bates, R.C. and A.M. Mercurio, *The epithelial-mesenchymal transition (EMT) and colorectal cancer progression.* Cancer Biol Ther, 2005. **4**(4): p. 365-70.

3.	Hayward, I.P. and R.H. Whitehead, *Patterns of growth and differentiation in the colon carcinoma cell line LIM 1863.* Int J Cancer, 1992. **50**(5): p. 752-9.

4.	Bates, R.C., et al., *Apoptosis induced by inhibition of intercellular contact.* J Cell Biol, 1994. **125**(2): p. 403-15.

5.	Bates, R.C. and A.M. Mercurio, *Tumor necrosis factor-alpha stimulates the epithelial-to-mesenchymal transition of human colonic organoids.* Mol Biol Cell, 2003. **14**(5): p. 1790-800.

6.	Korpal, M., et al., *The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2.* J Biol Chem, 2008. **283**(22): p. 14910-4.

7.	Miettinen, P.J., et al., *TGF-beta induced transdifferentiation of mammary epithelial cells to mesenchymal cells: involvement of type I receptors.* J Cell Biol, 1994. **127**(6 Pt 2): p. 2021-36.

8.	Longley, D.B., D.P. Harkin, and P.G. Johnston, *5-fluorouracil: mechanisms of action and clinical strategies.* Nat Rev Cancer, 2003. **3**(5): p. 330-8.

9.	*Efficacy of adjuvant fluorouracil and folinic acid in colon cancer. International Multicentre Pooled Analysis of Colon Cancer Trials (IMPACT) investigators.* Lancet, 1995. **345**(8955): p. 939-44.

10.	Johnston, P.G. and S. Kaye, *Capecitabine: a novel agent for the treatment of solid tumors.* Anticancer Drugs, 2001. **12**(8): p. 639-46.

11.	Giacchetti, S., et al., *Phase III multicenter randomized trial of oxaliplatin added to chronomodulated fluorouracil-leucovorin as first-line treatment of metastatic colorectal cancer.* J Clin Oncol, 2000. **18**(1): p. 136-47.

12.	Witta, S.E., et al., *Restoring E-cadherin expression increases sensitivity to epidermal growth factor receptor inhibitors in lung cancer cell lines.* Cancer Res, 2006. **66**(2): p. 944-50.
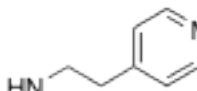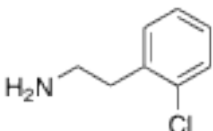
APPENDIX 1
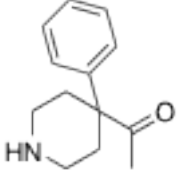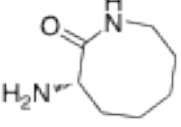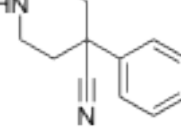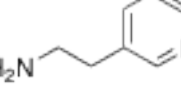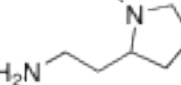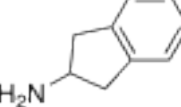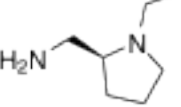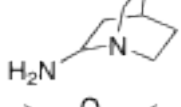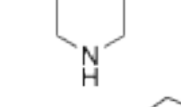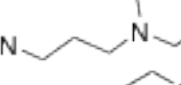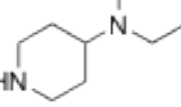
Table 1.
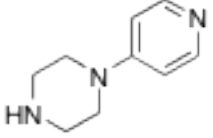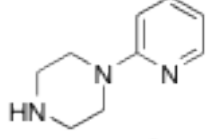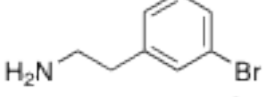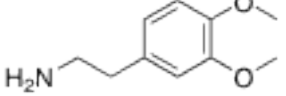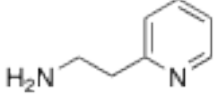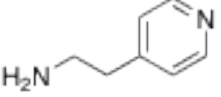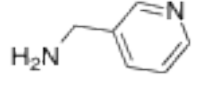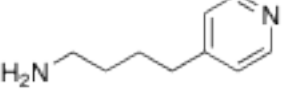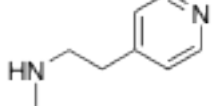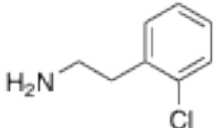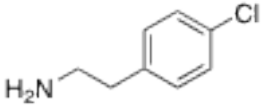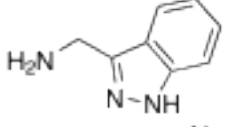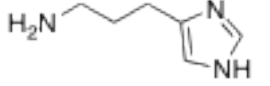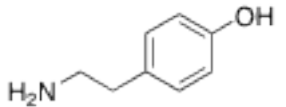
For X = OH: HNR_1R_2, EDC, HOBt, DIPEA, THF
For X = Cl: HNR_1R_2, DIPEA, DCM

Compound	R	WB[a] SW620	WB[a] H520	ICW[b] SW620	ICW[b] H520
1		5.68	5.55	2.19	4.79
5		7.30	10.06	5.03	
6		6.98	9.68	4.71	
7		5.05	10.01		
8		0.97	1.20		
9		4.29	8.05	3.88	
10		5.34	8.01	4.43	
11		0.96			
12		0.84			
13		0.92			

14		0.90		
15		0.75		
16		0.80		
17		4.97		5.41
18		0.86		
19		1.36	3.24	
20		1.13		
21		8.11	10.84	4.15
22		1.19		
23		1.12		
24		3.51	6.10	2.71
25		1.06		
26		1.25		

No.	Structure				
27		1.18			
28		1.13			
29		2.29	1.97	1.17	1.10
30		2.71	2.89	4.21	
31		12.90	4.55	4.49	
32		13.25	4.62	3.50	
33		2.50		0.72	1.65
34		8.91	9.20	4.44	
35		1.63		0.91	0.83
36		8.44	5.38	3.96	4.59
37				1.19	2.12
38				1.17	0.56
39				3.81	3.84
40				3.10	4.16

154

	R		
41		2.58	2.28
42		0.14	1.76
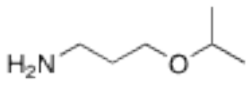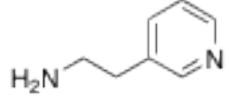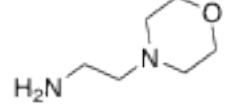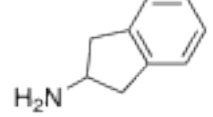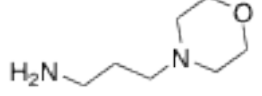
[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by Western blot for compound treatment at 10 μM. (TSA = 3.74).
[b] Data is fold change in E-cadherin restoration above a DMSO control as measured by In-Cell Western assay for compound treatment at 10 μM. (TSA = 2.37).
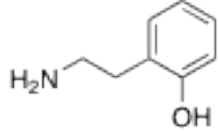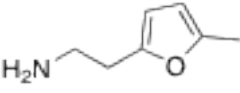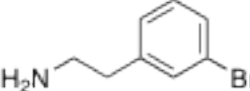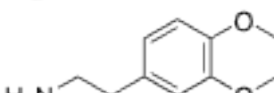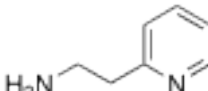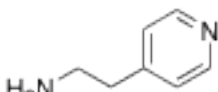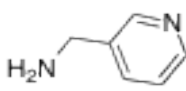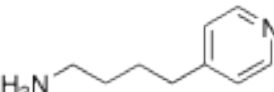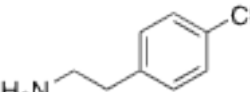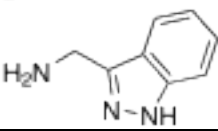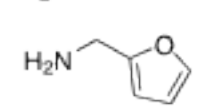
Table 2.

For X = OH: HNR_1R_2, EDC, HOBt, DIPEA, THF
For X = Cl: HNR_1R_2, DIPEA, DCM

Compound	R	WB[a] SW620	WB[a] H520	ICW[b] SW620	ICW[b] H520
43		2.43	3.74	1.58	
44		3.55	6.93	3.90	
45		4.40	8.87	6.46	
46		2.72	8.36	3.92	
47		1.13		1.25	
48		1.20			
49		1.18		1.61	

155

50		4.11	2.41	4.54	
51		3.17	3.23	5.89	
52		1.84	2.35		
53		2.82	3.57	3.64	
54		6.23	4.88	3.64	
55		8.60	5.66	3.02	
56		1.29			
57		9.6	11.65	6.33	
58		1.20			
59		7.98	2.41	4.18	3.64
60				2.21	2.15
61		0.93	-0.45		

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by Western blot for compound treatment at 10 μM. (TSA = 3.74).
[b] Data is fold change in E-cadherin restoration above a DMSO control as measured by In-Cell Western assay for compound treatment at 10 μM. (TSA = 2.37).
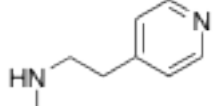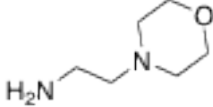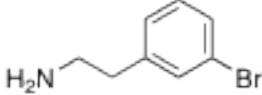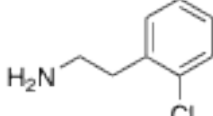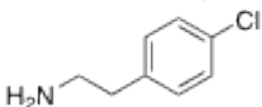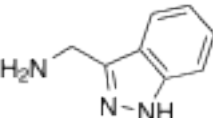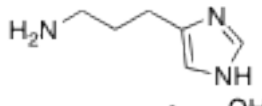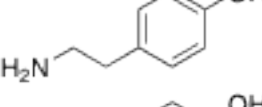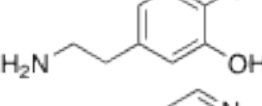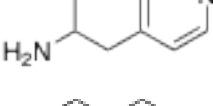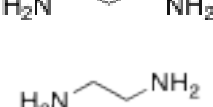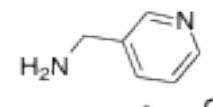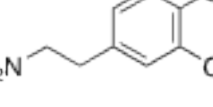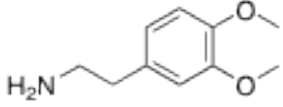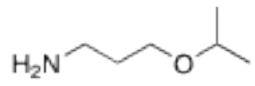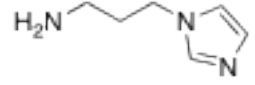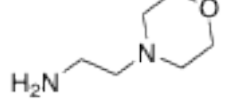
Table 3.
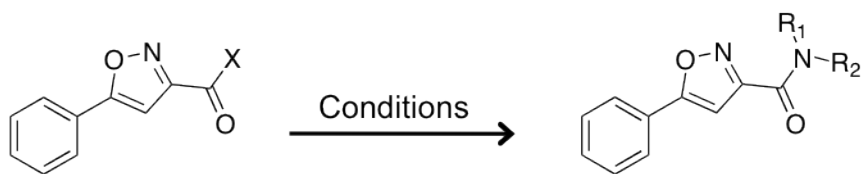
For X = OH: HNR$_1$R$_2$, EDC, HOBt, DIPEA, THF
For X = Cl: HNR$_1$R$_2$, DIPEA, DCM

Compound	R	WB[a] SW620	WB[a] H520	ICW[b] SW620	ICW[b] H520
62		3.71	6.40	2.59	
63		4.98	7.00	3.37	
64		4.66	6.78		
65		4.78	3.91	2.36	
66		1.29			
67		2.43	3.73	1.29	
68		4.99	5.37	2.26	
69		11.99	4.43	3.51	
70		13.69	4.64	1.16	
71		2.04			
72		9.6	8.79	4.36	
73		8.75	5.43	5.80	

157

74		1.29		
75		5.38	5.14	2.80
76		1.84		
77		2.52	0.64	0.91	1.42
78				2.24	1.87
79				1.36	-1.95
80				5.20	5.40
81				2.35	4.70
82				2.65	4.24
83				0.32	1.59
84				1.66	146
85				1.23	1.12
86		2.44	6.89	
87		1.58	1.04	
88		2.32	5.69	2.00

89		1.30
90		0.95
91		2.24 2.60

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by Western blot for compound treatment at 10 μM. (TSA = 3.74).
[b] Data is fold change in E-cadherin restoration above a DMSO control as measured by In-Cell Western assay for compound treatment at 10 μM. (TSA = 2.37).
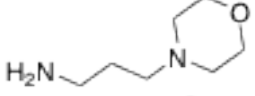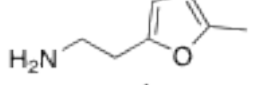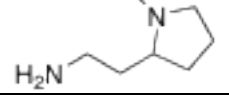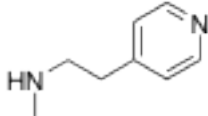
Table 4.
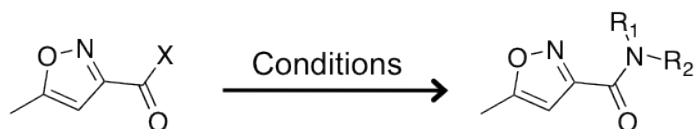
For X = OH: HNR$_1$R$_2$, EDC, HOBt, DIPEA, THF
For X = Cl: HNR$_1$R$_2$, DIPEA, DCM

Compound	R	WB[a] SW620	WB[a] H520	ICW SW620	ICW H520
92		1.08			
93		1.10			
94		0.87			
95		1.24			
96		0.84			
97		0.82			

159

98		1.37	
99		0.80	
100		0.83	
101		0.82	
102		0.96	
103		1.04	
104		2.11	0.89
105		1.38	

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by Western blot for compound treatment at 10 μM. (TSA = 3.74).
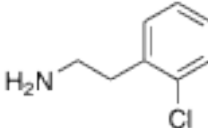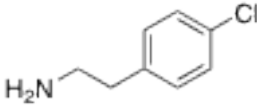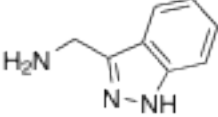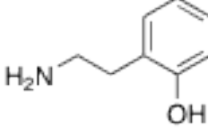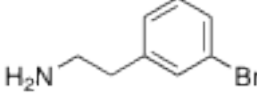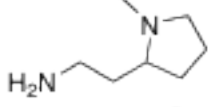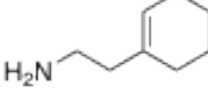
Table 5.
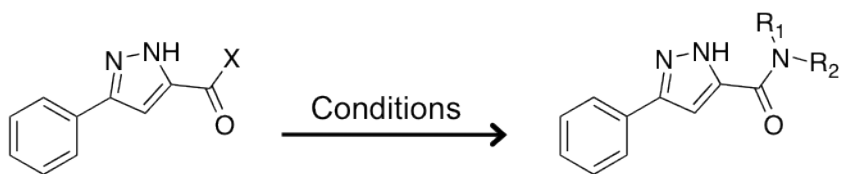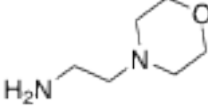

For X = OH: HNR$_1$R$_2$, EDC, HOBt, DIPEA, THF
For X = Cl: HNR$_1$R$_2$, DIPEA, DCM

Compound	R	WB[a] SW620	WB[a] H520	ICW[b] SW620	ICW[b] H520
106		2.10	2.93	1.28	
107		2.67	3.65	1.61	
108		1.39			
109		2.78	3.90	2.09	
110		4.16	4.05	1.73	
111		1.72		1.07	
112		3.50		1.71	
113		1.07			
114		1.29		0.87	
115		1.27		1.04	
116		7.23	8.69	5.68	
117		0.96		0.73	

161

118		1.03		
119		10.27	2.55	1.48
120			1.65	1.38
121			1.23	0.21
122		6.41	9.91	1.81
123		2.90	3.78	
124		1.44	1.16	0.95
125		1.67	1.21	
126		2.68	7.54	
127			1.19	0.96

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by Western blot for compound treatment at 10 μM. (TSA = 3.74).
[b] Data is fold change in E-cadherin restoration above a DMSO control as measured by In-Cell Western assay for compound treatment at 10 μM. (TSA = 2.37).

Table 6.
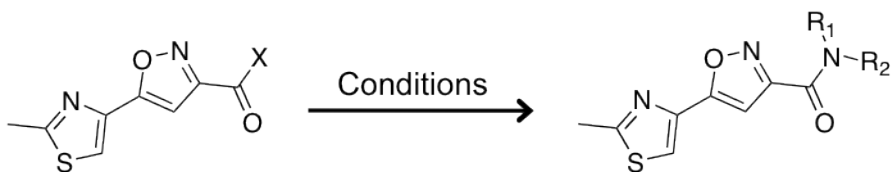
For X = OH: HNR$_1$R$_2$, EDC, HOBt, DIPEA, THF
For X = Cl: HNR$_1$R$_2$, DIPEA, DCM

Compound	R	WB[a] SW620	WB[a] H520	ICW[b] SW620	ICW[b] H520
128		1.01		1.23	
129		1.00		0.90	
130		1.23		1.52	
131		1.20		1.13	
132		1.14		1.01	
133		1.17			
134		0.96		0.96	
135		0.96		0.95	
136		1.06		0.99	
137		0.90			
138		0.94		1.12	

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by Western blot for compound treatment at 10 μM. (TSA = 3.74).
[b] Data is fold change in E-cadherin restoration above a DMSO control as measured by In-Cell Western assay for compound treatment at 10 μM. (TSA = 2.37).
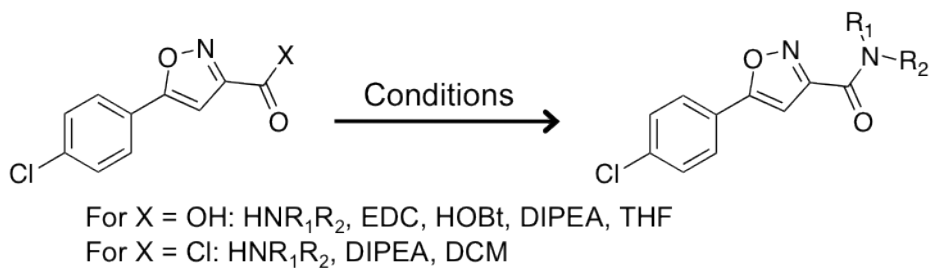
Table 7.

For X = OH: HNR$_1$R$_2$, EDC, HOBt, DIPEA, THF
For X = Cl: HNR$_1$R$_2$, DIPEA, DCM

Compound	R	WB[a] SW620	WB[a] H520	ICW[b] SW620	ICW[b] H520
139		1.94	0.73	1.32	
140		1.71	1.12	0.94	
141		2.09	1.03	1.01	
142		2.01	3.61	1.45	
143		1.17			
144		1.97	1.70	1.08	
145		1.00		1.28	
146		1.31		1.50	
147		1.07		1.14	
148		0.90		0.94	
149		1.18			

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by Western blot for compound treatment at 10 µM. (TSA = 3.74).
[b] Data is fold change in E-cadherin restoration above a DMSO control as measured by In-Cell Western assay for compound treatment at 10 µM. (TSA = 2.37).
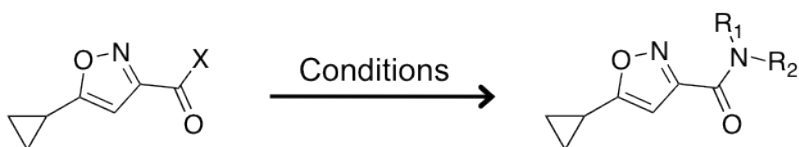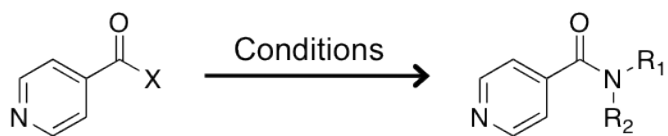
Table 8.

For X = OH: HNR$_1$R$_2$, EDC, HOBt, DIPEA, THF
For X = Cl: HNR$_1$R$_2$, DIPEA, DCM

Compound	R	WB[a] SW620	WB[a] H520	ICW SW620	ICW H520
150		1.03			
151		0.85			
152		1.15			
153		0.83			
154		0.78			
155		0.97			
156		0.90			
157		0.88			
158		0.91			
159		0.71			
160		8.30	0.95		

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by Western blot for compound treatment at 10 μM. (TSA = 3.74).
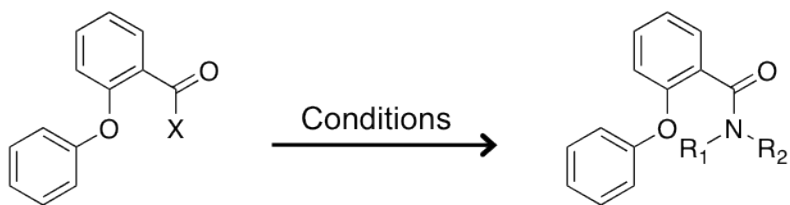
165

Table 9.

For X = OH: HNR$_1$R$_2$, EDC, HOBt, DIPEA, THF
For X = Cl: HNR$_1$R$_2$, DIPEA, DCM

Compound	R	WB[a] SW620	WB[a] H520	ICW SW620	ICW H520
161		0.96			
162		0.93			
163		0.83			
164		0.64			
165		0.74			
166		1.22			
167		0.87			
168		0.74			
169		0.63			
170		3.53	0.97		

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by Western blot for compound treatment at 10 μM. (TSA = 3.74).
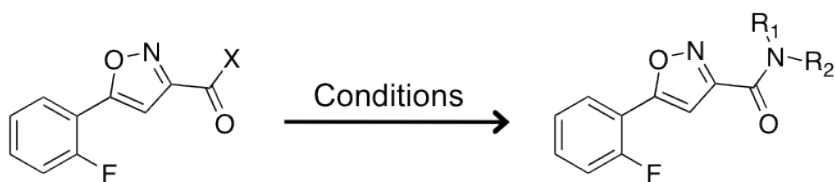
166

Table 10.

For X = OH: HNR$_1$R$_2$, EDC, HOBt, DIPEA, THF
For X = Cl: HNR$_1$R$_2$, DIPEA, DCM

Compound	R	WB[a] SW620	WB[a] H520	ICW SW620	ICW H520
171		1.56			
172		1.15			
173		0.97			
174		1.15			
175		1.33	1.10		
176		1.23			
177		1.26	1.47		
178		1.27			

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by Western blot for compound treatment at 10 μM. (TSA = 3.74).

Table 11.

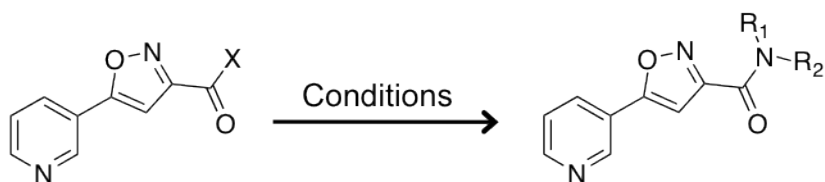

For X = OH: HNR_1R_2, EDC, HOBt, DIPEA, THF
For X = Cl: HNR_1R_2, DIPEA, DCM

Compound	R	WB[a] SW620	WB[a] H520	ICW[b] SW620	ICW[b] H520
179		1.88	2.07	1.08	
180		2.92	3.36	2.15	
181		3.78	3.26	1.91	
182		4.48	3.40	3.78	
183		3.54	3.17	3.01	
184		1.72	1.53	1.43	
185		4.20	3.54	1.50	
186		2.96	3.48	1.76	

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by Western blot for compound treatment at 10 μM. (TSA = 3.74).
[b] Data is fold change in E-cadherin restoration above a DMSO control as measured by In-Cell Western assay for compound treatment at 10 μM. (TSA = 2.37).

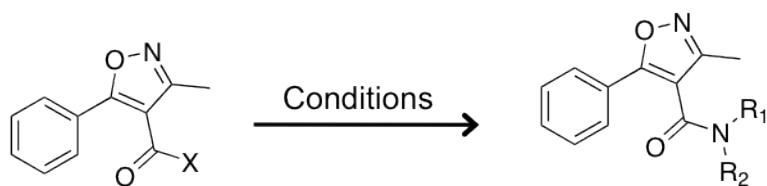**Table 12.**

For X = OH: HNR$_1$R$_2$, EDC, HOBt, DIPEA, THF
For X = Cl: HNR$_1$R$_2$, DIPEA, DCM

Compound	R	WB[a] SW620	WB[a] H520	ICW[b] SW620	ICW[b] H520
187		1.09		1.06	
188		1.08		0.94	
189		1.17		1.01	
190		1.79		1.14	
191		1.52		1.34	
192		1.48		1.19	
193		1.50		0.80	
194		1.70		0.90	

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by Western blot for compound treatment at 10 μM. (TSA = 3.74).
[b] Data is fold change in E-cadherin restoration above a DMSO control as measured by In-Cell Western assay for compound treatment at 10 μM. (TSA = 2.37).

169

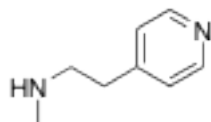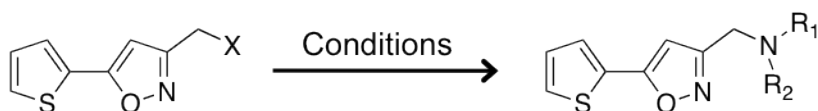**Table 13.**

For X = OH: HNR₁R₂, EDC, HOBt, DIPEA, THF
For X = Cl: HNR₁R₂, DIPEA, DCM

Compound	R	WB[a] SW620	WB[a] H520	ICW SW620	ICW H520
195		1.08			
196		1.08			
197		1.15			
198		1.01			
199		0.88			
200		1.10			
201		0.82			
202		1.22			
203		1.29			
204		1.25			
205		1.22			
206		1.30			

170

| 207 | | 1.10 |

Table 14.

For X = OH: HNR$_1$R$_2$, EDC, HOBt, DIPEA, THF
For X = Cl: HNR$_1$R$_2$, DIPEA, DCM

Compound	R	WB SW620	WB H520	ICW[a] SW620	ICW[a] H520
208				2.17	1.41
209				2.11	1.26
210				2.47	2.00
211				2.48	1.27
212				2.12	1.30
213				2.63	2.19

Compound	Structure	ICW[a] SW620	ICW[a] H520
214		1.03	1.09
215		1.46	1.11
216		1.67	1.28

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by In-Cell Western assay for compound treatment at 10 μM. (TSA = 2.37).
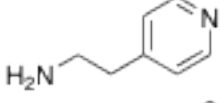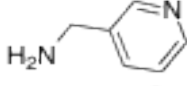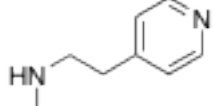
Table 15.
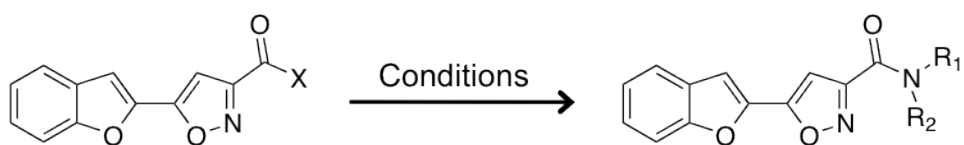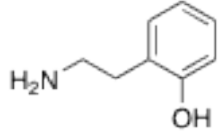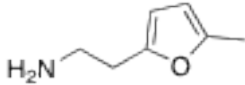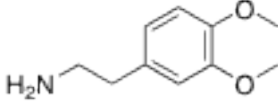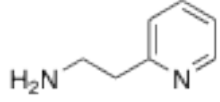

For X = OH: HNR$_1$R$_2$, EDC, HOBt, DIPEA, THF
For X = Cl: HNR$_1$R$_2$, DIPEA, DCM

Compound	R	WB SW620	WB H520	ICW[a] SW620	ICW[a] H520
217				2.10	1.55
218				2.22	1.27
219				2.35	1.18
220				2.05	1.40
221				2.02	1.60

222		0.87	1.01
223		1.63	2.12
224		1.91	1.82
225		2.08	1.42
226		2.36	1.39
227		1.61	1.18
228		1.57	1.04
229		1.29	2.02

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by In-Cell Western assay for compound treatment at 10 μM. (TSA = 2.37).
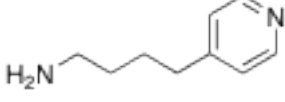
Table 16.
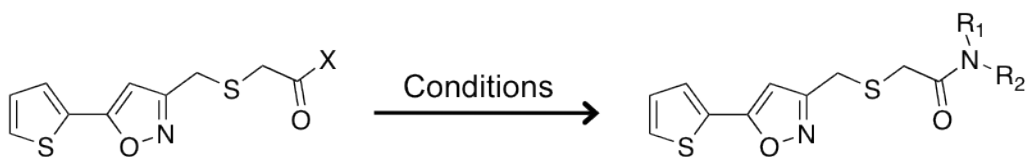
For X = OH: HNR$_1$R$_2$, EDC, HOBt, DIPEA, THF
For X = Cl: HNR$_1$R$_2$, DIPEA, DCM

Compound	R	WB SW620	WB H520	ICW[a] SW620	ICW[a] H520
230				1.50	1.19
231				0.65	1.44
232				1.44	1.51
233				1.06	1.26
234				1.12	1.22
235				0.88	1.06
236				1.24	1.08
237				1.38	1.16
238				1.27	1.42
239				0.96	1.12
240				1.08	1.28

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by In-Cell Western assay for compound treatment at 10 μM. (TSA = 2.37).
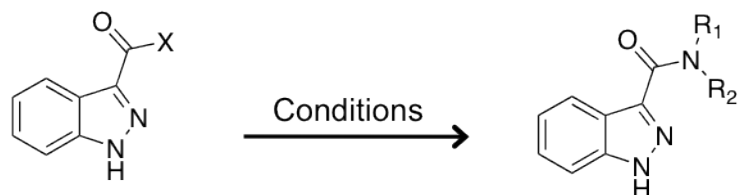
Table 17.

For X = OH: HNR$_1$R$_2$, EDC, HOBt, DIPEA, THF
For X = Cl: HNR$_1$R$_2$, DIPEA, DCM

Compound	R	WB SW620	WB H520	ICW[a] SW620	ICW[a] H520
241				1.38	1.34
242				1.41	1.05
243				0.86	0.61
244				0.79	0.41
245				1.23	1.36
246				1.37	1.24
247				0.75	0.21
248				0.89	2.32
249				1.20	1.12
250				1.15	0.97
251				0.91	0.07

Compound	R			
252		0.64	-0.38	
253		1.08	0.61	
254		1.04	0.27	

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by In-Cell Western assay for compound treatment at 10 μM. (TSA = 2.37).
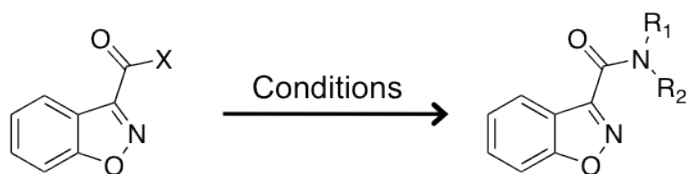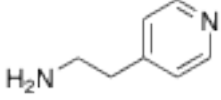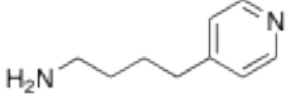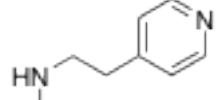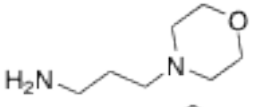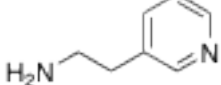
Table 18.

For X = OH: HNR$_1$R$_2$, EDC, HOBt, DIPEA, THF
For X = Cl: HNR$_1$R$_2$, DIPEA, DCM
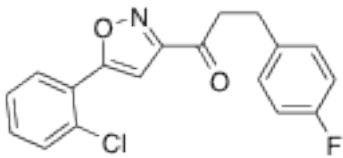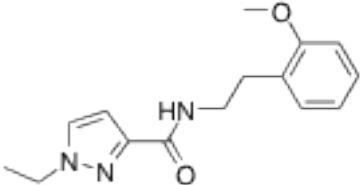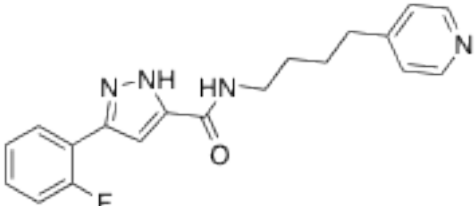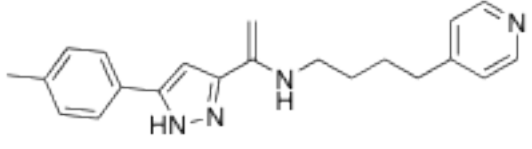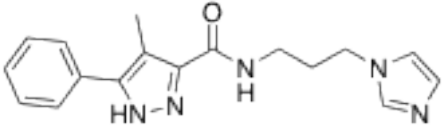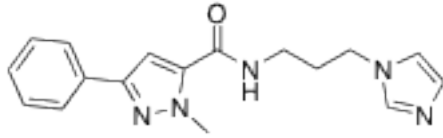
Compound	R	WB SW620	WB H520	ICW[a] SW620	ICW[a] H520
255				1.02	2.47
256				1.27	0.92
257				1.26	1.18
258				1.03	1.07
259				1.21	1.29

176

260		1.00	1.11
261		1.35	1.07
262		0.96	1.00
263		1.42	1.23
264		1.04	1.25
265		1.02	1.29
266		0.68	0.91
267		1.13	1.35
268		0.83	1.37

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by In-Cell Western assay for compound treatment at 10 μM. (TSA = 2.37).
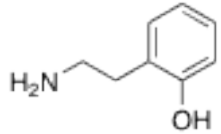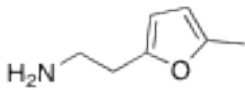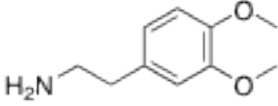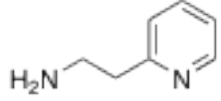
Table 19.

Compound	Compound

269
ICW[a] SW620: 0.86

270
ICW SW620: 0.91

271
ICW SW620: 1.97
ICW H520: 4.07

272
ICW SW620: 0.95
ICW H520: 2.30

273
ICW SW620: 0.87
ICW H520: 0.86

274
ICW SW620: 1.03
ICW H520: 1.06
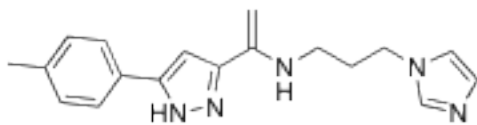
275
ICW SW620: 0.93
ICW H520: 1.13
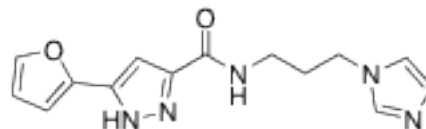
276
ICW SW620: 0.90
ICW H520: 1.38
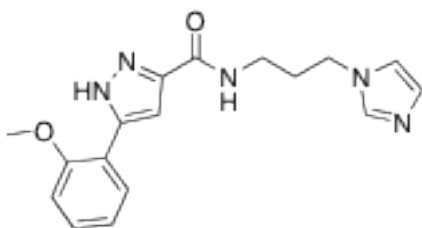
277
ICW SW620: 1.07
ICW H520: 1.41
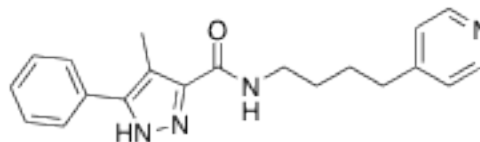
278
ICW SW620: 0.75
ICW H520: 1.05
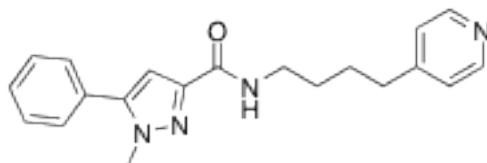
279
ICW SW620: 1.38
ICW H520: 0.67
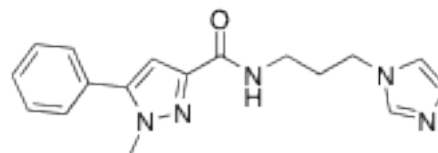
280
ICW SW620: 0.75
ICW H520: 1.65
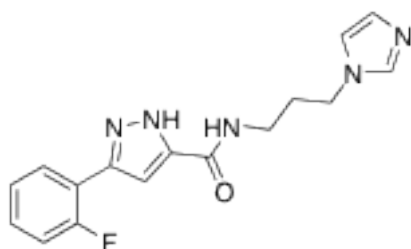
281
ICW SW620: 1.15
ICW H520: 2.04
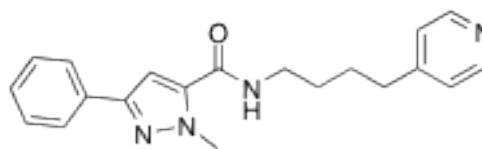
282
ICW SW620: 0.87
ICW H520: 1.31
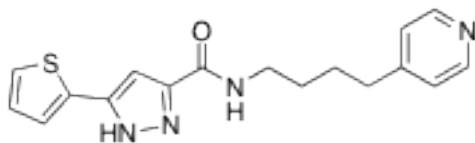
283
ICW SW620: 1.55
ICW H520: 2.82
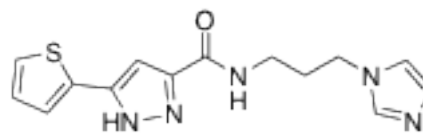
284
ICW SW620: 1.11
ICW H520: 2.21
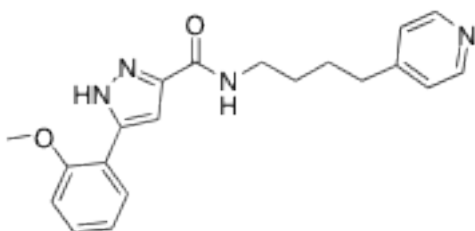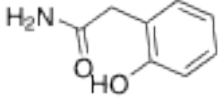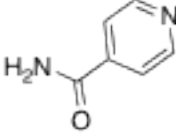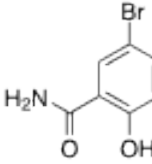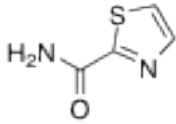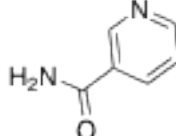
285
ICW SW620: 0.89
ICW H520: 0.98

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by In-Cell Western assay for compound treatment at 10 μM. (TSA = 2.37).

Table 20.

X = O, NH
n = 1, 2

Conditions: a) CDI, NH$_2$CH$_2$(CH$_2$)$_n$NHBOC, DCM, DIPEA; b) TFA, DCM
c) EDC, HOBT, or HAYU, DIPEA, DMF, DCM

Compound	X	n	R	ICW[a] SW620	EC$_{50}$ (µM)	ICW[a] H520	EC$_{50}$ (µM)
286	NH	1		1.37		1.24	
287	NH	1		1.45		1.04	
288	NH	1		1.77		1.10	
289	O	2		5.23	5.0	5.54	1.0
290	NH	1		1.29		1.22	
291	O	2		4.92	4.97	5.92	1.5
292	O	2		2.92	5.0	5.27	1.4
293	O	2		1.62		1.88	

181

294	O	1		3.40	3.7	3.51	5.6
295	O	1		2.79	8.8	2.87	5.1
296	O	1		1.18		1.14	
297	O	1		8.98	4.8	4.22	
298	O	1		3.05	7.1		

[a] Data is fold change in E-cadherin restoration above a DMSO control as measured by In-Cell Western assay for compound treatment at 10 μM. (TSA = 2.37).

182

Table 21. Displays the EC50 values calculated for selected analogs. Concentration response curves were developed from seven treatment concentrations from 30 to 0.1 µM using the In-Cell Western assay.

Compound	SW620 EC$_{50}$ (µM)	H520 EC$_{50}$ (µM)
1	10.6	5.4
39	5.6	
40	1.8	
41	5.8	
54	6.0	2.5
57	2.13	1.25
73	1.5	2.0
80	7.0	
81	1.5	
82	5.0	
116	4.5	1.6
289	5.0	1.0
291	4.97	1.5
292	5.0	1.4
294	3.7	5.6
295	8.8	5.1
297	4.8	
298	7.1	

APPENDIX 2

Table 1. Displays the various numbering schemed that have been used throughout the project, so that compounds can be easily identified in lab notebooks (Y/R code) and/or requested for further use if necessary (VU#).

Thesis #	VU ID	Lab Code
1	VU0075630	R21
5	VU0364639-1	1J
6	VU0364640-1	2J
7	VU0364641-1	3J
8	VU0364441-1	12Y
9	VU0152198-4	13Y
10	VU0364442-1	14Y
11	VU0364443-1	15Y
12	VU0364444-1	16Y
13	VU0364445-1	17Y
14	VU0364446-1	18Y
15	VU0364447-1	19Y
16	VU0364448-1	20Y
17	VU0364449-1	21Y
18	VU0364450-1	22Y
19	VU0364451-1	23Y
20	VU0364452-1	24Y
21	VU0075620-6	25Y
22	VU0364453-1	26Y
23	VU0364454-1	27Y
24	VU0233088-5	28Y
25	VU0364455-1	29Y
26	VU0364456-1	30Y
27	VU0364457-1	31Y
28	VU0364458-1	32Y
29	VU0365319-2	46Y
30	VU0210058-4	47Y
31	VU0405342-1	94Y
32	VU0405346-1	98Y
33	VU0229479-3	102Y
34	VU0409653-1	134Y
35	VU0409657-1	138Y

36	VU0413357-1	181Y
37	VU0413361-1	185Y
38		213Y
39	VU0449695-1	248Y
40	VU0449697-1	250Y
41	VU0449699-1	252Y
42	VU0449701-1	254Y
43	VU0142371-2	33Y
44	VU0365822-1	34Y
45	VU0365823-1	35Y
46	VU0109076-2	36Y
47	VU0285856-3	48Y
48	VU0403135-1	49Y
49	VU0403136-1	50Y
50	VU0403137-1	51Y
51	VU0403138-1	52Y
52	VU0403139-1	53Y
53	VU0326279-2	54Y
54	VU0405343-1	95Y
55	VU0405347-1	99Y
56	VU0228184-4	103Y
57	VU0409654-1	135Y
58	VU0409658-1	139Y
59	VU0413358-1	182Y
60	VU0413362-1	186Y
61		214Y
62	VU0365824-1	37Y
63	VU0365825-1	38Y
64	VU0329150-3	39Y
65	VU0403140-1	55Y
66	VU0403141-1	56Y
67	VU0180426-4	57Y
68	VU0403142-1	58Y
69	VU0405344-1	96Y
70	VU0405348-1	100Y
71	VU0180488-5	104Y
72	VU0409655-1	136Y

73	VU0409675-1	156Y
74	VU0409659-1	140Y
75	VU0180398-4	157Y
76	VU0409676-1	158Y
77	VU0413359-1	183Y
78	VU0413363-1	187Y
79		215Y
80	VU0449694-1	247Y
81	VU0449696-1	249Y
82	VU0449698-1	251Y
83	VU0449700-1	253Y
84	VU0451771-1	256Y
85	VU0451842-1	257Y
86	VU0180016-5	1Y
87	VU0180488-4	4Y
88	VU0181409-5	6Y
89	VU0362438-1	7Y
90	VU0362440-1	9Y
91	VU0362436-1	3Y
92	VU0362439-1	8Y
93	VU0293163-2	10Y
94	VU0404056-1	74Y
95	VU0404057-1	75Y
96	VU0365333-2	76Y
97	VU0364642-2	77Y
98	VU0404058-1	78Y
99	VU0405557-1	106Y
100	VU0405558-1	107Y
101	VU0365337-2	108Y
102	VU0405559-1	109Y
103	VU0365334-1	30J
104	VU0365335-1	31J
105	VU0365336-1	32J
106	VU0404063-1	88Y
107	VU0404064-1	89Y
108	VU0404065-1	90Y

109	VU0404066-1	91Y
110	VU0404067-1	92Y
111	VU0405345-1	97Y
112	VU0405349-1	101Y
113	VU0405350-1	105Y
114	VU0405560-1	110Y
115	VU0405561-1	111Y
116	VU0409656-1	137Y
117	VU0405562-1	112Y
118	VU0409660-1	141Y
119	VU0413360-1	184Y
120	VU0413364-1	188Y
121		216Y
122	VU0365327-1	23J
123	VU0365328-1	24J
124	VU0365329-1	25J
125	VU0365330-1	26J
126	VU0365332-1	28J
127	VU0451770-1	255Y
128	VU0365826-1	40Y
129	VU0365827-1	41Y
130	VU0365828-1	42Y
131	VU0365829-1	43Y
132	VU0403143-1	59Y
133	VU0403144-1	60Y
134	VU0403145-1	61Y
135	VU0403146-1	62Y
136	VU0403147-1	63Y
137	VU0403148-1	64Y
138	VU0403149-1	65Y
139	VU0365830-1	44Y
140	VU0365831-1	66Y
141	VU0075446-6	45Y
142	VU0404051-1	67Y
143	VU0404052-1	68Y
144	VU0404053-1	69Y
145	VU0186810-5	70Y

146	VU0404054-1	71Y
147	VU0404055-1	72Y
148	VU0075466-2	73Y
149	VU0409674-1	155Y
150	VU0404059-1	79Y
151	VU0404060-1	80Y
152	VU0365340-2	81Y
153	VU0404061-1	82Y
154	VU0245555-5	83Y
155	VU0404062-1	84Y
156	VU0365338-2	85Y
157	VU0365339-2	85Y
158	VU0347121-2	87Y
159	VU0364644-1	8J
160	VU0365341-1	37J
161	VU0405563-1	113Y
162	VU0405564-1	114Y
163	VU0405565-1	115Y
164	VU0364645-2	116Y
165	VU0187080-4	117Y
166	VU0240112-5	93Y
167	VU0364646-1	10J
168	VU0042509-5	11J
169	VU0364647-1	12J
170	VU0365342-1	39J
171	VU0365321-1	16J
172	VU0032890-5	17J
173	VU0365322-1	18J
174	VU0365323-1	19J
175	VU0365324-1	20J
176	VU0365325-1	21J
177	VU0365326-1	22J
178	VU0365320-1	15J
179	VU0405566-1	118Y
180	VU0405567-1	119Y
181	VU0405568-1	120Y
182	VU0405569-1	121Y

183	VU0405570-1	122Y
184	VU0405571-1	123Y
185	VU0405572-1	124Y
186	VU0405573-1	125Y
187	VU0405574-1	126Y
188	VU0405575-1	127Y
189	VU0405576-1	128Y
190	VU0405577-1	129Y
191	VU0405578-1	130Y
192	VU0405579-1	131Y
193	VU0405580-1	132Y
194	VU0405581-1	133Y
195	VU0409661-1	142Y
196	VU0409662-1	143Y
197	VU0409663-1	144Y
198	VU0409664-1	145Y
199	VU0409665-1	146Y
200	VU0409666-1	147Y
201	VU0409667-1	148Y
202	VU0409668-1	149Y
203	VU0409669-1	150Y
204	VU0409670-1	151Y
205	VU0409671-1	152Y
206	VU0409672-1	153Y
207	VU0409673-1	154Y
208	VU0413335-1	159Y
209	VU0413336-1	160Y
210	VU0413337-1	161Y
211	VU0413338-1	162Y
212	VU0413339-1	163Y
213	VU0413340-1	164Y
214	VU0413341-1	165Y
215	VU0413342-1	166Y
216	VU0413343-1	167Y
217	VU0413344-1	168Y
218	VU0413345-1	169Y
219	VU0413346-1	170Y

220	VU0413347-1	171Y
221	VU0413348-1	172Y
222	VU0413349-1	173Y
223	VU0413350-1	174Y
224	VU0413351-1	175Y
225	VU0413352-1	176Y
226	VU0413353-1	177Y
227	VU0413354-1	178Y
228	VU0413355-1	179Y
229	VU0413356-1	180Y
230		189Y
231		190Y
232		191Y
233		192Y
234		193Y
235		194Y
236		195Y
237		196Y
238		197Y
239		198Y
240		199Y
241		200Y
242		201Y
243		202Y
244		203Y
245		204Y
246		205Y
247		206Y
248		207Y
249		208Y
250		209Y
251		210Y
252		211Y
253		212Y
254		217Y
255		218Y
256		219Y

257		220Y
258		221Y
259		222Y
260		223Y
261		224Y
262		225Y
263		226Y
264		227Y
265		228Y
266		229Y
267		230Y
268		231Y
269	VU0362435-1	2Y
270	VU0362437-1	5Y
271	VU0449020-1	232Y
272	VU0449019-1	233Y
273	VU0449018-1	234Y
274	VU0449017-1	235Y
275	VU0449016-1	236Y
276	VU0449015-1	237Y
277	VU0449014-1	238Y
278	VU0449013-1	239Y
279	VU0449012-1	240Y
280	VU0449008-1	241Y
281	VU0449009-1	242Y
282	VU0449021-1	243Y
283	VU0449022-1	244Y
284	VU0449010-1	245Y
285	VU0449011-1	246Y
286	VU0452047-1	258Y
287	VU0452036-1	264Y
288	VU0452982-1	269Y
289	VU0452046-1	259Y
290	VU0452952-1	266Y
291	VU0452045-1	260Y
292	VU0452033-1	261Y
293	VU0452035-1	263Y

294	VU0452034-1	262Y
295	VU0452953-1	267Y
296	VU0452951-1	265Y
297	VU0452954-1	268Y
298	VU0452983-1	270Y

* Some compounds do not have VU# as they were not registered through the system at Vanderbilt University. These compounds can be found in a freezer in the lab or easily synthesized using the protocol provided in Chapter 2 and/or Appendix Table for which they are located.