2015-12-14

Antisense RNA-Mediated Epigenetic Regulation of Brain-Derived Neurotrophic Factor.

Roya Pedram Fatemi
University of Miami, roya.pedram@gmail.com

Follow this and additional works at: http://scholarlyrepository.miami.edu/oa_dissertations

Recommended Citation
http://scholarlyrepository.miami.edu/oa_dissertations/1553

This Open access is brought to you for free and open access by the Electronic Theses and Dissertations at Scholarly Repository. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of Scholarly Repository. For more information, please contact repository.library@miami.edu.
UNIVERSITY OF MIAMI

ANTISENSE RNA-MEDIATED EPIGENETIC REGULATION OF BRAIN-DERIVED NEUROTROPHIC FACTOR

By

Roya Pedram Fatemi

A DISSERTATION

Submitted to the Faculty of the University of Miami in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Coral Gables, Florida

December 2015
UNIVERSITY OF MIAMI

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

ANTISENSE RNA-MEDIATED EPIGENETIC REGULATION OF BRAIN-DERIVED NEUROTROPHIC FACTOR

Roya Pedram Fatemi

Approved:

Peter Buchwald, Ph.D.
Associate Professor of Molecular and Cellular Pharmacology

Claes Wahlestedt, M.D., Ph.D.
Professor of Psychiatry and Behavioral Sciences

Vladlen Z. Slepak, Ph.D.
Professor of Molecular and Cellular Pharmacology

Arun Malhotra, Ph.D.
Associate Professor of Biochemistry and Molecular Biology

Mohammad Ali Faghihi, M.D., Ph.D.
Assistant Professor of Psychiatry and Behavioral Sciences

Dean of the Graduate School

Sapna Deo, Ph.D.
Associate Professor of Biochemistry and Molecular Biology
Long noncoding RNAs (lncRNAs) regulate chromatin remodeling through their interactions with epigenetic enzymes during development and disease. The inhibition of the natural antisense transcript of Brain-derived neurotrophic factor (BDNF-AS), results in BDNF promoter de-repression and transcriptional upregulation, both in vitro and in vivo. Recently, we showed that BDNF-AS interacts with the histone methyltransferase enhancer of zeste homolog 2 (EZH2) to suppress BDNF mRNA and protein expression. BDNF is an important neurotrophin that is required for neural development and maintenance of the nervous system. Dysregulation of BDNF occurs in a number of neurological disorders, including: Alzheimer’s Disease, Parkinson’s Disease, Rett syndrome, and amyotrophic lateral sclerosis. Previous attempts to upregulate BDNF by administering the recombinant form in various parts of the central nervous system have failed, mostly due to the challenge of delivering BDNF to the correct cells and neural networks. Our approach to upregulating BDNF by modulating its interaction with an epigenetic enzyme is a highly specific target with potential therapeutic value. To achieve this, we developed a novel pharmacological assay to characterize the interaction between long noncoding RNAs and their epigenetic targets using Amplified Luminescent
Proximity Homogeneous Assay (AlphaScreen) technology. With this assay, we are able to quantify lncRNA-protein interactions rapidly for the purpose of high throughput screening, enabling drug discovery efforts for this novel class of drug targets. In this work, we present our assay development and screening findings, including the identification of potential small molecule modulators of lncRNA-protein interactions. Furthermore, we describe the application of this lncRNA-protein interaction assay to detect RNA requirements for EZH2 recruitment, a much debated and important question that lingers in the field. From our work, it is evident that BDNF-AS has several regions of RNA that are required for EZH2 recruitment, potentially due to the importance of this transcript in regulating BDNF. This work describes exploratory drug discovery for a novel class of drug targets as well as applications to understand the basic biochemistry governing lncRNA-protein interactions.
DEDICATION

This thesis is dedicated to my amazing parents, loving husband and two beautiful children, for giving me support, love, strength, motivation and for providing a constant source of joy and laughter in my life.
ACKNOWLEDGEMENTS

I would like to thank my wonderful mentors, Dr. Mohammad Faghihi and Dr. Claes Wahlestedt for their mentorship, support, patience, and for their trust in my abilities. They gave me freedom to learn and guidance when I asked for it and taught me many valuable lessons. Thank you to my thesis committee, Dr. Peter Buchwald (chair), Dr. Arun Malhotra, Dr. Vladlen Slepak and Dr. Sapna Deo for taking a genuine interest in my work and devoting their valuable time to helping me develop as a scientist.

The wonderful members of the Faghihi and Wahlestedt labs have been a tremendous source of guidance and mentorship, giving me valuable insights and much technical and professional guidance. I am truly grateful to Dr. Claude-Henry Volmar and Dr. Shaun Brothers for making themselves available to help me with all aspects of my work whenever I needed it.

Thank you to the wonderful people in the lab for being so smart, helpful and generous with their time and for being a great group to work with. Thank you also for years of friendship, laughter, and kindness. I am indebted to each and every person for making me a better scientist.

Thanks to my fellow classmates for being a source of support, a shoulder to lean on, and for years of great conversations and laughter getting through the difficult task of getting a Ph.D.

Many thanks to all of my teachers who always had faith in me, encouraged me and took a genuine interest in my success.
Thank you to my family—immediate and extended, for years of love, support, guidance, and happiness. You have been there for me every step of the way and I love you for it.

Finally, my work builds on the knowledge and hard work of all the researchers who came before me. A career in science is challenging and I am truly thankful to those who took this difficult road before me and the colleagues who share the same curiosities as myself whose work I have built upon. I am eternally grateful to stand on the shoulders of the many giants who have come before me.
# TABLE OF CONTENTS

LIST OF FIGURES........................................................................................................................................viii

LIST OF TABLES.............................................................................................................................................ix

PUBLICATION NOTE........................................................................................................................................x

Chapter 1. Introduction......................................................................................................................................1

1.1 Long noncoding RNAs (lncRNAs) add complexity to the human genome............................................1

1.2 Functions of long noncoding RNAs.......................................................................................................2

1.3 Long noncoding RNAs as epigenetic regulators.......................................................................................3

1.4 Long noncoding RNAs as potential therapeutic targets........................................................................4

1.4.1 Cancer...............................................................................................................................................5

1.4.2 Heart Disease.....................................................................................................................................7

1.4.3 Neurological disorders.......................................................................................................................8

1.5 Regulation of brain-derived neurotrophic factor (BDNF) transcription by an antisense RNA-mediated epigenetic mechanism.................................................................9

1.6 Polycomb Repressive Complex 2 (PRC2) and long noncoding RNAs.................................................10

1.7 BDNF therapeutics................................................................................................................................11

Chapter 2. Methods....................................................................................................................................13

2.1 In vitro transcription and biotinylation of RNA.....................................................................................13

2.2 RNA electrophoretic mobility shift assay (RNA EMSA)....................................................................14

2.3 LncRNA-EZH2 interaction AlphaScreen assay....................................................................................15

2.4 High throughput screen (HTS) of phytochemical compound library................................................17

2.5 HEK293 cell culture and drug treatment...............................................................................................17

2.6 RNA extraction, cDNA synthesis and quantitative real-time PCR.....................................................17

2.7 Enzyme linked immunosorbent assay....................................................................................................18

2.8 Software modeling of lncRNA structure...............................................................................................18

Chapter 3. Development of an assay to quantify lncRNA-protein interactions............................................19

3.1 Summary................................................................................................................................................19

3.2 Results..................................................................................................................................................19

3.2.1 BDNF-AS interacts directly with EZH2..........................................................................................19

3.2.2 LncRNA-protein interaction assay optimizations...........................................................................20
3.2.3 AlphaScreen detects BDNF-AS-EZH2 binding ........................................ 24
3.2.4 AlphaScreen detects HOTAIR-EZH2 binding ........................................ 25
3.2.5 Optimization for HTS ........................................................................ 26
3.2.6 HTS of phytochemical library ................................................................ 27
3.2.7 Secondary assays: BDNF-AS-EZH2 ...................................................... 31
3.2.8 Secondary assays: HOTAIR-EZH2 ......................................................... 34

3.3 Discussion .............................................................................................. 35
3.4 Future directions .................................................................................... 38

Chapter 4. RNA requirement for EZH2 interactions ...................................... 39
4.1 Summary .................................................................................................. 39
4.2 Results .................................................................................................... 39
  4.2.1 AlphaScreen measures HOTAIR-protein interactions .......................... 39
  4.2.2 BDNF-AS RNA requirement for EZH2 binding ................................. 42
4.3 Discussion .............................................................................................. 46
4.4 Future directions .................................................................................... 47

Chapter 5. Discussion .................................................................................... 48
  5.1 Designing IncRNA-directed therapeutics ............................................... 48
  5.2 Small molecule modulators of IncRNA-protein interactions .................. 51
  5.3 Conclusions .......................................................................................... 52

REFERENCES ................................................................................................. 53
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Mechanism of the BDNF-AS - EZH2 interaction</td>
<td>20</td>
</tr>
<tr>
<td>3.2</td>
<td>Schematic of the AlphaScreen adapted to quantify lncRNA-protein interactions</td>
<td>21</td>
</tr>
<tr>
<td>3.3</td>
<td>Test RNAs are <em>in vitro</em> transcribed and biotin-labeled</td>
<td>22</td>
</tr>
<tr>
<td>3.4</td>
<td>AlphaScreen assay conditions were optimized to maximize assay signal</td>
<td>23</td>
</tr>
<tr>
<td>3.5</td>
<td>Determination of long noncoding RNA and protein concentrations for screening</td>
<td>24</td>
</tr>
<tr>
<td>3.6</td>
<td>AlphaScreen bead concentrations were optimized for screening</td>
<td>25</td>
</tr>
<tr>
<td>3.7</td>
<td>RNA controls for the lncRNA-protein interaction AlphaScreen</td>
<td>25</td>
</tr>
<tr>
<td>3.8</td>
<td>Concentration-dependence of BDNF-AS and HOTAIR binding to EZH2</td>
<td>26</td>
</tr>
<tr>
<td>3.9</td>
<td>DMSO tolerance of lncRNA-protein interaction assay</td>
<td>28</td>
</tr>
<tr>
<td>3.10</td>
<td>Scatterplot of values from the phytochemical compounds screened for the inhibition of BDNF-AS - EZH2</td>
<td>29</td>
</tr>
<tr>
<td>3.11</td>
<td>Scatterplot of values from the phytochemical compounds screened for the inhibition of HOTAIR- EZH2</td>
<td>30</td>
</tr>
<tr>
<td>3.12</td>
<td>Chemical structure of BDNF-AS - EZH2 primary screen hits</td>
<td>32</td>
</tr>
<tr>
<td>3.13</td>
<td>Validated hits from primary screen of BDNF-AS- EZH2</td>
<td>33</td>
</tr>
<tr>
<td>3.14</td>
<td>Secondary assays to test the effect of ellipticine treatment on BDNF transcription and translation</td>
<td>33</td>
</tr>
<tr>
<td>3.15</td>
<td>Chemical structure of HOTAIR- EZH2 primary screen hits</td>
<td>34</td>
</tr>
<tr>
<td>3.16</td>
<td>Validated hits from the primary screen of HOTAIR- EZH2</td>
<td>35</td>
</tr>
<tr>
<td>4.1</td>
<td>The lncRNA-protein interaction AlphaScreen assay was validated with HOTAIR</td>
<td>41</td>
</tr>
<tr>
<td>4.2</td>
<td>Modeling the structure of BDNF-AS</td>
<td>44</td>
</tr>
<tr>
<td>4.3</td>
<td>BDNF-AS fragments were tested for their ability to bind EZH2</td>
<td>45</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table 1.1</th>
<th>LncRNAs have been implicated in a number of diseases.</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 3.1</td>
<td>Assay parameters calculated to determine day-to-day variation in assay signal and HTS suitability.</td>
<td>27</td>
</tr>
<tr>
<td>Table 3.2</td>
<td>List of initial hits from the primary screen for BDNF-AS- EZH2</td>
<td>31</td>
</tr>
<tr>
<td>Table 3.3</td>
<td>List of initial hits from the primary screen for HOTAIR- EZH2</td>
<td>34</td>
</tr>
</tbody>
</table>
PUBLICATION NOTE

Chapter 1

Versions of material in this chapter have been published in:


Chapters 2-5

Versions of material in this chapter have been published in:


Chapter 1. Introduction

1.1 Long noncoding RNAs (IncRNAs) add complexity to the human genome

For many years, a rigid view of the central dogma of molecular biology prevailed and the idea that RNA simply acts as an intermediary that is faithfully translated into protein remained uncontested. However, findings from the Human Genome Project indicate that humans do not possess substantially greater genes than lower organisms as previously thought. This ‘g-value paradox’ describes the disproportionately low number of human protein coding genes as compared to lower organisms. Large-scale transcriptomics efforts launched following the Human Genome Project, notably: Functional Annotation of the Mouse (FANTOM) and Encyclopedia of DNA Elements (ENCODE) have yielded unexpected results and have helped to make sense of the ‘transcriptional noise’ that was thought to comprise an overwhelming majority of the human genome. These studies show that the genome is pervasively transcribed and that many genes are transcribed from both sense and antisense strands of DNA, resulting in protein products, non-protein coding transcripts, or combinations of the two depending on the gene locus.

Non-protein coding RNA transcripts (ncRNAs) are being recognized as functionally important RNAs that regulate a number of protein products. Noncoding RNAs are divided into two broad classes based on length: (i) small regulatory ncRNAs that are less than 200 nucleotides long, e.g. small nuclear RNAs, piwi-interacting RNAs, micro RNAs, and small interfering RNAs and (ii) long noncoding RNAs (IncRNAs) that range from 200 base pairs up to 100 kilobases, e.g. natural antisense transcripts (NATs) and long intergenic noncoding RNAs (lincRNAs). IncRNAs are RNA Polymerase II...
transcripts that may or may not have a 5’ cap, be polyadenylated, and spliced similar to messenger RNAs. Recent evidence and multiple examples show that many of these transcripts are evolutionarily conserved in position and expression patterns across vertebrate genomes \(^8,^{15-17}\). Furthermore, lncRNA promoters are tightly regulated to ensure cell-, tissue- and development-specific expression \(^10,^{11,17-23}\).

NATs are transcribed from the opposite strand as that of the protein coding messenger RNA (typically the sense strand) to which they are complementary and have been shown to perform regulatory functions in cells by altering transcription, mediating RNA-DNA interactions or by modifying RNA-RNA interactions in the nucleus and cytoplasm \(^{24-26}\). Furthermore, NATs can function through concordant or discordant regulatory mechanisms. In concordant regulation, the downregulation of the antisense transcript results in a reduction in sense transcript expression. In discordant regulation, the down-regulation of the noncoding transcript results in the subsequent upregulation of the cognate sense transcript \(^7\). Although our understanding of the diverse functions of NATs remains incomplete, to date, they have been implicated in regulating gene expression at the transcriptional and post-transcriptional level in various development and disease contexts \(^{27}\).

### 1.2 Functions of long noncoding RNAs

LncRNAs have been separated into several broad classes in terms of their mechanisms of regulation of mRNA transcription and translation: decoys, regulators of translation, enhancers and modular scaffolds that guide chromatin modifying enzymes to specific genomic loci. LncRNAs can regulate apoptosis and the cell cycle; the lncRNA Growth arrest-specific 5 (GAS5) is induced during growth factor starvation and binds the
glucocorticoid receptor through its hairpin domain, acting as a decoy that physically blocks the transcription of metabolic genes \(^{28,29}\). The lncRNA metastasis-associated lung adenocarcinoma transcript 1 (MALAT-1) localizes to nuclear speckles to associate with serine/arginine-rich (SR) proteins involved in splicing \(^{30,31}\). Another important function of lncRNAs is to regulate translation, e.g. beta-site APP cleaving enzyme 1-antisense (BACE1-AS), physically masks miRNA-binding sites on its cognate mRNA BACE1, forming a stable duplex to prevent mRNA degradation \(^{32,33}\). LncRNAs can also enhance the expression of neighboring genes, for example, ncRNAs arising from the T-cell acute lymphocytic leukemia 1 (SCL/TAL1) and Snail family zinc finger 1 and 2 (SNAI1 and SNAI2) loci act as positive regulators of gene expression \(^{34,35}\).

### 1.3 Long noncoding RNAs as epigenetic regulators

LncRNAs are believed to confer specificity to ubiquitously expressed epigenetic complexes based on the observation that the enzymatic subunits of these complexes contain RNA but not DNA binding domains. Furthermore, a number of molecular studies have shown that lncRNAs interact with chromatin remodeling complexes. Specifically, lncRNAs function as modular scaffolds to position protein complexes to certain genomic loci \(^{36-39}\), resulting in changes in gene expression \emph{in cis} (at the same gene locus) \(^{40}\) or \emph{in trans} (at different loci) \(^{41}\) to alter chromatin. Chromatin has classically been described as having two main states: the compact, enzymatically inaccessible heterochromatin state and euchromatin that has a loose, open structure ripe for transcription. The histone core of chromatin consists of eight proteins that wrap 147 base pairs of DNA to form a nucleosome, the basic structural component of chromatin. Histone proteins can be chemically modified at their histone tails to alter chromatin structure and DNA
accessibility depending on the type of amino acid and the post-translational modification that is deposited \(^{42-45}\). Multiple reports have shown that antisense RNAs interact with epigenetic enzymes to form RNA-protein complexes that can interact with histones, DNA, or other RNAs and chromatin modifying enzymes \(^{21,37,40,46}\).

The most extensively studied lncRNA is arguably the X-inactive specific transcript (\(Xist\)), a roughly 17 kb mammalian-specific lncRNA that regulates X-chromosome inactivation (through an epigenetic mechanism that involves interactions with polycomb proteins), by silencing one of the two X-chromosomes in females \(^{47,48}\). LncRNAs are also involved in imprinting; potassium channel, voltage gated KQT-like subfamily Q, member 1 antisense transcript (\(Kcnq1ot1\)) and acute insulin response 2 (\(Air\)) act as scaffolds to recruit chromatin-modifying enzymes. The 91-kb lncRNA \(Kcnq1ot1\) interacts with two chromatin modifying enzymes: histone-lysine N-methyltransferase EHMT2 (G9a) and Polycomb repressive complex-2 (PRC2) during imprinting to produce lineage-specific transcriptional silencing patterns \(^{49-51}\). The lncRNA \(Air\) silences gene expression at the Insulin-Like Growth Factor 2 Receptor/ Solute Carrier Family 22 (Organic Cation Transporter), Members 2 and 3 (\(Igf2r/Slc22a2/Slc22a3\)) locus of paternal alleles by targeting G9a to chromatin \(^{52,53}\).

**1.4 Long noncoding RNAs as potential therapeutic targets**

The potential role of long noncoding RNAs in various disease contexts is currently under investigation. The evidence observed thus far suggests that these transcripts play a role in the switch from normal to disease state (**Table 1.1**). Knowledge of the diverse mechanisms of action of lncRNAs and their role in disease progression provides us with the opportunity to explore their unique therapeutic potential.
Table 1.1 *LncRNAs have been implicated in a number of diseases.*

<table>
<thead>
<tr>
<th>Transcript</th>
<th>Target</th>
<th>Mechanism</th>
<th>Disease Relevance</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAS5</td>
<td>Glucocorticoid receptor (GR)</td>
<td>Decoy</td>
<td>Cancer, Autoimmune disease</td>
<td>29, 30</td>
</tr>
<tr>
<td>APOA1-AS</td>
<td>APOA1</td>
<td>Scaffold, Guide</td>
<td>Cardiovascular disease</td>
<td>82</td>
</tr>
<tr>
<td>BDNF-AS</td>
<td>BDNF</td>
<td>Scaffold, Guide</td>
<td>Neurodegenerative disorders</td>
<td>41</td>
</tr>
<tr>
<td>H19</td>
<td>E-cadherin</td>
<td>Scaffold, Guide</td>
<td>Cancers</td>
<td>76-78</td>
</tr>
<tr>
<td>HOTAIR</td>
<td>HOXD 8-11, 13</td>
<td>Scaffold, Guide</td>
<td>Cancers</td>
<td>59-68</td>
</tr>
<tr>
<td>KCNO1OT1</td>
<td>KCNO1</td>
<td>Scaffold, Guide</td>
<td>Beckwith-Wiedemann Syndrome, adrenal neuroblastoma</td>
<td>45, 46</td>
</tr>
<tr>
<td>ANRIL</td>
<td>Ink4b/ARF/INK4a</td>
<td>Scaffold, Guide</td>
<td>Cancers</td>
<td>73, 72</td>
</tr>
<tr>
<td>PCAT-1</td>
<td>BRCA2, CENPF, CENPE</td>
<td>Scaffold, Guide</td>
<td>Prostate cancer</td>
<td>73</td>
</tr>
<tr>
<td>PINT</td>
<td>p53 pathway genes</td>
<td>Scaffold, Guide</td>
<td>Cancers</td>
<td>79</td>
</tr>
<tr>
<td>PTENq1q2-a isoform</td>
<td>PTEN</td>
<td>Scaffold, Guide</td>
<td>Cancers</td>
<td>75</td>
</tr>
<tr>
<td>TUG1</td>
<td>HOX87</td>
<td>Scaffold, Guide</td>
<td>Cholangiocarcinoma</td>
<td>80</td>
</tr>
</tbody>
</table>

LncRNAs can regulate a number of gene targets and are involved in the switch from normal development to disease state.

1.4.1 Cancer

Hox antisense intergenic RNA (*HOTAIR*) is a 2.2kb lncRNA transcribed from the homeobox C cluster (*HOXC*) locus that negatively regulates the expression of *HOXD* body patterning genes in *trans*. *HOTAIR* interacts with PRC2 to silence genes by depositing repressive histone 3 lysine 27 trimethyl marks (H3K27me3) on the *HOXD* locus. Importantly, *HOTAIR* is overexpressed in metastatic tumors, in particular, breast, ovarian, esophageal, gastric as well as a number of other cancers. The current understanding, at least in breast cancer, is that *HOTAIR* overexpression promotes tumor metastasis by re-targeting PRC2 to metastasis suppressor genes. Furthermore, knockdown of *HOTAIR* inhibits cancer invasiveness, highlighting its potential as a target for cancer therapeutics.

The lncRNA Cyclin-Dependent Kinase Inhibitor 2B antisense RNA 1 (*ANRIL*) was first identified as an antisense transcript overlapping the tumor suppressor locus that encodes for two cyclin-dependent kinase inhibitors, p15*INK4b* and p16*INK4a*, and a regulator of the p53 pathway, p14*ARF* (*INK4b/ARF/INK4a*). *ANRIL* negatively regulates...
its cognate transcript, INK4b, in cis, by an epigenetic mechanism that induces heterochromatin formation in melanoma-neural system tumor syndrome. Prostate cancer tissues have significantly elevated levels of ANRIL and CBX7 (a chromobox protein (CBX) that is part of the polycomb repressive complex 1 (PRC1)) that correlates with a significant reduction in INK4a. ANRIL associates with CBX7 and point mutations of the CBX7 chromodomain have been shown to disrupt lncRNA binding and H3K27me3, resulting in de-repression of the INK4b/ARF/INK4a locus.

Some lncRNAs are unique to specific diseases, e.g. prostate and colorectal cancers, while others are overexpressed in several cancers. The prostate cancer-associated lncRNA transcript-1 (PCAT-1) is overexpressed solely in metastatic prostate cancer and knockdown of this transcript reduces proliferation in prostate cancer cells. Suppressor of zeste 12 homolog (SUZ12), a component of the repressive PRC2 complex, binds to the PCAT-1 promoter to suppress transcription of metastatic genes. Therefore, targeting PCAT-1 could de-repress genes that help prevent prostate cancer progression.

The important tumor suppressor gene phosphatase and tensin homolog gene (PTEN) is silenced in a number of cancers and is regulated in part by antisense lncRNA transcripts that arise from the PTEN pseudogene (PTENpg1) locus. The a isoform of PTENpg1 RNA functions in trans and interacts with enhancer of zeste homolog 2 (EZH2) and DNA methyltransferase 3a (DNMT3a) to negatively regulate the expression of PTEN, making the a isoform of PTENpg1 a potential target in several cancers. The imprinted maternally expressed H19 lncRNA positively regulates tumor growth via a PRC2 dependent mechanism. H19 physically interacts with EZH2 to reduce E-cadherin and promote metastasis in oral tongue squamous cell carcinoma and bladder cancer.
The 7.1 kb lncRNA taurine upregulated 1 (TUG1) is induced by p53 and interacts with PRC2 to repress 120 genes, most of which are involved in cell cycle regulation. Interestingly, one of the genes repressed by TUG1 is homeobox B7 (HOXB7), a known oncogene. Similarly, the p53-induced transcript (PINT) lncRNA acts as a tumor suppressor that reduces cell proliferation by regulating the expression of genes involved in p53 signaling via a PRC2-dependent mechanism. These p53-associated lncRNAs silence oncogenic and cell proliferation-promoting genes. Therefore, these lncRNAs have the potential to be exploited for their potential therapeutic function in regulating p53 signaling. The many examples of lncRNAs implicated in cancer suggests that these transcripts play an important role in disease development and progression and thus their functions and potential as therapeutic targets warrant further study.

1.4.2 Heart disease

Apolipoprotein A-1 (APOA-1) is the major protein component of high-density lipoprotein (HDL) which has been shown to protect against coronary heart disease. APOA1-AS regulates APOA1 in cis, and knockdown of APOA1-AS in vivo significantly increases endogenous APOA1, making APOA1-AS a potential therapeutic target for cardiovascular disease. A novel lncRNA, Braveheart (Bvht), was identified to be important in specifying lineage commitment in mice via a PRC2-dependent mechanism. Bvht interacts with SUZ12 during multiple stages of cardiomyocyte differentiation and targeting this interaction might be a means by which to promote cardiomyocyte differentiation in stem cells.
1.4.3 Neurological disorders

Neurodegenerative disorders are characterized by the progressive loss of specific neuronal subtypes in the central nervous system (CNS), resulting in a number of devastating diseases, e.g., Alzheimer’s disease (AD)\textsuperscript{80,81}, Parkinson’s disease (PD)\textsuperscript{82} and amyotrophic lateral sclerosis (ALS), whose patients display significant cognitive, motor and behavioral deficits. Neurotrophins have been shown to play an important role in neuronal survival, outgrowth, and differentiation\textsuperscript{83}. Deficits in brain-derived neurotrophic factor (BDNF)\textsuperscript{84} and depletion of ephrin B2 receptor (EPHB2) through proteosomal degradation\textsuperscript{85} have been observed in AD. Ephrin B2 ligand (EFNB2) and neurotrophic tyrosine kinase receptor type 3 (NTRK3) are proteins involved in signaling pathways whose downstream effectors have an impact on neuronal health, making them promising therapeutic candidates. Interestingly, antisense transcripts to BDNF, EPHB2, EFNB2, and NTRK3 have been identified. It is estimated that there are over 20,000 brain-expressed long noncoding transcripts that exhibit cell-type specific spatial and temporal expression patterns in the adult mouse brain, suggesting that the antisense transcripts of these neurotrophins are potentially involved in the modulation of sense transcription\textsuperscript{18}.

An analysis of genes both enriched in the brain and implicated in autism revealed that 40% of disease genes had overlapping antisense transcription\textsuperscript{86}. An antisense transcript to \textit{SYNGAP1} (\textit{SYNGAP1}-AS), was differentially expressed in several brain regions of autism patients compared to healthy control subjects\textsuperscript{86}. \textit{SYNGAP1} is important for learning, memory and cognition and its dysregulation has been implicated in autism\textsuperscript{87}.

Gene repression by lncRNAs is prevalent in the brain and occurs at several gene loci whose protein products are involved in nervous system development and maintenance as
well as synapse formation and maturation. As potent regulators of transcription, inhibition of these transcripts could result in locus-specific upregulation of target genes.

1.5 Regulation of brain-derived neurotrophic factor (BDNF) transcription by an antisense RNA-mediated epigenetic mechanism

BDNF, one of the most abundant growth factors in the brain, is critical for the development, survival and maintenance of neurons in the nervous system. Furthermore, BDNF is severely downregulated in a number of neurodegenerative disorders. The BDNF locus gives rise to a 2.2 kb antisense transcript, BDNF-AS, that discordantly regulates BDNF mRNA and protein levels in cis. Blocking or knockdown of BDNF-AS correlates with a reduction in H3K27me3 at the BDNF promoter region and a reduction in EZH2 occupancy, suggesting that BDNF-AS exerts its repressive effect on the BDNF locus by interacting with PRC2 to silence BDNF expression in cis. A follow-up study showed that BDNF-AS interacts directly with EZH2, suggesting that BDNF-AS or its site of interaction with PRC2 could be viable therapeutic targets to upregulate BDNF. This mechanism of specific gene upregulation might be extended to other neuroprotective genes, e.g. glial-derived neurotrophic factor (GDNF) and ephrin B2 receptor (EPHB2), that were also shown to be discordantly regulated by their cognate antisense transcripts. Importantly, one of the first examples of in vivo endogenous gene upregulation was described by blocking mouse Bdnf-AS. These findings have major implications in treating neurodegenerative and neurodevelopmental diseases where BDNF is dysregulated.
1.6 Polycomb Repressive Complex 2 (PRC2) and long noncoding RNAs

PRC2 is an epigenetic enzyme composed of four core subunits: EZH2, SUZ12, embryonic ectoderm development (EED), and retinoblastoma binding protein 48 (RBP48). PRC2 has been shown to act as a dimer that binds a single RNA transcript to promote the recruitment of other PRC2 complexes to cooperatively nucleate gene repression. It is postulated that if a chromatin recruitment event does not occur, that RNA functions as a decoy to prevent PRC2 from interacting with active genes \(^94\).

EZH2 is the catalytic domain of the PRC2 complex and contains the SET domain that has histone methyltransferase activity \(^95\). EZH2 is a 746 amino acid long protein that is able to modify histone tails to deposit a repressive epigenetic mark (H3K27me3) at gene promoters. The crystal structure of EZH2 has not been solved, however the crystal structure of the SET domain is known. EZH2 is fairly disordered and has several domains, including: a) the SET domain (amino acids 612-727), b) the CXC domain (a stretch of evolutionarily conserved cysteine residue repeats unique to SET proteins, amino acids 503-605), and c) a putative noncoding RNA binding domain (amino acids 342-370) \(^95,96\). Although not catalytic, association with EED is crucial for RNA recognition and SUZ12 is required for histone methyltransferase activity \(^94,97,98\). One advantage in targeting EZH2 is that because the ncRBD is separate and distinct from the SET domain, it is possible to modify EZH2 function without permanently altering its enzymatic activity.

The binding of RNAs to EZH2 has been reported to be nonspecific in nature, however, some reports indicate that EZH2 can bind specific RNAs with much greater affinity \(^37,48,99,100\). Studies indicate that EZH2 does interact with several important long
noncoding RNAs to help regulate important cellular processes. For example, upon its interaction with Xist, EZH2 helps initiate X-chromosome inactivation, an important event in early fetal development. EZH2 also interacts with Kcnq1ot1, a lncRNA that is involved in maternal imprinting of a voltage-gated potassium channel linked to Beckwith-Wiedemann syndrome.

1.7 BDNF therapeutics

Previous research has focused on upregulating BDNF, however, these efforts have been met with limited success. In diseases such as schizophrenia and depression, there is an indirect effect on BDNF levels and so direct changes in BDNF levels might not be effective, especially since the mechanisms behind BDNF alterations are unclear. Furthermore, it is difficult to target BDNF, mainly because BDNF levels are tightly regulated within the neurons and neural circuits involved in learning and memory. Importantly, different diseases exhibit brain-region-specific alterations in BDNF, therefore increasing BDNF in the CNS in a non-specific fashion could result in neuronal excitotoxicity, impaired learning and memory and have other damaging side effects, e.g. tumor growth and cardiovascular disease. As such, therapeutics must be carefully targeted such that BDNF is only delivered to localized regions in the CNS and periphery. One report shows the feasibility of BDNF as a therapeutic in Huntington’s Disease. Knock-in of the BDNF receptor tropomyosin receptor kinase B (TrkB) in striatal cells, coupled with recombinant BDNF administration reduced apoptosis and increased neural survival. This report highlights the potential of recombinant BDNF delivery in a specific therapeutic context.
Administration of recombinant BDNF in humans systemically has not been successful in part because of difficulties determining dosage (it is unclear how much BDNF is delivered to neurons and how much is blocked by the blood-brain barrier), poor *in vivo* half-life and limited efficacy \(^{84,105}\). High doses of BDNF reduced ALS phenotypes for patients in the advanced stages of the disease \(^{106}\) and resulted in a series of clinical trials where recombinant BDNF was administered intrathecally in the spinal cord of ALS patients \(^{104,107,108}\). Despite initial reports of detectable increases in BDNF in the cerebrospinal fluid, later trials with smaller numbers of patients did not reproduce these findings \(^{104,107}\). Intranasal delivery of BDNF has also been attempted and it is reported that there is an increase in this neurotrophin following treatment, however these reports offer limited data in rodents and have not been substantiated in humans. \(^{109-111}\)

To overcome limitations in the delivery of recombinant BDNF, recent efforts have focused on mimicking the function of BDNF with peptides to selectively target TrkB receptors \(^{112}\). However, the specificity and effectiveness of BDNF mimetics remains to be seen.
Chapter 2. Methods

2.1 In vitro transcription and biotinylation of RNA

Plasmids containing human BDNF-AS (NR_033315.1) fragments downstream of the T7 promoter sequence and the SP6 promoter sequence immediately downstream of the 3’ end of the gene in the pMA vector were custom synthesized using GeneArt gene synthesis services (ThermoFisher Scientific, Waltham, MA, USA). Human BDNF-AS (NR_033315.1) was PCR amplified from the pMA-BDNF-AS vector with the following primers: T7 Forward 5’ TAATACGACTCACTATAG 3’ and SP6 Reverse 5’ ATTTAGGTGACACTATA 3’. Human HOTAIR was PCR-amplified from the pcDNA3.1-HOTAIR vector generously provided by Dr. Howard Chang using forward primer 5’ TAATACGACTCACTATAGGACTCGC 3’ and reverse primer 5’ TTGAAAATGCATCCAGATATTAATATATCTACA 3’. Fragments of human HOTAIR and E. coli MBP transcripts were produced using PCR primers as previously described. The pFB1.HMBP.EED vector containing the MBP gene was a generous gift from Dr. Thomas Cech. The pRL-TK vector containing the Renilla luciferase control transcript (Promega Corp, Madison, WI, cat # E2241) was linearized with XhoI, run on an agarose gel and the linearized fragment was excised and purified (Qiagen, Valencia, CA, gel extraction kit cat # 28704) prior to in vitro transcription. The BDNF-AS and HOTAIR PCR products were used as templates for T7 in vitro transcription following the manufacturers instructions (Applied Biosystems T7 MEGAScript kit cat # AM 1333, Carlsbad, CA). RNA yield was maximized by incubating overnight at 37°C in a water bath. The size and purity of RNAs was confirmed by applying test samples (50-100 ng) onto an Agilent RNA 6000 Nano chip (Agilent Technologies, Santa Clara, CA, cat #
Prior to applying the RNA onto the chip, RNAs were heated at 70°C for 2 minutes and cooled to room temperature to relax secondary structure. The chip was then run on an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). RNA was 3’ biotinylated using the Pierce RNA 3’ end biotinylation kit (Thermo Fisher Scientific, Waltham, MA, cat # 20160). Biotinylation reactions resulted in the fusion of a biotin-labeled CTP fused to the 3’ end of the RNA using T4 RNA ligase. Biotinylation reactions were extended overnight at 16°C and labeling efficiency was confirmed using the Pierce Chemiluminescent detection module (Thermo Fisher Scientific, Waltham, MA, cat # 89880). Briefly, a dot blot was performed to estimate the biotinylation efficiency of samples as compared to serial dilutions of a known biotin IRE control transcript. Equimolar control and sample transcripts were diluted, spotted onto a nylon membrane, crosslinked, and treated with streptavidin-horseradish peroxidase. The membrane was developed using the chemiluminescent module on the FluorChem E imager software version 4.1.1 (Bio-Techne, Minneapolis, MN). To estimate labeling efficiency, a non-saturated spot of the sample was compared to a known dilutions of the control transcript. Transcripts with >75% labeling efficiency were used for experiments.

### 2.2 RNA electrophoretic mobility shift assay (RNA EMSA)

RNA EMSA was performed using the LightShift® Chemiluminescent RNA EMSA kit (Thermo Fisher Scientific, Waltham, MA, cat # 20158) as per the manufacturers instructions using 3’ end biotinylated RNA (1 nM). Purified Human protein (EZH2 NM_004456) containing a C-terminal myc/FLAG tag was purchased from Origene technologies (Rockville, MD). RNA was incubated in a 37°C water bath for 30 min to allow the RNA to fold in EMSA binding buffer (10 mM HEPES pH 7.3, 20 mM KCl, 4
mM MgCl$_2$, 1 mM DTT) before EZH2 was added to the samples as indicated. The RNA was allowed to interact with EZH2 for 20 min at 37°C before the samples were run on a 0.5% agarose gel (Sigma-Aldrich, St. Louis, MO, cat # A0576) for 2 hrs at 4°C and 90V in 0.5 X TBE. The samples were then electrophoretically transferred onto a positively charged nylon membrane (Thermo Fisher Scientific, Waltham, MA, cat # 77016) in 0.5 X TBE for 30 min at 4°C and 400 mA. RNA was crosslinked to the nylon membrane for 5 min with a UV lamp equipped with 254 nm bulbs. The Chemiluminescent detection module (Thermo Fisher Scientific, Waltham, MA, cat # 89880) was used per manufacturers instructions to detect biotin-labeled RNA. Briefly, the nylon membrane was blocked with blocking buffer and then treated with a streptavidin-horseradish peroxidase conjugate that was detected by chemiluminescence on the FluorChem E imager software version 4.1.1 (Bio-Techne, Minneapolis, MN). Labeling efficiency was measured by comparing the signal intensity of newly labeled transcripts to known dilutions of a biotinylated iron response element (IRE) control RNA. Non-saturated spots of the test sample were compared to known dilutions of the control to determine labeling efficiency. Test RNAs with similar labeling efficiency (> 75%) were used for experiments to reduce variability.

2.3 LncRNA- EZH2 interaction AlphaScreen assay

AlphaScreen assays were performed using RNA described above and human protein (EZH2 NM_004456, LSD1 NM_001009999, PFKP NM_002627) containing a C-terminal myc/FLAG tag purchased from Origene technologies (Rockville, MD). Briefly, vectors containing the gene of interest were overexpressed in HEK293 cells. Cells were lysed and passed through an anti-FLAG column to capture the protein of interest
followed by conventional chromatography steps. Purified protein samples were run on an SDS-PAGE gel and stained with Coomassie blue to confirm protein size. Optimal concentrations of RNA, EZH2, acceptor and donor beads were determined by cross-titrating each reagent. RNA (1 nM to 1 pM as indicated) and EZH2 (4 nM) were tested in a 40 µl reaction (10 µl per each reaction component: RNA, protein, AlphaScreen Acceptor and AlphaScreen Donor beads (PerkinElmer, Waltham, MA, cat # 6760613) and were plated onto white 384-well OptiPlates (PerkinElmer, Waltham, MA, cat # 600790). Prior to performing any experiments, RNA was allowed to fold at 37°C for 15 min. RNA, protein, AlphaScreen beads and compounds were diluted in the following assay buffer: HEPES (30 mM) pH=7.4, NaCl (100 mM), MgCl₂ (2 mM), NP-40 (0.01%), Guanidinium HCl (10 mM), E.coli tRNA (50 µM, Sigma-Aldrich, St. Louis, MO, cat # R1753) and the plate was sealed and spun at 1000 rpm, then incubated at 37°C while shaking at 1000 rpm for 30 minutes. The assay buffer is a modification of the buffer used in a previous report of RNA-protein interactions measured by AlphaScreen. In order to find optimal buffer conditions, individual components of the buffer, including magnesium chloride, guanidinium HCl, E. coli tRNA, bovine serum albumin as well as range of buffer pHs, incubation times, and orders of reagent addition were tested to produce optimal assay signal. AlphaScreen anti-FLAG-coated Acceptor beads diluted in 1X PBS (20 µg/ml) were added to the plate and incubated for 60 minutes at room temperature. Subsequently, AlphaScreen streptavidin-coated donor beads (20 µg/ml) diluted in 1X PBS were added to the plate and incubated for 60 min at room temperature. The plate was re-sealed and spun at 1000 rpm briefly following each reagent addition. Bead additions were performed in the dark room under a green light due to the
photosensitivity of the beads. The assay plate was then read on an EnVision 2104 Multilabel Plate Reader (PerkinElmer, Waltham, MA) pre-programmed from the manufacturer with the AlphaScreen module (excitation of the donor beads at 680 nm and Alpha acceptor bead emission was measured at 570 nm).

2.4 High throughput screen (HTS) of phytochemical compound library

The Prestwick phytochemical compound library (Prestwick Chemical, Illkirch-Graffenstaden, France), was screened at 10 µM (0.1% DMSO) in duplicate. BDNF-AS or HOTAIR RNA (0.3 nM) and EZH2 (4 nM) were then added to the compound plate, the plate was sealed and spun at 1,000 rpm after final addition. Bead dilutions and additions were performed as described above. Data was plotted using GraphPad Prism software (San Diego, CA).

2.5 HEK293 cell culture and drug treatment

HEK293 cells were used to test the effect of Ellipticine (CAS 519-23-3) because these cells express EZH2, BDNF, BDNF-AS and contain the machinery required to inhibit BDNF. Cells were plated overnight in DMEM supplemented with 10% FBS in 6-well plates (500,000 cells per well) in a 37°C incubator (5% CO₂) and were treated the next day morning with DMSO (0.01%) or Ellipticine (1 µM) for 48 hours.

2.6 RNA extraction, cDNA synthesis and quantitative real-time PCR

Cells were washed with 1X PBS, lysed and RNA was extracted by passing cell lysates through RNeasy mini kit columns (Qiagen, Valencia, CA, cat # 74106) according to manufacturers instructions, including on-column DNase treatment. RNA (800 ng) was reverse transcribed using the qScript™ cDNA SuperMix kit (Quanta Biosciences,
Gene expression was measured by real-time PCR using human BDNF (Life Technologies, Grand Island, NY, cat # 4331182- assay ID Hs02718934_s1) and human GAPDH (Life Technologies, Grand Island, NY, cat # 4326317- RefSeq NM_002046.3) Taqman primer-probe assays and Taqman gene expression mastermix (Life Technologies, Grand Island, NY, cat # 4369016) in a 10 µl reaction. Real-time PCR data was measured using QuantStudio Flex 6 software from Life Technologies (Grand Island, NY) and the delta Ct method was used to determine relative gene expression. Data was analyzed using a two-tailed t-test on GraphPad Prism software (San Diego, CA).

2.7 Enzyme linked immunosorbent assay

The media from cells treated with compound (1 ml) was collected for ELISA experiments. The ELISA kit for human BDNF from Promega Corporation (Madison, WI, cat # G7611) was used following the manufacturers instructions. The average absorbance (measured at 450 nm) in the ellipticine-treated samples was subtracted from the background and normalized to DMSO-treated samples.

2.8 Software modeling of lncRNA structure

The BDNF-AS sequence was entered into Mfold and Sfold softwares openly available online to derive the most thermodynamically favorable secondary structure using default settings. catRAPID software was used to model the interaction propensity of BDNF-AS and EZH2 using default settings. The interaction propensity of EZH2 (746 amino acids) with the full length BDNF-AS transcript was modeled.
Chapter 3. Development of an assay to quantify lncRNA-protein interactions

3.1 Summary

We utilized AlphaScreen (a technology previously reported for the screening of small molecules targeting other RNA-protein interactions \(^{113}\)) to assay, quantify, and screen for small molecule modulators of lncRNA-protein interactions. In this study, the interaction of two lncRNAs (BDNF-AS and HOTAIR) with EZH2, were tested against a phytochemical library to look for small molecule inhibitors that can alter the expression of downstream target genes. Ellipticine, a low-micromolar inhibitor of BDNF-AS- EZH2 interactions, was identified as a compound that upregulates BDNF transcription. This work shows the feasibility of using HTS to identify modulators of a novel drug target, lncRNA-protein interactions, with small molecule therapies.

3.2 Results

3.2.1 BDNF-AS interacts directly with EZH2

Although our previous report infers that EZH2 and BDNF-AS interact at the chromatin level, the direct physical association has not been shown \(^{40}\). We confirmed this interaction with RNA EMSA. In this assay, biotinylated BDNF-AS was incubated with increasing concentrations of EZH2 that prevented the migration of biotinylated BDNF-AS through an agarose gel. The shift in the migration of biotinylated BDNF-AS in the presence of EZH2 was reversed with the addition of excess unbiotinylated BDNF-AS that displaced the biotinylated RNA bound to EZH2 \(^{92}\). These findings support the notion that
there is a direct physical interaction between BDNF-AS and EZH2 (Figure 3.1) that is also confirmed with AlphaScreen.

Figure 3.1 Mechanism of the BDNF-AS- EZH2 interaction. BDNF-AS interacts with and helps guide EZH2 to the BDNF locus, enabling EZH2 to deposit repressive methyl marks at the BDNF promoter.

3.2.2 LncRNA-protein interaction assay optimizations

AlphaScreen technology offers a rapid method of quantifying lncRNA-protein interactions using a nonradioactive amplified luminescent proximity homogeneous bead-based detection method. In the assay, the lncRNA and protein of interest are allowed to interact before AlphaScreen acceptor and donor beads are added to the reaction to measure the association of the binding partners. Upon excitation at 680 nm, the Alpha streptavidin-coated donor beads (that contain the photosensitizer phthalocyanine) convert molecular oxygen to an excited singlet oxygen (not a free radical) with a short (4 µsec) half-life. If the RNA and protein are bound, the singlet oxygen can diffuse up to 200 nm to make contact with a thioxene derivative on the Alpha donor bead, resulting in an amplified chemiluminescent emission between 520- 620 nm (Figure 3.2). In the
absence of a nearby acceptor bead (when the RNA and protein do not interact), the singlet oxygen falls to the ground state and does not produce a measurable signal. The signal amplification observed with AlphaScreen facilitates assay miniaturization for screening purposes.

**Figure 3.2 Schematic of the AlphaScreen adapted to quantify lncRNA-protein interactions.** Following the incubation of biotinylated BDNF-AS with FLAG-tagged EZH2, anti-FLAG-coated acceptor beads and streptavidin-coated donor beads are added to each well. Excitation of the donor beads at 680 nm results in the excitation of ambient oxygen that excites nearby acceptor beads (in the case of an RNA-protein interaction), resulting in a measurable emission at 570 nm that is used to quantify the assay.

Due to the sensitive nature of the AlphaScreen, the purity of RNA and protein samples were critical for experiments. RNA preparation was optimized (as described in methods) to produce a single RNA species. The secondary structure of *in vitro* transcribed RNAs was relaxed and samples were run on the bioanalyzer to confirm the size of the synthesized transcripts (**Figure 3.3 A**). We observed strong bands of the correct size for both BDNF-AS and HOTAIR RNAs. The bioanalyzer is run under non-denaturing conditions and as such, some secondary structures that were not denatured with heat persisted, accounting for the weaker bands seen in the HOTAIR gel. After confirming their size, the lncRNAs were biotinylated. A biotin label was fused to the 3’ end of the transcript to reduce interactions that could possibly result in changes in RNA
structure. The ligation of the biotin to RNA was optimized to ensure high efficiency (> 75%) labeling (as described in the Methods) with dot blot (Figure 3.3 B). Test RNAs with similar labeling efficiency were used for experiments to reduce variability. Purified human myc/FLAG-tagged EZH2 (size and purity confirmed as described in methods), was obtained from Origene technologies (Rockville, MD).

![Figure 3.3](image)

**Figure 3.3** Test RNAs are in vitro transcribed and biotin-labeled. (A) In vitro transcribed RNAs are heated at 70°C (to relax secondary structure) before being tested on the bioanalyzer under non-denaturing conditions to confirm the presence of RNAs of the correct size prior to biotin-labeling (nts- nucleotides). (B) Equimolar concentrations of 3’ end biotinylated RNA are spotted onto a nylon membrane, crosslinked and treated with streptavidin-horseradish peroxidase before being compared to a control biotinylated IRE RNA (75% labeling efficiency) to estimate the extent of biotin labeling in the test samples (as described in the Methods).

Next, we tested several reagent conditions and assay parameters to maximize assay signal. Individual components of the buffer, including magnesium chloride (required for RNA folding), guanidinium HCl (to prevent RNase activity), *E. coli* tRNA (to prevent non-specific binding of RNA to protein), bovine serum albumin (to prevent non-specific binding of protein to RNA), as well as range of buffers, buffer pHs, incubation times and temperatures for each section of the protocol, and orders of reagent addition were tested. To increase the specificity of the assay and block non-specific RNA interactions, we tested the effect of various tRNA concentrations on a single point
in the assay (Figure 3.4 A). Based on these data, we chose to use 50 µM tRNA in the assay buffer so as to prevent non-specific RNA binding without dampening the assay signal. We also tested the effect of increasing AlphaScreen bead incubation times on the assay signal and observed that increasing the bead incubation period to two hours improved the assay signal (Figure 3.4 B). RNA and protein concentrations for screening were determined by cross-titrating 1:2 serial dilutions of BDNF-AS into fixed concentrations of EZH2 as indicated using 20 µg/ml beads (Figure 3.5). Based on the binding isotherms generated, future assays were performed with 4 nM EZH2 to produce a robust assay signal. AlphaScreen beads were cross-titrated to determine appropriate acceptor and donor bead concentrations with the following reagent conditions: BDNF-AS (0.3 nM) and EZH2 (4 nM), (Figure 3.6). Ultimately, 20 µg/ml acceptor and 20 µg/ml donor beads were selected for experiments as these conditions produced a strong signal that did not change much with the addition of excess of either bead.

Figure 3.4 AlphaScreen assay conditions were optimized to maximize assay signal. (A) To reduce non-specific RNA interactions and increase the specificity of the AlphaScreen signal, *E. coli* tRNA was titrated into a single concentration of BDNF-AS and EZH2 (0.3 nM and 4 nM, respectively) and the AlphaScreen was performed as described in the Methods (n=3). (B) We tested the effect of reagent incubation times on the AlphaScreen signal. AlphaScreen was performed as described in the Methods (BDNF-AS concentrations indicated, EZH2 (4 nM)), except after the addition of AlphaScreen beads, incubation periods were 30, 60 or 120 minutes prior to being read (n=3).
3.2.3 AlphaScreen detects BDNF-AS- EZH2 binding

Several experiments were performed to confirm the specificity of the signal detected in the AlphaScreen assay. To show specific binding, the BDNF-AS- EZH2 signal was displaced with increasing concentrations of unbiotinylated BDNF-AS RNA (1-, 10-, and 100-fold excess). We observed that a 10-fold increase in unbiotinylated BDNF-AS was able to reduce the assay signal by 40% and 100-fold excess of unlabeled RNA completely quenched the signal.

We tested a non-human RNA to see if this transcript is able to specifically bind to EZH2 with the same potency as BDNF-AS. The Renilla luciferase pRL-TK vector was used to generate a negative control transcript to test for EZH2 binding. The Renilla RNA produced a low, concentration-dependent signal in the AlphaScreen (Figure 3.7 A), indicating some non-specific binding confirmed with RNA EMSA (Figure 3.7 B). This
finding corroborates previous reports that EZH2 is ‘promiscuous’, having the ability to bind many RNAs, even non-mammalian transcripts with low affinity\textsuperscript{99,100,122}.

**Figure 3.6** AlphaScreen bead concentrations were optimized for screening. AlphaScreen was performed (as described in the Methods) with BDNF-AS and EZH2 (0.3 nM and 4 nM, respectively). Anti-FLAG-coated acceptor beads and streptavidin-coated donor beads are added to each well at the indicated concentrations (n=3).

**Figure 3.7** RNA controls for the lncRNA-protein interaction AlphaScreen. A Renilla luciferase transcript was tested for its ability to interact with EZH2 as a control in the (A) AlphaScreen (n=3) and (B) RNA EMSA assays (as described in the Methods). AlphaScreen is highly sensitive and was able to detect concentration-dependent nonspecific binding that was confirmed in the RNA EMSA assay.

### 3.2.4 AlphaScreen detects HOTAIR- EZH2 binding

*HOTAIR* is an important biological target that has been implicated in several cancers\textsuperscript{54,61,123}. To study this lncRNA-protein interaction, we used the same assay conditions as
BDNF-AS to measure HOTAIR- EZH2 interactions. Increasing concentrations of HOTAIR RNA produced an increase in assay signal, confirming reports that EZH2 and HOTAIR are binding partners. Binding isotherms for the interaction of BDNF-AS or HOTAIR with EZH2 show an RNA concentration-dependent increase in assay signal when either transcript is incubated with EZH2 (Figure 3.8).

![Figure 3.8](image)

**Figure 3.8** Concentration-dependence of BDNF-AS and HOTAIR binding to EZH2. The lncRNA-protein interaction assay (as described in the methods) is performed with biotinylated lncRNAs BDNF-AS or HOTAIR titrated into EZH2 (4 nM) to produce a concentration-dependent AlphaScreen signal (n=3).

### 3.2.5 Optimization for HTS

In order to determine the suitability of our assay for screening, we calculated the day-to-day variability of several parameters: Z-factor, coefficient of variation and half maximal effective concentration (EC$_{50}$). We calculated Z-factor, a measure of assay suitability for screening from the binding isotherms generated in Figure 3.8. EZH2 (4 nM) and RNA (0.3 nM) consistently produced a signal with Z-factors greater than 0.5, indicating that the assay has little variation at this point. The coefficient of variation for this point was at or below 10%, which is suitable for HTS (Table 3.1). EC$_{50}$ values were
calculated using non-linear regression and consistent EC<sub>50</sub> values (0.33 nM ± 0.065) were observed (Table 3.1).

### Table 3.1 Assay parameters calculated to determine day-to-day variation in assay signal and HTS suitability.

<table>
<thead>
<tr>
<th></th>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
<th>Mean ± SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Z-factor</strong></td>
<td>0.71</td>
<td>0.63</td>
<td>0.68</td>
<td>0.67 ± 0.02</td>
</tr>
<tr>
<td><strong>Coefficient of variation (%)</strong></td>
<td>5.73</td>
<td>10.09</td>
<td>9.24</td>
<td>8.35 ± 1.33</td>
</tr>
<tr>
<td><strong>EC&lt;sub&gt;50&lt;/sub&gt; (nM)</strong></td>
<td>0.22</td>
<td>0.45</td>
<td>0.32</td>
<td>0.33 ± 0.07</td>
</tr>
</tbody>
</table>

Assay parameters calculated over three individual days of experiments to monitor assay performance and suitability for HTS. Screening optimizations included measuring Z-factor, coefficient of variation and EC<sub>50</sub> of the same experiment over several days. This consistency (Z-factor > 0.5 and coefficient of variation < 10%) suggests that the assay is suitable for HTS at EZH2 (4 nM) and BDNF-AS (3 nM).

As most compound libraries are dissolved in DMSO, the stability of this assay in the presence of increasing concentrations of DMSO was tested. It was observed that both BDNF-AS- EZH2 and HOTAIR- EZH2 assays were able to tolerate up to 1% of DMSO without altering the assay signal (Figure 3.9). Screening was performed at 0.1% DMSO, a concentration that did not affect either the BDNF-AS- EZH2 or HOTAIR- EZH2 interactions. Collectively, these data suggest that both IncRNA- EZH2 assays are suitable for small molecule screening.

### 3.2.6 HTS of phytochemical library

Phytochemical compounds have many pharmacophores, a great degree of stereochemistry and are natural metabolites, making them excellent candidates for the screening of potential bioactive drugs<sup>125</sup>. Because of these characteristics, a primary screen was performed using the Prestwick phytochemical library (Prestwick Chemical,
Figure 3.9 DMSO tolerance of lncRNA-protein interaction assay. DMSO was tested over a 100-fold range (0.1-10%) for both lncRNAs of interest. The assay components were stable in DMSO up to 1% and the AlphaScreen counts were reduced in a DMSO-concentration dependent manner beyond this point (n=3).

Illkirch-Graffenstaden, France). The 320-compound library was screened at 10 μM (0.1% DMSO) against BDNF-AS- EZH2 and HOTAIR- EZH2 and the scatterplots for percent inhibition of the library were generated (Figures 3.10 and 3.11). The data generated from this small-scale screen shows the potential of screening larger compound libraries in the lncRNA-protein interaction assay.

There are no known compound inhibitors of the RNA-protein interactions of interest, therefore, 100% inhibition represented the absence of RNA in the assay. This value indicates the background signal from interactions between donor and acceptor beads in the absence of a binding partner. The negative control consisted of a lncRNA and EZH2 in 0.1% DMSO in the absence of any compound (0% inhibition). The average Z-factor for the BDNF-AS- EZH2 screen was 0.65 and the hit rate (using a cutoff 70% inhibition as a hit) was 3.1%. For HOTAIR- EZH2, the average Z-factor was 0.85 and the hit rate was 1.25%. It was also observed that some compounds enhance the interaction between lncRNAs and EZH2, which are shown as negative values on the graphs. These
compounds seem to strengthen the lncRNA-protein interaction and although not explored by our group, could have therapeutic value in other disease contexts. As expected, one of

![Figure 3.10](image)

**Figure 3.10** Scatterplot of values from the phytochemical compounds screened for the inhibition of BDNF-AS - EZH2. BDNF-AS (0.3 nM) and EZH2 (4 nM) were incubated with 10 μM (0.1% DMSO) test compound and AlphaScreen was performed as described in the Methods. The average Z-factor was 0.65 and the hit rate (cutoff of 70% inhibition as a hit) was 3.1%. Data is the average of duplicates expressed as percent inhibition. The point circled in red indicates biotin, a compound that validates our screen.

the natural compounds present in the compound library, biotin, produced a hit in both screens (**Figures 3.10** and 3.11). Free biotin will saturate the streptavidin-coated donor beads and since free biotin does not interact with EZH2, will inhibit the assay signal. This compound serves as a positive control and as an assay validation. These data suggest that
Figure 3.11 Scatterplot of values from the phytochemical compounds screened for the inhibition of HOTAIR-EZH2. HOTAIR (0.3 nM) and EZH2 (4 nM) were incubated with 10 µM (0.1% DMSO) test compound and AlphaScreen was performed as described in the Methods. The average Z-factor was 0.85 and the hit rate was 1.25%. Data is the average of duplicates expressed as percent inhibition. The point circled in red indicates biotin, a compound that validates our screen.

Possible modulators of lncRNA- EZH2 interactions exist and that they can be identified with our assay.
3.2.7 Secondary assays: BDNF-AS- EZH2

From the initial screen, 10 putative compound hits for BDNF-AS- EZH2 (Table 3.2 and Figure 3.12) were identified. Interestingly, only one compound (biotin) was identified as a common inhibitor for both transcripts. BDNF-AS- EZH2 compound hits were assayed to establish potency against BDNF-AS- EZH2 and validated hits for BDNF-AS- EZH2 inhibition (myricetin, gossypol, ellipticine and biotin, Figure 3.13) were tested in secondary cell culture experiments. HEK293 cells were selected as a model to study inhibitors of BDNF-AS- EZH2 as these cells express EZH2, BDNF-AS and the repressed target of this interaction, BDNF. This cell model was also used in initial studies to study the mechanism of BDNF-AS. All validated compounds were tested in a concentration-dependent manner (0.1, 0.3, 1 and 3 µM) to determine toxicity after 48-hour drug treatment. We determined concentrations for treatment with each drug by

Table 3.2 List of initial hits from the primary screen for BDNF-AS- EZH2.

<table>
<thead>
<tr>
<th>BDNF-AS-EZH2 Compound Hits</th>
<th>% Inhibition</th>
<th>Molecular weight (g/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solanine alpha</td>
<td>102.9</td>
<td>868.06</td>
</tr>
<tr>
<td>Syrosingopine</td>
<td>72.4</td>
<td>666.7</td>
</tr>
<tr>
<td>Myricetin</td>
<td>69.4</td>
<td>318.24</td>
</tr>
<tr>
<td>Roseoflavin</td>
<td>76.37</td>
<td>404.42</td>
</tr>
<tr>
<td>Curcumin</td>
<td>97.2</td>
<td>368.38</td>
</tr>
<tr>
<td>9-Methoxyellipticine</td>
<td>73.89</td>
<td>276.34</td>
</tr>
<tr>
<td>Gossypol</td>
<td>94.26</td>
<td>518.56</td>
</tr>
<tr>
<td>Biotin</td>
<td>129.43</td>
<td>244.31</td>
</tr>
<tr>
<td>Ellipticine</td>
<td>116.7</td>
<td>246.34</td>
</tr>
<tr>
<td>Ergotamine tartrate</td>
<td>95.96</td>
<td>656.7</td>
</tr>
</tbody>
</table>

Compounds with >70% inhibition were considered hits and were selected for follow-up studies. Compounds in blue were not further tested due to limited availability.
visually surveying cells for toxicity and monitoring their overall health during treatment. After determining non-toxic treatment concentrations, cells were treated for 48 hours and then changes in target gene expression were measured. One compound, ellipticine, (1µM) was able to upregulate BDNF transcription (~ 3-fold, p <0.0001) after 48 hours, normalized to GAPDH (Figure 3.14 A). The effect of ellipticine treatment on BDNF levels after 24 and 72 hours was also measured, however no significant changes in BDNF expression were observed. To test whether increased BDNF transcription also resulted in increased BDNF protein, an ELISA assay was performed using the media from ellipticine-treated HEK293 cells. BDNF is a secreted protein and therefore was expected to be present in the media on compound-treated cells.

![Chemical structures of the compounds](image1)

**Figure 3.12** Chemical structure of BDNF-AS- EZH2 primary screen hits. Hits from the BDNF-AS- EZH2 screen were selected for validation.

A small (10%), but statistically significant (p <0.001) upregulation in BDNF protein was observed (Figure 3.14 B). Although ellipticine was not able to upregulate BDNF
appreciably at the protein level, an increase in BDNF transcription following drug treatment was observed, indicating that ellipticine treatment does indeed have an effect on the target of BDNF-AS- EZH2.

**Figure 3.13** Validated hits from the primary screen of BDNF-AS- EZH2. The compounds listed in Table 3.2 underwent concentration-response testing (as described in the legend to Figure 3.10, concentrations as indicated). Concentration-response curves for the validated primary BDNF-AS- EZH2 screen hits are shown (n=3).

**Figure 3.14** Secondary assays to test the effect of ellipticine treatment on BDNF transcription and translation. (A) HEK293 cells were treated for 48 hours with ellipticine (1 µM) before cells were lysed, RNA was extracted and gene expression was measured (**** p<0.0001, n=3). (B) HEK293 cells were treated for 48 hours with ellipticine (1 µM) before the media on the cells was used to measure changes in secreted BDNF protein with ELISA (***p <0.001, n=3).
3.2.8 Secondary assays: *HOTAIR*- EZH2

All three hits for *HOTAIR*- EZH2 (Table 3.3 and Figure 3.15) were tested and validated. For secondary cell culture experiments, HeLa cells were chosen as a cell model as they express EZH2, *HOTAIR* and repressed *HOTAIR* target gene *HOXD11*. The validated *HOTAIR*- EZH2 compounds were tested in a concentration-dependent manner (0.1, 0.3, 1 and 3 μM) to determine toxicity in cells following 48-hour drug treatment. Upon visual inspection to survey overall cell health, concentrations to be tested in HeLa cells were selected and changes in *HOXD11* expression were measured. These experiments validated one hit, camptothecin, for *HOTAIR*- EZH2. However, camptothecin also altered the expression of two housekeeping genes: *beta actin* and *GAPDH* and was not further studied.

**Table 3.3** List of initial hits from the primary screen for *HOTAIR*- EZH2.

<table>
<thead>
<tr>
<th><strong>HOTAIR-EZH2 Compound Hits</strong></th>
<th><strong>% Inhibition</strong></th>
<th><strong>Molecular weight (g/mol)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Coraline chloride hydrate</td>
<td>84.01</td>
<td>417.89</td>
</tr>
<tr>
<td>Ellagic Acid</td>
<td>61.9</td>
<td>302.2</td>
</tr>
<tr>
<td>Camptothecin (S, +)</td>
<td>74.35</td>
<td>348.4</td>
</tr>
<tr>
<td>Biotin</td>
<td>107.02</td>
<td>244.31</td>
</tr>
</tbody>
</table>

Compounds with >70% inhibition were considered hits and were selected for follow-up studies. All of these compounds underwent concentration-response validation testing.

**Figure 3.15** Chemical structure of *HOTAIR*- EZH2 primary screen hits. Hits from the *HOTAIR*- EZH2 screen were selected for validation.
3.3 Discussion

In the past decade, next generation sequencing platforms and multicenter transcriptomics efforts have helped to increase the inventory of functional IncRNAs at an incredible rate.

![Graph](image)

**Figure 3.16** *Validated hits from the primary screen of HOTAIR- EZH2.* The compounds listed in Table 3.3 underwent concentration-response testing (as described in the legend to Figure 3.11, concentrations as indicated). Concentration-response curves for the primary HOTAIR- EZH2 screen hits are shown (n=3).

However, increasing knowledge of the disease-relevance of IncRNAs and the observation that they can exert their functions by acting through epigenetic enzymes, necessitates thorough studies on the structural components of these binding partners\(^{101,126,127}\). Furthermore, it prompts us to view these interactions as novel therapeutic targets. Understanding the nature of RNA-ligand interactions is important and has been highlighted\(^ {113}\), however work in this arena has been slow, partially due to difficulties in targeting RNAs as a single RNA sequence can give rise to multiple metastable structures\(^ {128}\). Several low throughput assay methods exist to measure these interactions,
however this is the first report of an assay being used to study lncRNA-protein interactions for small molecule screening purposes.

Our AlphaScreen assay performed well with Z-factors greater than 0.5 and acceptable assay variation (<10%), indicating that the assay can be scaled to screen large compound libraries. Although AlphaScreen is a convenient and commonly used screening technology, one well-known limitation of this assay is the ‘hook effect’. At the ‘hook’ or peak assay signal, either donor or acceptor bead is saturated with RNA or protein. Increasing RNA or protein concentration above the ‘hook’ point results in a decrease in assay signal, producing a bell-shaped curve as opposed to a classical saturation curve \(^{117}\). Our experiments showed that reducing concentrations of the protein was not able to overcome the observed hook while maintaining a consistent assay signal. Although the assay can be used successfully for screening purposes, due to the limitations incurred by the persistence of the hook effect, it is not possible to approximate a binding constant for the lncRNA- EZH2 interactions (Figure 3.5). Despite this limitation, we were able to perform a primary screen and confirm several hits in secondary assays. In this scenario, low throughput methods, such as RNA EMSA could be used to determine binding constants for lncRNA-protein interactions, while this assay has the distinct advantage of being used for large-scale screening.

A small natural compounds library was tested to show the feasibility of compound screening using the lncRNA-protein AlphaScreen interaction assay. Natural compounds have many pharmacophores, a great degree of stereochemistry and are natural metabolites, facilitating transport to their intended sites of action and making them excellent candidates for screening of potential bioactive drugs \(^{125}\). It was observed that
one compound, ellipticine, was able to upregulate BDNF mRNA, and to a lesser degree BDNF protein. Isolated from Apocyanaceae plants, ellipticine is a naturally occurring alkaloid that functions as a DNA topoisomerase II inhibitor and has derivatives with promising anti-tumor and anti-HIV activities \(^{129-132}\). Although ellipticine and its derivatives possess desirable anti-cancer effects, particularly in aggressive breast cancers, hydrophobicity and toxicity remain problematic. Alternative delivery methods to reduce \textit{in vivo} toxicity have been attempted with some success, suggesting that future derivatives of this compound have therapeutic potential \(^{133-135}\).

The main screening hit for \textit{HOTAIR}- EZH2 was camptothecin, a quinoline alkaloid isolated from the \textit{Camptotheca acuminata} tree whose derivatives possess many medical benefits. Camptothecin is a DNA topoisomerase I inhibitor whose most well-known derivative to date is topotecan, a potent chemotherapeutic agent with FDA approval for use in ovarian, cervical, and most recently small cell lung cancer. Camptothecin itself is mostly water insoluble and has many side effects, which is why current research focuses on derivatives of this quinolone for therapeutic uses. Importantly, this compound is still widely researched and is still being pursued by many groups for the development of cancer therapies \(^{136,137}\).

The identified hits from the screens are compounds whose scaffolds exhibit therapeutic potential that require further development to produce potential small molecule leads. This work, however, serves as a proof of concept that lncRNA-protein interactions can be assayed and pharmacologically targeted. Small molecule therapies that could upregulate BDNF mRNA, if able to cross the blood-brain barrier, could possibly be used to treat a number of neurological diseases where BDNF is downregulated \(^{93,138}\).
Furthermore, *HOTAIR*-EZH2 was screened. *HOTAIR* is significantly upregulated in breast 54, colorectal 123, and pancreatic 61 cancers and plays a functional role in the progression of these cancers through its interaction with Polycomb repressive complex 2. Therefore, compounds that arise from future screening of this target could be potential cancer therapeutics.

Our method, compared to previously established methods of measuring lncRNA-protein interactions is more rapid, requires less steps and is amenable to large-scale screening of many potential lncRNA-protein targets with diverse small molecule libraries. HTS can lead to two possible types of hits: (1) compounds that directly block the RNA-binding pocket of EZH2 and (2) compounds that bind to RNA to change the secondary or tertiary structure in a way that prevents its binding to the intended protein target. In the second scenario, identified hits are highly lncRNA-specific and might act selectively to block the function of EZH2 targets. Nonspecific blocking of EZH2, as previously attempted in cancer, might have mixed and opposing effects on several oncogenes and oncosuppressors.

### 3.4 Future directions

Introducing a new level of specificity by targeting lncRNA-protein interactions might help identify potent and specific therapeutics for a number of disorders. Although the study of highly specific interactions could be challenging, targeting individual dysregulated lncRNAs and their protein partners could pave the road for precision medicine. As such, further optimizations and modified compound libraries are required to fully explore lncRNA-protein interactions and small molecules that could modulate these drug targets.
Chapter 4. RNA requirement for EZH2 interactions

4.1 Summary

Our lncRNA-protein interaction AlphaScreen assay was employed to gain insights into the RNA requirement for binding to EZH2. The ability of the assay to detect these interactions was further confirmed for accuracy by testing experimentally validated segments of HOTAIR. Next, structure modeling software and experimental data were employed to fragment BDNF-AS and determine which segments of this transcript enable it to interact with EZH2. Although the affinity for EZH2 is variable among different fragments of BDNF-AS, EZH2 is able to bind multiple RNA fragments. These data suggest that several structural elements in BDNF-AS enable it to interact with EZH2, highlighting the importance of the antisense transcript in fine-tuning the activity of this indispensable neurotrophin.

4.2 Results

4.2.1 AlphaScreen measures HOTAIR-protein interactions

HOTAIR is a well-studied lncRNA whose structure was recently elucidated. Prior to the publication of the secondary structure of HOTAIR, several groups used various biochemical approaches to understand which segments of this transcript are required to interact with epigenetic enzymes. In one report, immunoprecipitation of different size fragments of in vitro transcribed HOTAIR with EZH2 by randomly truncating the 3’ end of the transcript showed that the first 300 nucleotides of HOTAIR are indispensable for its interaction with EZH2. In the same report, the 3’ end of
HOTAIR was shown to be a requirement for its interaction with another epigenetic enzyme, lysine-specific demethylase 1A (LSD1) \(^{37}\). Other groups studied the HOTAIR-EZH2 interaction using RNA EMSA and showed that the first 300 nucleotides of HOTAIR is sufficient to bind EZH2 \(^{98-100,140}\).

The use of AlphaScreen as a methodology to measure RNA-protein interactions for small molecule modulators has been reported \(^{92,113}\). Here, we applied this tool to rapidly determine which segments of RNA are required for lncRNA-protein interactions. Our assay overcomes several disadvantages observed with traditional RNA-protein interaction assays because it requires significantly less starting material, provides an accurate readout within hours, is highly sensitive, and can be used to perform large-scale experiments. In addition to previous work on the validity of this assay for the detection of lncRNA-protein interactions \(^{92}\), we were able to replicate previous findings using fragmented HOTAIR transcripts (Figure 4.1 A). Reports show that HOTAIR binds to EZH2 with high affinity (K\(_D\) ~ 28 nM), even though EZH2 interacts with many RNAs, even nonspecific RNAs, for example, MBP, with low affinity (K\(_D\) ~ 380 nM) \(^{99,140}\). We sought to replicate these findings with fragments of HOTAIR in the AlphaScreen assay with EZH2 (Figure 4.1 B) and LSD1 (Figure 4.1 C), two proteins reported to bind different segments of HOTAIR. EZH2 is able bind HOTAIR full length, 1-300, and 300-2146 with roughly the same affinity as shown by the binding isotherms in Figure 4.1 B. HOTAIR 1-300 antisense exhibits low affinity for EZH2. These findings corroborate previous reports \(^{37,99,140}\) and especially the most recent study describing HOTAIR structure, indicating that domain 1 (nucleotides 1-530) has several important structural elements which increase its affinity for EZH2 two-fold over fragment 1-300 alone \(^{139}\).
Figure 4.1 The lncRNA-protein interaction AlphaScreen assay was validated with HOTAIR. (A) Schematic representing the various HOTAIR fragments synthesized for experiments. AlphaScreen (as described in the Methods) was performed to produce binding isotherms for the interaction of (B) EZH2 (4 nM), (C) LSD1 (4 nM), and (D) PFKP (4 nM) with increasing concentrations of indicated RNAs (n=3).

Fragments 1-300 and 300-2146 (which contain 230 nucleotides required to bind HOTAIR) bind to EZH2 with similar affinity, and AlphaScreen was sensitive enough to detect this, whereas EMSA was not. As an RNA control, the affinity of a non-mammalian transcript, maltose binding periplasmic protein (MBP) for EZH2 was tested and some non-specific binding, similar to a previous report, was observed. We confirmed that two MBP fragments (1-300 and 1-800) bind EZH2 non-specifically in our assay (Figure 4.1 B), similar to another non-mammalian transcript (Renilla luciferase, Figure 3.7). The ability of various HOTAIR fragments to interact with LSD1 was tested and it was
observed that none of the fragments containing the 5’ end of HOTAIR had a strong affinity for this enzyme as a previous report suggests (Figure 4.1 C)\textsuperscript{37}. Protein phosphofructokinase-platelet 1 (PFKP), a kinase involved in glycolysis, served as a negative protein control. Binding isotherms for all transcripts tested in Figures 4.1 B and C indicate that these transcripts do not have an affinity for PFKP (Figure 4.1 D). Collectively, these observations strongly suggest that EZH2 interacts with various fragments of HOTAIR with high affinity and that AlphaScreen can be utilized to detect and measure these interactions. Therefore, we used this newly developed tool to better understand the structure of BDNF-AS.

\textbf{4.2.2 BDNF-AS RNA requirement for EZH2 binding}

BDNF-AS is a 1.5 kb lncRNA whose structure is currently unknown. We hypothesized that certain sequence motifs or secondary structure elements of the transcript enable it to interact with EZH2. Previous data suggest that the sense-antisense overlap region might be critical for its interaction with EZH2\textsuperscript{40}. This overlap region is present in all splice variants of BDNF-AS in both human and mouse. Moreover, blocking the first 113 nucleotides of the 225-nucleotide overlap using tiling antisense oligonucleotides (AntagoNATs), resulted in an upregulation in BDNF mRNA. However, there was no significant increase in BDNF mRNA when inhibiting the second half of the overlap region of BDNF-AS with antagoNATs\textsuperscript{40}. Therefore, we hypothesize that the BDNF sense-antisense overlap region might be important for mediating its interaction with EZH2. We used lncRNA-protein interaction modeling software to predict segments that bind to EZH2 and tested these fragments in the AlphaScreen assay.
To fragment *BDNF*-AS, software tools were employed to produce potential secondary structures for *BDNF*-AS. The goal of structure modeling was to identify signature stem loop structures similar to those identified for other EZH2-interacting lncRNAs X-inactive specific transcript (*Xist*) and *HOTAIR*\(^{48,98}\). Mfold\(^{114}\) and Sfold\(^{115}\) were used to model *BDNF*-AS to derive the most thermodynamically favorable RNA structure predictions (Figure 4.2 A). RNA stem loops were previously reported to be required for interacting with EZH2\(^{48,98}\), however it was not possible to discern those same structures in the *BDNF*-AS transcript. This was not unexpected as software tools are limited to energy folding calculations that often produce structures not observed in nature and do not take into account pseudoknots, which can result in important changes in RNA secondary structure.

To supplement these computational methods, an open-source software, catRAPID, was used to produce a table of results with specific amino acid and RNA regions where a lncRNA and protein can potentially interact based on secondary structure prediction, van der Waals interactions and hydrogen bonding\(^{116}\). The catRAPID predictions were used as a guide to study the structure of *BDNF*-AS (Figure 4.2 B). Notably, catRAPID accurately predicted 6 of 20 amino acid stretches that could interact with lncRNAs. These predicted protein sites are consistent with the putative noncoding RNA binding domain previously identified for amino acids 342-370 of EZH2\(^{96}\).

We used software models and experimental data to fragment *BDNF*-AS into segments of various lengths to determine whether sequence, secondary structure, or a mix of both, are responsible for the interaction of *BDNF*-AS with EZH2 (Figure 4.3 A and
B). A more systematic approach was also employed and *BDNF*-AS was fragmented to synthesize a series of successive 3’ end truncated fragments (Figure 4.3 B). We observed

**Figure 4.2** Modeling the structure of *BDNF*-AS. (A) Sfold was used to model the secondary structure of *BDNF*-AS by determining the minimal free energy structure. (B) CatRapid was used to predict the RNA and protein regions on *BDNF*-AS and EZH2 that interact. The blue box highlights the predicted protein regions that correlate with the noncoding RNA binding domain of EZH2 (amino acids 342-370).

that full length *BDNF*-AS binds to EZH2 with the same affinity as *BDNF*-AS missing the sense-antisense overlap region (*BDNF*-AS overlap deletion) (Figure 4.3 C). Furthermore, the 225-nucleotide *BDNF*-AS sense-antisense overlap segment (*BDNF*-AS nucleotides 306-526) is sufficient to bind EZH2, although with slightly less affinity than the full length or overlap deletion fragments (Figure 4.3 C). However, when truncating the overlap fragment into two equal parts (overlap A and B), the ability of either fragment to bind to EZH2 was inhibited (Figure 4.3 C).
Figure 4.3  BDNF-AS fragments were tested for their ability to bind EZH2. (A) Schematic of the BDNF-AS fragments based on the conserved sense-antisense region of the transcript (nucleotides 306-526 of BDNF-AS). (B) Schematic of the 3’ end deletion fragments of BDNF-AS. (C) Fragments based around the overlap region in (A) were tested in the AlphaScreen assay (n=3). (D) Truncation fragments in (B) were tested in the AlphaScreen assay. LncRNA-protein interaction assays were performed as described in the Methods with the indicated RNAs (n=3).

Comparing the 3’ end truncated fragments of BDNF-AS, fragments 1-1000, 1-500 and 1-257 all bind EZH2 with similar affinity (Figure 4.3 D). Interestingly, fragments 1-1000 and 1-500 contain the sense-antisense overlap region that is not present in fragment 1-257. Lastly, two fragments were synthesized to test the catRAPID-predicted results. BDNF-AS 580-700 was predicted to bind to EZH2, therefore, we synthesized this short fragment as well as a longer transcript, BDNF-AS 300-700, a region just downstream of the BDNF sense-antisense overlap region (nucleotides 306-526), to verify that nearby
structural or sequence elements were not disrupted due to fragmentation. However, the catRAPID-predicted RNA regions ($BDNF$-AS 580-700 and 300-700) do not show affinity for EZH2 (Figure 4.3 D). These findings indicate that EZH2 has a high affinity for binding RNA fragments with varying lengths and sequences.

4.3 Discussion

We utilize the lncRNA-protein interaction AlphaScreen assay in conjunction with software modeling and experimental evidence to glean insights into the sequence or structural elements of $BDNF$-AS that enable it to interact with EZH2. The evidence indicates that several regions of $BDNF$-AS exhibit strong affinity for EZH2, regardless of nucleotide sequence. We observe that the experimentally-derived sense-antisense overlap region alone is able to bind EZH2 and that 3’ end truncated forms that contain the overlap bind EZH2 with similar affinity to that of transcripts that do not contain the overlapping region. This evidence supports the notion that secondary structure motifs in $BDNF$-AS are responsible for EZH2 interactions. These findings are consistent with reports that RNA secondary structure is predominantly responsible for the ability of many diverse lncRNAs to bind EZH2 $^{141,142}$. For example, mutating the sequence but maintaining the structure of stem-loops in the A-repeat region of $Xist$ did not affect the ability of the transcript to bind to EZH2 in EMSA assays $^{141,142}$. An RNA immunoprecipitation following by sequencing (RIP-seq) analysis of polycomb-associated RNAs in mouse embryonic stem cells showed that ~9,800 diverse RNA transcripts can associate with PRC2, most likely due to structure rather than sequence similarity $^{46}$. In a more recent report, photoactivatable-ribonucleoside- Enhanced crosslinking and immunoprecipitation (PAR-CLIP) $^{143}$, was used to identify RNA- EZH2 interactions $^{122}$. This report concluded
that EZH2 binds nascent RNAs promiscuously \textit{in vivo} without the need for specific structure or sequence elements, but did not rule out that specific lncRNAs could have evolved to exploit this property \textsuperscript{122}. Finally, a recent review of interspecies conservation amongst different lncRNAs from ENCODE-compiled data determined that in the case of lncRNAs, structure is the main functional unit and evolutionary constraint, highlighting the need to further understand the structural motifs of functional lncRNAs \textsuperscript{144}.

It is plausible that \textit{BDNF}-AS has several regions that are responsible for its interaction with EZH2, such that the ablation of one does not significantly impact the binding ability of the rest of the transcript. This might, in part, be due to RNA folding. Furthermore, it suggests that this lncRNA has evolved such that fragmentation or ablation of specific segments does not inhibit its ability to bind EZH2. Considering the importance of regulating BDNF in development and disease, it is possible that it is evolutionary required that this transcript have several regions that bind EZH2 to conserve function and fine-tune expression.

\textbf{4.4 Future directions}

The lncRNA-protein interaction assay is able to accurately predict interacting partners, but in the case of \textit{BDNF}-AS, requires structural insights to make further conclusions about this transcript. Structure information would make it possible to synthesize a minimal EZH2-binding region of \textit{BDNF}-AS to rapidly screen for small molecule inhibitors. Recent insights into \textit{HOTAIR} structure will further drug discovery efforts and will act as a primer for targeting other lncRNAs with small molecules \textsuperscript{139}. 
Chapter 5. Discussion

5.1 Designing IncRNA-directed therapeutics

Despite the multiple ways that genes can be silenced, methods for upregulating the expression of therapeutic genes remains a challenge \(^{145}\). IncRNAs represent attractive therapeutic targets as they regulate the expression of a limited set of genes (a single gene in the case of cis-acting natural antisense transcripts), whereas proteins typically contribute to multiple cellular processes and molecular pathways, making it difficult to predict all their targets and avoid side effects. Although IncRNAs lack sequence conservation, they are structurally conserved, making them stable and predictable physical targets \(^{144,146}\). Recent advances in RNA structure determination for the repeat RNA that causes spinocerebellar ataxia type 10 and HOTAIR will allow for rational design of small molecules targeting these transcripts \(^{139,147}\). Efforts are already underway to target microRNA transcripts \(^{148,149}\) as well as RNAs arising from repeat expansions in Fragile X syndrome\(^{150}\), frontotemporal dementia and amyotrophic lateral sclerosis \(^{151}\). Although the ‘druggability’ of RNAs with small molecules is actively studied, there is still much work to be done in this exciting field \(^{152,153}\). IncRNAs represent appealing pharmacological and therapeutic targets, however inhibiting these transcripts in vivo remains a challenge \(^{154}\).

IncRNAs have mainly been targeted using antisense oligonucleotides (ASOs), siRNAs and viral vectors that contain shRNAs. Knockdown of disease genes, in vitro, usually involves the use of siRNAs that are unstable and non-penetrating to target cells in vivo. Modified antisense oligonucleotides are stable, nuclease-resistant, short (13-25 nucleotides), single-stranded DNA oligonucleotides that are complementary to RNAs of
interest, that have been used to induce RNase H-dependent cleavage of lncRNA targets. Various modified nucleotides have been tested both in vitro and in vivo and these modifications are continually undergoing enhancements in chemistry. Modified oligonucleotides targeting antisense lncRNAs (AntagoNATs), usually contain a mixture of 2’-O-methyl RNA and Locked Nucleic Acid (LNA) modifications at their 5’ and 3’ ends to protect against nuclease cleavage and phosphothioate backbone to enhance their cellular uptake.\(^{155}\)

Knockdown of lncRNAs in vivo has led to successful upregulation of therapeutic target genes and proteins resulting in measurable phenotypic responses.\(^{156}\) Continuous in vivo infusion of antagoNATs targeting mouse Bdnf-AS produced a locus-specific upregulation in Bdnf mRNA and protein. Mice treated with Bdnf antagoNATs display a marked increase in proliferating cells and neuronal survival in the brain as compared to controls, indicating that an increase in functional Bdnf was achieved. This effect is made all the more significant because of the difficulty in upregulating neurotrophins in vivo, even through small molecule modulators of neurotrophin receptors.\(^{157}\) Repeated administration of AntagoNATs targeting monkey ApoA1-AS resulted in the upregulation of HDL levels. Nuclear-localized lncRNA Malat1 was successfully targeted for downregulation in skeletal muscle with the systemic administration of a gapmer, that might have therapeutic applications in cancer.\(^{158}\) Subcutaneous administration of antisense oligonucleotides targeting MALAT1 also effectively inhibited human lung cancer cell proliferation in a mouse xenograft model.\(^{159}\) In the past few years, around 100 ASOs and 40 RNAi-based therapies have entered clinical trials, including over 20 in advanced clinical trials (Phase II or III).\(^{160}\) Two oligonucleotide drugs, Fomivirsen and
Mipomersen\textsuperscript{162}, have received FDA-approval to treat Cytomegalovirus retinitis and high blood cholesterol, respectively. Additionally, several ASO-based therapeutics are in stage III clinical trials, including ASOs that induce exon skipping in Duchenne muscular dystrophy\textsuperscript{163}. Although the use of ASOs for gene upregulation seems promising, poor intracellular uptake and chemistry-dependent toxicity are still major concerns when developing therapeutics\textsuperscript{164}.

RNA interference approaches have been successful in downregulating the expression of lncRNAs \textit{in vitro} to promote upregulation of their target genes\textsuperscript{165}. \textit{In vivo}, RNAi-based approaches require a delivery vehicle such as liposomes, nanoparticles or viruses to protect siRNA or shRNA vectors from nuclease degradation, prevent their accumulation in the liver and enhance cellular uptake\textsuperscript{165}. There have been multiple successful attempts to target an mRNA with RNAi \textit{in vivo}\textsuperscript{166}, and currently, 40 RNAi therapies are in clinical trials\textsuperscript{166}. Development of new RNAi delivery vehicles for the inhibition of lncRNAs and locus-specific gene upregulation holds promise. Viral vectors are effective in delivering shRNAs \textit{ex vivo} and \textit{in vivo} to produce stable and specific knockdown of target RNAs and have been used multiple times \textit{in vitro} and \textit{in vivo} to knockdown lncRNAs to upregulate target genes. Lentiviral vectors for gene therapy have successfully been used for transducing hematopoietic stem cells in the course of clinical trials for treating two rare genetic disorders: metachromatic leukodystrophy (MLD)\textsuperscript{167} and Wiskott–Aldrich syndrome (WAS)\textsuperscript{168}. Immune cells genetically modified with lentiviral vectors have been used in multiple paradigms of cancer immunotherapy\textsuperscript{169}. Moreover, \textit{ex vivo} gene therapy with lentiviral transduction of shRNAs to stably downregulate target disease mRNA has been used in transplantation of CD34(+) cells to treat HIV\textsuperscript{170}. Overall, antisense
oligonucleotides and RNAi therapies targeting repressive IncRNAs are promising but require additional developments in oligonucleotide chemistry and use of vehicles for safe and effective delivery.

### 5.2 Small molecule modulators of IncRNA-protein interactions

Chromatin-modifying enzymes are ubiquitously expressed proteins that lack rigidity in protein structure, making them difficult therapeutic targets using a small molecule approach \(^{171}\). However, these chromatin-modifying enzymes usually require IncRNAs to form complexes with other effector proteins to be targeted to specific genomic loci. Targeting IncRNA-protein interactions would increase the specificity of compounds, reduce off-target effects, and provide reversible inhibition of chromatin modifying enzymes at non-catalytic domains of the protein \(^{172}\).

Because of the malleability of proteins and RNAs, the most efficient site to target IncRNAs with small molecules should ideally be at the binding cleft where the IncRNA and chromatin-modifying enzyme interact. To achieve this, modeling the conformation of proteins and RNAs to understand the unique conformations these molecules undertake when they interact is key. Several studies have focused on high throughput approaches to understand RNA structure and RNA-protein interactions on a large-scale \(^{173-179}\). However, X-ray crystallography and NMR spectroscopy should be incorporated to produce models to better understand how small molecules would both fit and bind in the IncRNA-protein binding cleft. This approach has been successfully used in modeling a single protein or protein-protein interactions, but should also be extended large-scale to IncRNA-protein interactions. It should be noted that these approaches have important limitations, e.g. the length of time and difficulty in producing crystal structures and the
limitations in nucleic acid length and protein size in NMR, making it difficult to produce a structure for most full-length lncRNAs. Biochemical studies are necessary to accurately model the minimum regions required to allow lncRNAs to interact with specific domains of chromatin modifying enzymes. Quantitative assay methods should be employed in conjunction with biochemical approaches and structural modeling so that compound screening can be performed and hit compounds can be optimized using accurate models. The use of small molecules would overcome many of the challenges observed in current ncRNA-directed therapies.

LncRNA-protein interactions represent an important class of molecular effectors within the cell as well as a new and unexplored set of highly specific druggable targets.

5.3 Conclusions

Although a wealth of transcriptomics data exists indicating the potential presence of thousands of lncRNAs, validation of the function of these transcripts remains a challenge and to date, the function and mechanisms of only a few of these transcripts have been elucidated. However, in order to design effective therapeutics, we must study the function of lncRNAs, elucidate their mechanisms, determine their structures, and implement approaches to quantitatively assay for drugs that modulate their interactions.

The data presented in this dissertation provides strong evidence that long noncoding RNA-protein interactions can be quantified by means of a pharmacological assay. Furthermore, our lncRNA-protein interaction assay is the first quantitative assay that is amenable to high throughput screening for small molecule modulators. This work provides key insights into lncRNA-protein interactions and highlights the need for RNA secondary structure studies to better understand these interactions.
REFERENCES


