ENGINEERING OF FUNCTIONALLY ASYMMETRIC SIRNA DUPLEXES

By

Phillip Angart

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Chemical Engineering - Doctor of Philosophy

2017

ABSTRACT

ENGINEERING OF FUNCTIONALLY ASYMMETRIC SIRNA DUPLEXES

By

Phillip Angart

Short interfering RNAs (siRNAs) are a promising nucleic acid-based therapeutic strategy that offer an alternative to traditional small-molecule based drugs for the treatment of a variety of otherwise-untreatable diseases. These small, chemically synthesized RNAs can transiently inhibit the expression of target genes through the use of a native eukaryotic pathway called RNA interference (RNAi). The design of highly active siRNAs is not straightforward. Our strategy to improve siRNA design criteria focuses on the idea that siRNAs enter the RNAi pathway as duplexes but only become functional once one of the siRNA strands has been selected. Ensuring high selectivity of the intended strand is essential for proper siRNA functionality. To address this challenge, this research aimed to (1) understand how features of the siRNA control the relative activities of the two strands of the siRNA duplex, termed functional asymmetry and (2) better understand the mechanism of siRNA asymmetric strand selection.

To improve our understanding of the characteristics that drive siRNA strand selection and strand activity, we began our investigation with two characteristics, relative terminal hybridization stability ($\Delta\Delta$G) and $5'$ terminal nucleotide (TN) Rank. These characteristics have been previously shown to cooperatively predict siRNA activity. Our analysis indicates that these characteristics are also predictive of siRNA functional asymmetry. A comparative analysis between individual strand loading and activity shows that the $\Delta\Delta$G reflects a combination of duplex hybridization stabilities (ΔGs) that are important in the

formation of RISC and its kinetics. The TN was found to influence siRNA strand activity post-siRNA loading, suggesting a role in maximizing RISC half-life. Taken together, these results indicate that siRNA activity is influenced by siRNA-protein interactions that occur both pre- and post-siRNA strand selection.

The mechanism of siRNA loading and the protein-RNA interactions responsible for driving differential strand loading are not fully understood. Based on evidence from other systems, we hypothesized that asymmetric strand selection was driven, at least in part, by preferential binding of different segments of siRNAs by RNAi pathway proteins. Here we demonstrated that one RNAi protein, PACT, preferentially localizes to one siRNA terminus of a duplex known to be functionally asymmetric. This indicates that PACT may thereby influence siRNA asymmetric strand loading during initiation of RNAi.

The work presented here identified characteristics of siRNAs duplexes that were found to drive siRNA strand selection and activity. Collectively, these results inform siRNA design and address biological questions about the functional aspects of the RNAi pathway. Future work in this area will continue to investigate how characteristics of siRNAs impact their functionality within the RNAi pathway, ultimately leading towards a highly refined set of rules for siRNA design. A more guided approach to siRNA design will facilitate the development of siRNAs as a therapeutic platform.

To my family

ACKNOWLEDGMENTS

I am grateful to all of the people who have helped in the preparation of this dissertation and my graduate education.

Thank you to my advisor, Professor S. Patrick Walton for guidance, support, and friendship. I am grateful for the opportunity I have been given.

To my committee members, Professors Christina Chan, Charles Hoogstraten, and Timothy Whitehead for their integral part in developing the work presented here. I appreciate both your time and effort in guiding me through this process.

To the undergraduates I have had the pleasure of working with and mentoring: Amanda Lindgren, Rebecca Jacobs, Kwasi Adu-Berchie, Sarah Thorwall, and Rebecca Carlson. Thank you for your efforts, without which this work would not be possible.

Thank you to the past and present members of the Cellular and Biomolecular Engineering lab, who have been friends as well as coworkers. A special note of thanks to Aritro Nath, Hyun Cho, Amanda Malefyt and Betul Bilgin for your mentorship in the beginning of my graduate career.

To all my friends, who have listened to me talk about my research and never understood it, at least you tried. You are too numerous to name but I am thankful for your support.

Finally, to my brother, mother, and father, none of this would have been possible without your love and support.

$\sim$Thanks.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

A	adenine
Ago	Argonaute
ANOVA	analysis of variance
bp	base pair
B. Choshinensis	*Brevibacillus choshinensis*
BDNF	brain derived neurotrophic factor
BSA	bovine serum albumin
C	cytidine
C3PO	component 3 promoter of RISC
cDNA	coding DNA
CHO	Chinese hamster ovary
clp1	cleavage and polyadenylation factor I subunit
CV	column volume
$\Delta\Delta$G	relative terminal hybridization stability
ΔG	hybridization stability
DGCR8	DiGeorge syndrome chromosomal region 8
dNTP	deoxynucleotide triphosphate mixture
DNA	reoxyribonucleic acid
ds	double-stranded
DTT	diothiothreitol
E. coli	*Escherichia coli*

EDTA	ethylenediaminetetraacetic acid
EGFP	enhanced green fluorescent protein
EMSA	electrophoetic mobility shift assay
G	guanine
HEK	human embryonic kidney
His	histidine
HRP	horse radish peroxidase
HSP	heat shock protein
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl -D-1-thiogalactopyranoside
isomiR	miRNA isoform
MBP	maltose binding protein
Medipal	Medipal Merlin-Dicer-PACT liaison domain
miRNA	microRNA
mRNA	messenger RNA
N	any nucleotide
ND	not determined
NGF	nerve growth factor
NS3	non-structural protein 3
nt	nucleotide
OAS1	$2'$-$5'$-oligoadenylate synthetase 1
OD_{600}	optical density at 600 nm

p75	low-affinity nerve growth factor receptor
PACT	PKR activator
PAGE	polyacylamide gel electrophoresis
PAZ	Piwi-Argonaute-Zwille
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PIWI	P-element induced wimpy testis
PKR	protein kinase R
pp-luc	*Photinus pyralis* (firefly) luciferase
pre-miRNA	precursor-miRNA
pre-RISC	precursor-RISC
RBD	RNA binding domain
RBP	RNA binding protein
RIG-I	retinoic acid-inducible gene I
RISC	RNA-induced silencing complex
RK_{13}	rabbit kidney epithelial cells
RNA	ribonucleic acid
RNAi	RNA interference
RT-qPCR	reverse transcriptase-quantitative polymerase chain Reaction
SD	standard deviation
SDS-PAGE	sodium dodecyl sulfate-PAGE
SF21	*Spodoptera frugiperda* ovarian cells

shRNA	small hairpin RNA
siRNA	short-interfering RNA
sod	Cu, Zn superoxide dismutase
ss	single-stranded
TAR	Trans-activating response
TLR	Toll-like receptor
T_m	melting temperature
TN	$5'$ terminal nucleotide
TRBP	TAR RNA binding protein
TrkB	tropomyosin-related kinase B
U	uridine
UTR	untranslated region
VEGF	vascular endothelial growth factor
W	A or U nucleotide

Chapter 1

Introduction

1.1 Significance

Biologics are a class of pharmaceuticals that utilize biological macromolecules to achieve a therapeutic effect. As alternatives to traditional small molecule drugs, biologics have the potential to provide treatments for previously untreatable conditions. Antisense technologies are biologics that utilize nucleic acids to target RNA (and occasionally DNA) within a cell to alter gene expression. One antisense technology, RNA interference (RNAi) therapeutics, leverages the ability of a native eukaryotic regulatory mechanism to incorporate chemically-synthesized, small interfering RNAs (siRNAs) to direct gene silencing against a therapeutic target. siRNAs provide a transient, highly targeted method of post-transcriptional gene silencing. siRNAs are routinely used in research studies of eukaryotic biological processes, though transitioning the technology to the clinic has proven challenging. Principal among the challenges are the delivery of siRNAs to the cells of interest and, the focus of this work, the selection of highly active siRNAs with good specificity.

Significant effort has been spent identifying factors that enhance siRNA activity and specificity. siRNA activity is influenced by, among other factors, strand selection, the structure of the mRNA target region, base preferences, overall siRNA G/C content, and siRNA duplex thermodynamics. In comparison, siRNA specificity depends on strand selection,

immunogenicity, and uniqueness of the target sequence. Current siRNA design depends upon achieving a balance among these factors; siRNA design algorithms typically weight these factors based upon analyses of siRNA activity data (Naito and Ui-Tei, 2012). As it stands, the rules for selecting active, specific siRNAs continue to evolve.

siRNAs enter the RNAi pathway as duplexes but must be single-stranded in order to function, requiring selection of one siRNA strand for incorporation into the active complex. Asymmetric strand selection is critical for proper functionality of siRNAs against an intended target and occurs by recognition of specific features of the siRNA by the proteins of the RNAi pathway. The work described here focuses on understanding the siRNA characteristics that lead to a functionally asymmetric and active siRNA duplex. We also investigated the Protein Kinase R (PKR) Activator (PACT), which is postulated to facilitate siRNA loading, for its ability to act as an asymmetry sensor in the RNAi pathway.

1.2 Background

RNAi is a central post-transcriptional regulatory and defense mechanism in many eukaryotic organisms and is essential to understanding the regulatory landscape of eukaryotic cells. Studies of RNAi have informed mechanisms of developmental disorders, disease parthenogenesis, and cancer progression. It is also remarkable for the relative ease in which exogenous siRNAs, once delivered to the cytoplasm, can enter the RNAi pathway and constitute targeted gene silencing, creating a therapeutic effect.

RNAi was first observed in petunias when the gene that produces anthocyanin (the compound giving petunias their purple color) was overexpressed, in an attempt to enhance the flowers' color (Napoli and Lemieux, 1990). Instead, many flowers turned white instead

of deeper purple. It was later determined that the petunias turned white due to activation of the RNAi pathway, which recognized the overexpressed gene as non-native and silenced its expression, concomitantly silencing the native gene (Napoli and Lemieux, 1990). The discovery of both the RNA trigger and explanation of the post-transcriptional silencing effect was subsequently published, earning the senior authors the Nobel Prize in Physiology and Medicine in 2006 (Fire et al., 1998). In 2001, RNAi was characterized in mammals, with the simultaneous discovery of multiple native, small regulatory RNAs (Elbashir et al., 2001a; Lagos-Quintana, 2001). Native small RNAs have been found to be central in development, cancers, infections, and other diseases and remain an active area of research (Carthew and Sontheimer, 2009; Wittrup and Lieberman, 2015).

1.2.1 Small RNAs of RNAi

The RNAi pathway processes and utilizes small RNAs to enact sequence-specific post-transcriptional gene silencing and translational repression (Carthew and Sontheimer, 2009). The small RNA trigger confers specificity of the active protein-RNA complex towards a messenger RNA (mRNA) target with a complementary sequence (Meister et al., 2004). For enacting RNAi, small RNAs can be grouped into two categories, siRNAs and microRNAs (miRNAs). siRNAs are fully base paired $\sim$21 nucleotide (nt) duplexes with 19 internal base pairs (bp), $5'$ phosphates, and $3'$ dinucleotide overhangs (Figure 1.1) (Bernstein et al., 2001; Zamore et al., 2000; Nykänen et al., 2001; Elbashir et al., 2001c). miRNAs, while similar in overall architecture, typically contain bulges and mismatches within the duplex and only partially base pair with their target (Figure 1.1) (Bartel, 2009). The pathway itself is made up of relatively few proteins, but the diversity in native small RNAs enables

the regulation of $> 60\%$ of human genes (Friedman et al., 2008).

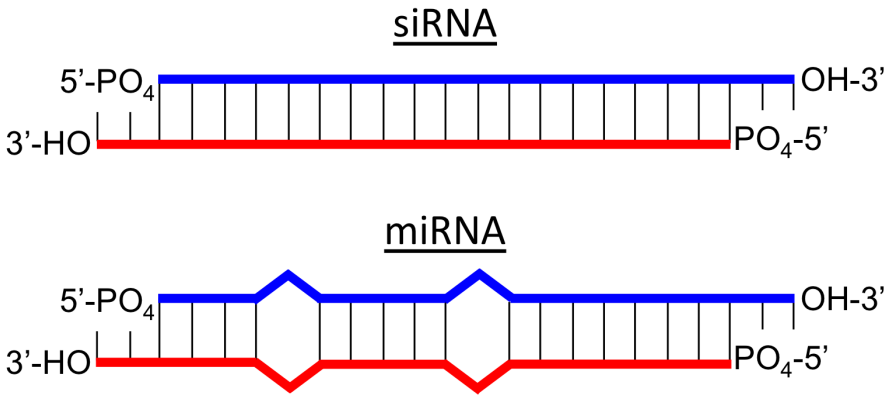


Figure 1.1: Anatomy of siRNAs and miRNAs. (Top) The canonical siRNA structure with ∼19 bp, and 3′ dinucleotide overhangs. (Bottom) Approximate miRNA structure, similar to siRNAs, with ∼19 internal bases and 3′ dinucleotide overhangs, but also containing bulges and mismatches within the duplex.

1.2.2 Details of the RNA Interference Pathway

In mammals, miRNA biogenesis begins in the nucleus with the transcription of a primary-miRNA, which is then processed by the Drosha-DiGeorge syndrome chromosomal region 8 (DGCR8) complex to form a precursor-miRNAs (pre-miRNA) (Gregory et al., 2004; Han et al., 2004) (Figure 1.2). Pre-miRNAs are exported to the nucleus by exportin-5 (Yi et al., 2003) where they are further processed into miRNAs by the Dicer complex, composed of Dicer, and/or TRBP and PACT, by cleavage of the pre-miRNA hairpin (Yoda et al., 2010). miRNAs are then loaded into one of the 4 Argonaute (Ago) proteins, forming pre-RNA induced silencing complex (pre-RISC) (Yoda et al., 2010). Once the passenger strand is removed, either by duplex unwinding or by cleavage and degradation of one of the strands, mature RISC is formed (Matranga et al., 2005; Yoda et al., 2010; Kawamata et al., 2009). Strand selection depends on its orientation within Ago, dictating which strand will be

loaded (guide strand) and which strand will be degraded (passenger strand) (Matranga et al., 2005; Rand et al., 2005). The guide RNA-Ago complex composes the minimal RISC, capable of targeting complementary mRNA for gene knockdown via Watson-Crick base pairing (Rivas et al., 2005). miRNA programmed RISC bind with partial complementarity to the $3'$ UTR of their target mRNA, which, through a variety of mechanisms, causes either translational repression or transcript decay (Reviewed in Bartel (2009) and Ameres and Zamore (2013)).

Endogenous siRNA biogenesis begins with long dsRNA or short-hairpin RNA (shRNA), which directly enter the Dicer complex and are cleaved into siRNA duplexes. While siRNAs are predominately loaded into Ago1 and Ago2, only Ago2 is capable of efficiently forming mature RISC (Su et al., 2009; Gu et al., 2011; Yoda et al., 2010; Meister et al., 2004) through endonucleolytic cleavage and removal of the fully complementary passenger strand (Matranga et al., 2005; Rand et al., 2005; Yoda et al., 2010). siRNA programmed RISC binds to complementary mRNA and, in a similar manner to passenger strand cleavage and removal, cleaves the mRNA between the nucleotides complementary to the 10^{th} and 11^{th} bases of the guide strand (Wang et al., 2008; Elbashir et al., 2001b). The cleaved mRNA is then released and degraded, allowing RISC to target another mRNA (Haley and Zamore, 2004). Ago2 is the only human Ago protein with an active ribonuclease domain making it essential to siRNA-mediated knockdown (Yoda et al., 2010).

Exogenous siRNAs are typically designed to mimic the mature endogenous siRNA structure (Elbashir et al., 2001a; Zamore et al., 2000). Once siRNAs are delivered to the cytoplasm, the $5'$ termini are rapidly phosphorylated by the protein Clp1 (Weitzer and Martinez, 2007) and are then loaded into a protein complex to form pre-RISC (Yoda et al.,

2010; Sakurai et al., 2011). The only protein required for pre-RISC is one of the Ago proteins (Yoda et al., 2010), although, by itself, Ago2 is unable to form mature RISC (Ye et al., 2011). It is unclear as to what proteins form pre-RISC in living cells; however, recombinant protein based assays have shown that pre-RISC activity can be formed *in vitro* with Ago2 and either TRBP, Dicer, HSP70/HSP90, or C3PO (Willkomm et al., 2016; Bernard et al., 2015; Iwasaki et al., 2010; Liu et al., 2009; Ye et al., 2011).

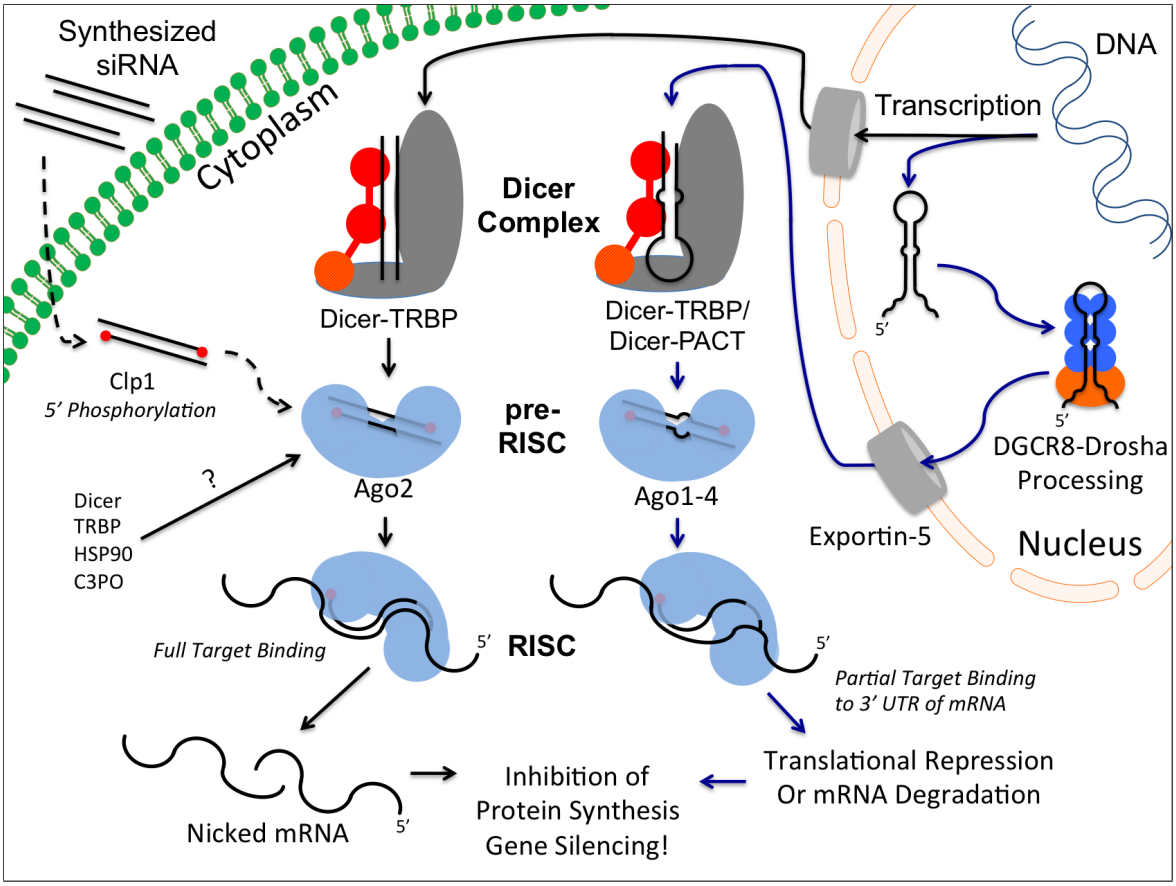


Figure 1.2: RNAi Mechanism. Endogenous miRNA biogenesis (blue lines) and siRNA biogenesis (black lines). Exogenous siRNAs enter the pathway through the pre-RISC (dash black lines), the components of which are not completely known. Not depicted are pre-miRNAs that are directly loaded into Ago2 constituting pathway "cross-talk" (Yang et al., 2010; Cheloufi et al., 2010), the localization of RNAi proteins (Stalder et al., 2013), regulatory elements of RNAi (Ha and Kim, 2014), and the mechanisms of miRNA mediated gene silencing (Bartel, 2009), and mechanisms of siRNA entry into the cell (Malefyt et al., 2012).

1.2.3 Ago2

Ago2 is the core catalytic protein of siRNA-mediated RISC (Rivas et al., 2005). It is one of four human Ago proteins but it is the only one capable of loading siRNAs and cleaving its mRNA target through its active PIWI domain (Figure 1.3) (Meister et al., 2004; Yoda et al., 2010). Ago2 interacts directly with the siRNA phosphate backbone, $5'$ terminal nucleotide (TN), and the 2 nt of the $3'$ end of the guide strand, which, before siRNA strand separation, composed the $3'$ overhang of the guide strand (Frank et al., 2010; Ma et al., 2004; Parker et al., 2009).

Ago2, as part of RISC, splits the guide RNA into different domains, with each domain having a different role in RISC function (Figure 1.4). RISC exposes nucleotides 2-4 of the guide strand to the solvent (1° seed), using these bases to scan for complementary targets. Once a complementary 3 nt sequence is found, bases 5-8 engage the mRNA target (2° seed) (Salomon et al., 2015; Chandradoss et al., 2015). Nucleotides 2-8 of the guide strand are the minimum sequence required to maintain a stable association between RISC and its target; this region of the guide strand is referred to as the seed region (Parker et al., 2009; Ameres et al., 2007; Lambert et al., 2011). Ago proteins have an adenine binding pocket, which, in circumstances where the nucleotide opposite the 1^{st} nucleotide of the guide strand is an adenine, alters the minimum required seed region to nucleotides 2-7 (Bartel, 2009). Seed region binding initiates significant conformational changes to allow binding of the remaining guide strand to the target mRNA followed by cleavage of the mRNA between the 10^{th} and 11^{th} bases of the guide strand (part of the central domain) (Elbashir et al., 2001b; De et al., 2013; Wang et al., 2008; Schirle and MacRae, 2012). Only nucleotides 2-16 are involved in full target recognition, with nucleotides 13-16 acting to stabilize the RISC-

mRNA interaction ($3'$ supplement) (Wee et al., 2013). The first nucleotide of the guide RNA interacts directly with the MID domain of Ago2 (Figure 1.3) and does not engage in base pairing with the target (TN; Figure 1.4) (Frank et al., 2010). The $5'$ nucleotide binding pocket interacts directly with the $5'$ phosphate and the phosphate backbone of nucleotides 1-4 (Suzuki et al., 2015; Frank et al., 2010); it also contains a loop structure that discriminates against the TN sequence, showing preference for uridine and adenine over guanine and cytidine nucleotides. Functionally, these binding preferences affect the half-life of RISC (De et al., 2013). This same binding pocket facilitates the loading of thermodynamically unstable small RNA termini (Suzuki et al., 2015) (See Chapter 2). The $3'$ tail of the guide strand interacts directly with the PAZ and N-terminal domains (Figure 1.3) and does not engage in target binding (Ma et al., 2004; Lingel et al., 2004; Kwak and Tomari, 2012).

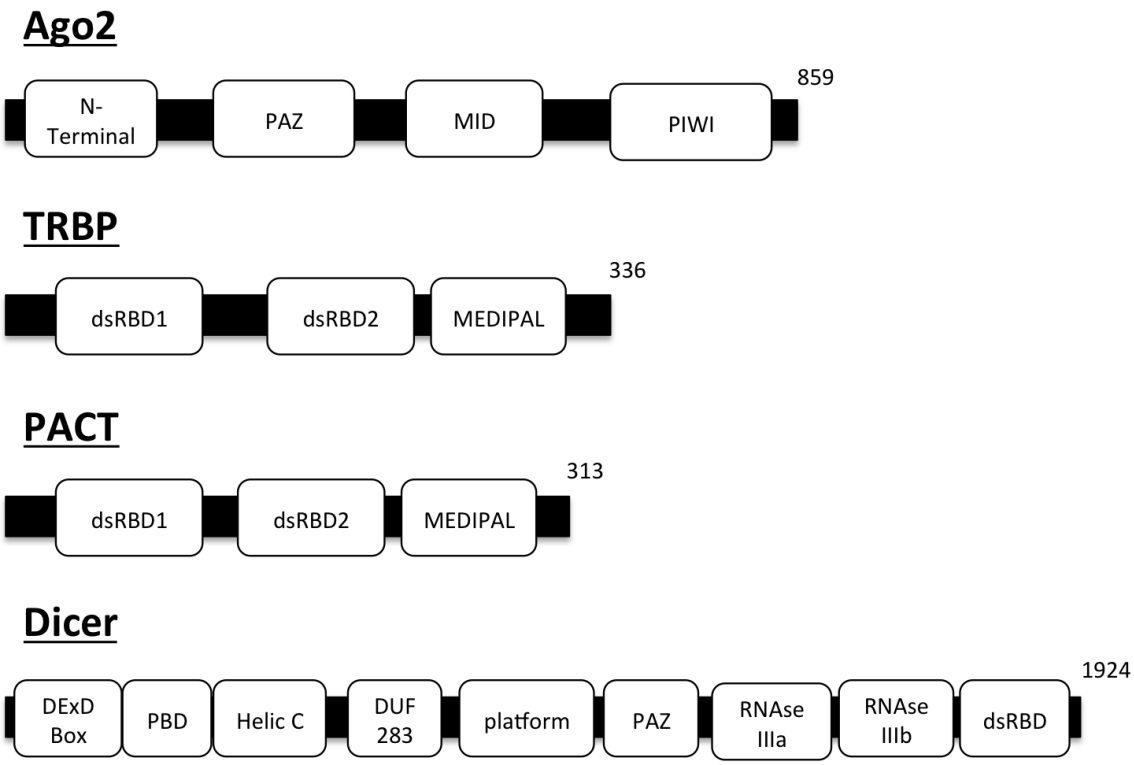


Figure 1.3: Domains of RNAi Proteins. Depicted are the domains of the essential RNAi cytoplasmic proteins with the number on the C-terminus denoting the number of amino acids in each protein.

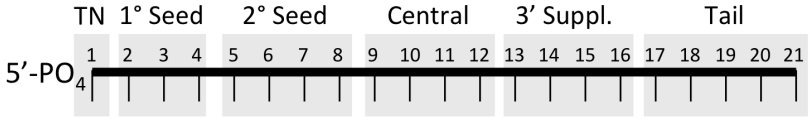


Figure 1.4: Functional Domains of Guide RNA in RISC. Guide RNA-mRNA interactions vary depending on the region of the guide strand and are modulated by Ago2. These interactions have been mapped in *Drosophila* and mouse Ago2.

1.2.4 TRBP and PACT

TRBP and PACT are dsRBPs with similar architectures, containing two double stranded RNA binding domains (dsRBDs) and a third domain that binds to a number of other proteins (Figure 1.3), including Dicer (Laraki et al., 2008; Daniels et al., 2009). The 2 dsRBDs

of TRBP and PACT act independently of one another, with dsRBD2 having greater affinity for dsRNA compared to dsRBD1 (Daviet et al., 2000). Within the context of the Dicer complex (Figure 1.2), TRBP and PACT act both independently and coordinately to facilitate the generation of siRNAs and miRNAs, controlling substrate loading and isomiR generation in (Lee and Doudna, 2012; Wilson et al., 2015; Kim et al., 2014). These proteins also have differences in substrate binding preferences, with PACT showing affinity for pre-miRNA like structures and TRBP having affinity for both dsRNA and pre-miRNAs (Lee et al., 2013).

Both TRBP and PACT function in contexts outside of the RNAi, playing a role in translational control, cell growth and cell cycle, and PKR activation/suppression (Daniels and Gatignol, 2012). In each case, TRBP and PACT interact with each other, as well as the proteins PKR and Merlin; collectively these proteins constitute a signaling network between cellular RNA and transcription (Daniels and Gatignol, 2012). These proteins are also regulated post-transcriptionally through phosphorylation, ubiquination, and SUMOylation effecting their function within the RNAi pathway (Paroo et al., 2009; Chen et al., 2015; Lee et al., 2004, 2005).

1.3 siRNA Design

siRNAs are designed to be fully complementary with their target RNA; still, sequence selection is not necessarily straightforward. This is due to the number of critical inter- and intramolecular interactions within the RNAi pathway that determine the eventual activity of the siRNA in silencing its target. The following sections discuss factors that influence siRNA sequence selection and additional modifications that can be made to siRNAs to

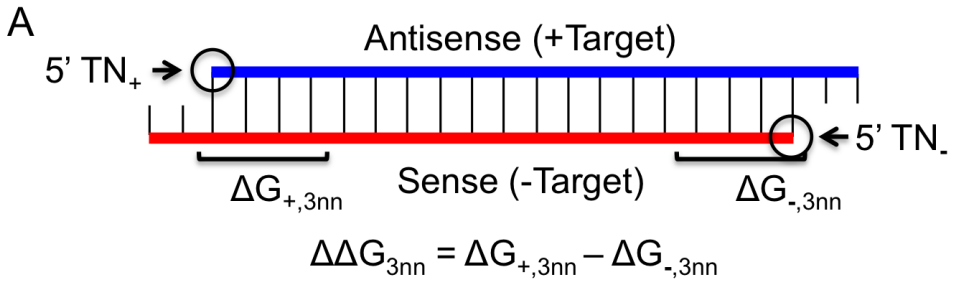
improve efficacy.

1.3.1 Functional Asymmetry

For siRNAs to silence the intended target, they must be properly oriented to ensure incorporation of the intended guide strand into RISC (Schwarz et al., 2003; Tomari et al., 2004; Gredell et al., 2010; Noland et al., 2011; Sakurai et al., 2011; Yoda et al., 2010). Incorporation of the unintended strand (i.e., the intended passenger strand) leads to the formation of a RISC that cannot cleave the intended target (reducing the activity of the siRNA therapeutic) and potentially causing off-target effects (reducing the specificity of the siRNA) (Schwarz et al., 2003; Khvorova et al., 2003; Grimm et al., 2006). Selection of an siRNA target region largely dictates the sequence of both the guide and passenger siRNA strands; because of this reduction in the degrees of freedom, it is useful to refer to siRNA activity both in terms of its absolute activity and relative activity over the passenger strand. We will refer to the difference in the activity of one siRNA strand relative to the other as *functional asymmetry* (Schwarz et al., 2003). Functional asymmetry is a function of multiple factors, including RISC half-life, RISC turnover rate, and biased incorporation of one siRNA strand into RISC (Frank et al., 2010; Haley and Zamore, 2004; Tomari et al., 2004; Suzuki et al., 2015; De et al., 2013).

Two factors hypothesized to influence siRNA functional asymmetry are the relative terminal thermodynamic hybridization stability ($\Delta\Delta$G) and the relative $5'$ terminal nucleotide (TN) sequence (Figure 1.5A), which can be ranked according to their predicted activity, TN Rank (Figure 1.5A and 1.5B) (Walton et al., 2010; Malefyt et al., 2013). Both of these characteristics are localized to the termini of the siRNA. The siRNA termini are

critical for authentication of the siRNA structure not only with the characteristic 2 nt $3'$ overhangs and $5'$ phosphates (Elbashir et al., 2001a; Nykänen et al., 2001), but also because these are the only regions of the siRNA guide known to interact with Ago2 and Dicer (Lima et al., 2009; Kini and Walton, 2007) in a manner other than the generic phosphate backbone interactions (Frank et al., 2010; Lingel et al., 2004; Suzuki et al., 2015; Schirle and MacRae, 2012).



B

TN Rank	5' TN+: 5' TN-
1	U:G
2	A:G
3	U:C
4	U:A
5	A:C
6	U:U
7	C:G
8	A:A
9	G:C
10	G:G
11	A:U
12	C:C
13	G:A
14	G:U
15	C:A
16	C:U

Figure 1.5: $\Delta\Delta G_{3nn}$ and TN Rank Definitions. (A) siRNA asymmetry characteristics marked on an siRNA with the top strand (blue) as the desired guide strand (Antisense) and the bottom strand (red) as the desired passenger strand (Sense). (B) TN rankings (from Malefyt et al. (2013)). Arrow denotes TN Ranks predicted to favor greater antisense (+) strand activity (Blue) or favor greater sense (-) strand activity (Red) (+/- Notation described in Figure 2.2).

1.3.1.1 Relative Terminal Hybridization Stability ($\Delta\Delta G$)

Biased strand incorporation occurs by the recognition of differences in the termini of the siRNAs by various proteins of the RNAi pathway (Tomari et al., 2004; Schwarz et al., 2003;

Khvorova et al., 2003; Gredell et al., 2010; Walton et al., 2010; Noland et al., 2011; Betancur and Tomari, 2012; Koh et al., 2013). Based on early studies of functional asymmetry in *Drosophila*, it was concluded that a $\Delta\Delta G_{4nn}$, the difference in hybridization energy between the two ends of the siRNA (similar to Figure 1.5A), was predictive of functional asymmetry and further attributed to biased strand loading (Tomari et al., 2004). These studies determined that the strand whose $5'$ terminus is less stably hybridized (more positive hybridization free energy) is preferentially loaded into RISC (Schwarz et al., 2003; Tomari et al., 2004; Gredell et al., 2010; Noland et al., 2011). More recent analyses have found that functional asymmetry in mammalian systems is more strongly predicted by the TN, with the $\Delta\Delta G_{4nn}$ being a second order effect to TN (Walton et al., 2010; Malefyt et al., 2013; Suzuki et al., 2015) (See Chapter 2). When both factors are combined, a $\Delta\Delta G_{3nn}$ is more informative than a $\Delta\Delta G_{4nn}$ definition (Walton et al., 2010).

1.3.1.2 $5'$ Terminal Nucleotide Sequence (TN)

The TN strongly predicts siRNA activity, with $5'$ uridine and adenine nucleotides being strongly favored over $5'$ cytidine and guanine nucleotides (Frank et al., 2010; Seitz et al., 2011). The TN contributes to RISC stability via a direct interaction with the nucleotide specificity loop within Ago2 (Frank et al., 2010; De et al., 2013). The increased affinity of some guide strands relative to others may result in an increase in RISC stability, the half-life of RISC, and subsequently higher strand activity (Frank et al., 2010; De et al., 2013). TN Rank, an extension of TN, is a measure of siRNA activity based on the TN of both the passenger and guide strands (Figure 1.5A and 1.5B) (Walton et al., 2010; Gredell et al., 2010; Malefyt et al., 2013). Despite the partial collinearity between TN Rank and $\Delta\Delta G_{3nn}$, these features are independent predictors of siRNA activity (Walton et al., 2010;

Malefyt et al., 2013; Suzuki et al., 2015).

1.3.2 mRNA Target Region

Strong mRNA secondary structure precludes siRNA binding through steric hindrance of RISC binding and cleavage (Brown et al., 2005; Ameres et al., 2007). *In silico* methods exist to predict, relatively accurately, mRNA secondary structures (Vickers et al., 2003; Ameres et al., 2007; Bohula et al., 2003; Overhoff et al., 2005; Schubert et al., 2005; Shao et al., 2007; Yoshinari et al., 2004); however, mRNA structure in a living cell is dynamic and predictive approaches can only provide partial mRNA structural information. Secondary structure information can inform siRNA design and, in general, regions of the mRNA with low overall secondary structure should be targeted. (Ameres et al., 2007; Overhoff et al., 2005; Shao et al., 2007).

1.3.3 Immunogenicity

siRNAs can cause a number of immunologic and cytotoxic responses, some of which are sequence-specific and others that arise due to the dsRNA structure (Sledz et al., 2003; Samuel-Abraham and Leonard, 2010; Jackson and Linsley, 2010; Robbins et al., 2007). In fact, immunogenic responses can be the cause of the therapeutic effect, rather than specific silencing mediated by the siRNA (Haussecker, 2012; Schlee et al., 2006). In general, the recognition of specific sequence motifs that leads to immune activation is mediated by toll-like receptor (TLR) recognition of pathogen-associated molecular patterns (Heil et al., 2004; Judge et al., 2005; Jackson et al., 2006a; Diebold et al., 2004, 2006; Goodchild et al., 2009; Weber et al., 2012; Kleinman et al., 2008; Reynolds et al., 2006; Kariko et al., 2004; Forsbach

et al., 2008). Immunostimulatory motifs include: GUCCUUCAA (Hornung et al., 2005), UGUGU (Judge et al., 2005), UGU Judge2005, UCA Jurk2011, GU-rich sequences (Heil et al., 2004), AU-rich sequences (Forsbach et al., 2008), and U-rich sequences (Goodchild et al., 2009). UGGC has been shown to be cytotoxic via non-immune mechanisms (Fedorov et al., 2006). OAS1, PKR, and RIG-I are cytoplasmic receptors that interact with dsRNAs, like siRNAs, to enact an innate immune response (Kodym et al., 2009; Manche et al., 1992; Bevilacqua and Cech, 1996; Marques et al., 2006; Kato et al., 2008; Samuel-Abraham and Leonard, 2010; Gantier and Williams, 2007). These receptors detect RNA structural features rather than sequence motifs. The only exception is OAS1, which is activated by a NNWW$(N)_9$WGN motif, where W can be either a A or U nucleotide and N is any nucleotide (Kodym et al., 2009).

1.3.4 Non-Specific Effects

miRNA-like targeting occurs when siRNAs have unintended seed region complementarity with the $3'$ UTR of an mRNA, resulting in translational repression or transcriptional silencing of an untargeted transcript (Doench et al., 2003; Lambert et al., 2011; Gu et al., 2011; Bartel, 2009; Lewis et al., 2003; Lin et al., 2005; Lai, 2002). mRNAs that are natively regulated by miRNAs are also more susceptible to miRNA-like off-targeting due to extended $3'$ UTRs and the preexistence of miRNA target sites (Schultz et al., 2011). Avoiding miRNA-like targeting effects is complicated by an inability to accurately predict seed sequences (Didiano and Hobert, 2006; Schultz et al., 2011; Lin et al., 2005). miRNA-like targeting is influenced by the surrounding sequence of the target, the position of the target in the mRNA, and the repetitiveness of the target sequence (Doench and Sharp,

2004; Lin et al., 2005; Broderick et al., 2011; Bartel, 2009). As such, the best way to account for miRNA-like targeting is to avoid seed sequences that have already been identified (http://www.mirbase.org) (Kozomara and Griffiths-Jones, 2011). miRNA-mediated silencing is dose dependent (Bartel, 2009); as siRNA potency increases and lower doses of siRNAs are required, miRNA-like off-target effects from siRNAs should concomitantly decrease.

1.3.5 Other siRNA Design Criteria

Other factors shown to influence siRNA activity include the G/C content (i.e., the overall duplex stability) of the siRNA (Reynolds et al., 2004; Vert et al., 2006; Khvorova et al., 2003), the secondary structure of the guide strand (Patzel et al., 2005; Köberle et al., 2006), internal repeats (Reynolds et al., 2004), palindromic sequences (Hossbach et al., 2006), and positional base preferences along the siRNA (Reynolds et al., 2004; Ui-Tei et al., 2004; Jagla et al., 2005; Huesken et al., 2005; Ladunga, 2006; Shabalina et al., 2006; Amarzguioui and Prydz, 2004; Gong et al., 2006; Takasaki et al., 2004; Holen, 2006; Takasaki, 2009; Katoh and Suzuki, 2007; Vermeulen et al., 2005). More recently, additional structural criteria, even to the level of tertiary structure (Sciabola et al., 2012), have been identified as valuable in predicting siRNA activity. While each of these factors may be important in siRNA design, either their overall influence is thought to be minimal compared to the other selection criteria or there is a lack of consensus on how to implement the feature as a selection criterion. As such, the use of the mentioned criteria are generally considered as a second order set of rules.

1.3.6 Non-Canonical siRNA Structural Designs

Using non-canonical siRNA structures in large part changes the point of entry into the RNAi pathway or changes the way in which the RNAi proteins interact with the non-canonical siRNA (Snead and Rossi, 2012). From the bottom up, ssRNAs are capable of being loaded by Ago2, albeit inefficiently, to form an active RISC *in vitro* (Martinez et al., 2002; Holen et al., 2003; Birmingham et al., 2006; Rivas et al., 2005) and *in vivo* when chemically modified (Lima et al., 2012; Haringsma et al., 2012). By only providing one siRNA strand to enter RISC, proper strand selection is no longer a design characteristic. Other structures are designed to bias strand incorporation by altering the length of the passenger strand (asymmetric interfering RNAs and asymmetric short-duplex siRNAs (Chu and Rana, 2008; Sun et al., 2008; Chang et al., 2009)), the length of the $3'$ overhang (fork-siRNAs (Hohjoh, 2004)), or by assembling a duplex with a segmented passenger strand (small internally segmented interfering RNAs (Bramsen et al., 2007)). A more recently tested structure utilizes a bulge at the second position of the guide strand, creating a perturbation at the first base of the seed region (Dua et al., 2011); the use of this modification to the canonical siRNA structure was shown to decrease off-targeting by miRNA-like activity. Other non-canonical structures exploit the ability for longer RNAs to enter RISC more efficiently; these structures are called Dicer substrate RNAs (Amarzguioui et al., 2006; Collingwood et al., 2008; Tanudji et al., 2010; Foster et al., 2012).

1.3.7 Incorporation of Chemical Modifications

Chemical modifications in siRNAs can increase their stability (specifically with regards to nuclease degradation), minimize immunogenicity, and, to a certain extent, improve the

activity of the siRNA (Turner et al., 2006; Volkov et al., 2009; Hong et al., 2010; Robbins et al., 2007; Hornung et al., 2005; Goodchild et al., 2009; Allerson et al., 2005; Bramsen et al., 2009; Kenski et al., 2012). Chemical modifications are also necessary for the use of siRNAs in clinical applications. Chemical modifications of the siRNA focus on changing the phosphodiester backbone, ribose sugar, nucleotide base, and 2′-OH ribose group. Effective use of chemical modifications requires the substitution of the new chemical moiety at a position within the siRNA where the additional group or structural alteration does not inhibit normal siRNA function. In general, the rationale behind chemical modifications is to incorporate small perturbations in the siRNA structure to prevent recognition and/or binding of the siRNA by nucleases and the immune receptors for RNA. The most common modifications include altering the 2′-OH group to a 2′-O-CH_3 or 2′-F to prevent recognition of the RNA by nucleases and TLR7 and TLR8 (Braasch et al., 2003; Chiu and Rana, 2003; Allerson et al., 2005; Manoharan et al., 2011; Cekaite et al., 2007; Robbins et al., 2007; Tluk et al., 2009; Fucini et al., 2012). TLR3 interacts with dsRNA, and, while TLR3 activation by canonical unmodified siRNAs has been shown in mice, it may not be an issue in humans (Reynolds et al., 2006; Kariko et al., 2004; Kleinman et al., 2008; Weber et al., 2012). Unfortunately, no absolute rules exist defining which chemical modifications are most useful and how they are best applied.

The 3′ overhangs are also a common location for chemical modifications for two reasons: i) they may provide a site of attack for endoribonucleases and ii) chemical modifications, even bulky ones, are typically well tolerated at these positions (Haupenthal et al., 2006). Phosphorothioate and phosphorodithioate modifications to the backbone of the siRNA generate siRNAs with nuclease resistance, but the number and positions of modifications

are important in retaining siRNA activity (Braasch et al., 2003; Amarzguioui, 2003; Yang et al., 2012).

1.4 Approach and Specific Aims

The work described here aimed to improve the design of siRNAs for both higher activity and specificity by focusing on the refinement of the features that determine siRNA functional asymmetry. This dissertation details two approaches taken towards understanding siRNA duplex characteristics and molecular interactions that lead to a highly functional asymmetric siRNA duplex.

The specific aims addressing the question of siRNA functional asymmetry are outlined below:

1. Understand the interrelation of TN Rank and $\Delta\Delta G_{3nn}$ on siRNA strand loading and siRNA functional asymmetry.

TN Rank and $\Delta\Delta G_{3nn}$ are characteristics of siRNA duplexes are effective predictors of siRNA activity. These characteristics were shown here to also predict of siRNA functional asymmetry. We further explored how these characteristics impact siRNA strand loading and found that $\Delta\Delta G_{3nn}$ term, while useful for predicting siRNA functional asymmetry, actually contains predictive information about the siRNA duplex energies that are important for siRNA loading and for describing siRNA-mRNA interactions. Discrimination of the TN was also found to occur post-loading. These findings provide a more accurate description of how duplex thermodynamics influence both siRNA loading and activity.

2. Characterize the extent to which PACT can sense asymmetric characteristics of siRNA duplexes

The proteins TRBP and PACT are mammalian homologs of the protein R2D2, a protein credited with sensing siRNA duplex asymmetry in *Drosophila*. TRBP has been previously shown to sense siRNA duplex asymmetry (Gredell et al., 2010). Here we investigated the asymmetry sensing of the PACT protein, showing that it too can asymmetrically localize to siRNA termini of functionally asymmetric siRNA duplexes. Furthermore, we were able to show that PACT terminus localization was not identical to TRBP for all of the sequences examined, suggesting that TRBP and PACT do not function in an identical manner. Lastly, we found that PACT siRNA terminus localization can be predicted by $\Delta\Delta G_{4nn}$.

Chapter 2

Terminal duplex stability and nucleotide identity differentially control siRNA loading and activity in RNA interference

2.1 Abstract

Efficient short interfering RNA (siRNA)-mediated gene silencing requires selection of a sequence that is complementary to the intended target and possesses sequence and structural features that encourage favorable functional interactions with the RNA interference (RNAi) pathway proteins. Here, we investigated how terminal sequence and structural characteristics of siRNAs contribute to siRNA strand loading and silencing activity, and how these characteristics ultimately result in a functionally asymmetric duplex in cultured HeLa cells. Our results reiterate that the most important characteristic in determining siRNA activity is the $5'$ terminal nucleotide identity. Our findings further suggest that siRNA loading is controlled principally by the hybridization stability of the $5'$ terminus (Nucleotides: 1-2) of each siRNA strand, independent of the opposing terminus. Post-loading, RISC spe-

cific activity was found to be improved by lower hybridization stability in the $5'$ terminus (Nucleotides: 3-4) of the loaded siRNA strand and greater hybridization stability towards the $3'$ terminus (Nucleotides: 17-18). Concomitantly, specific recognition of the $5'$ terminal nucleotide sequence by human Ago2 improves RISC half-life. These findings indicate that careful selection of siRNA sequences can maximize both the loading and the specific activity of the intended guide strand.

2.2 Introduction

The RNA interference (RNAi) pathway is a central regulatory mechanism in mammalian cells, where endogenous small RNAs, once incorporated into a ribonucleoprotein complex, target complementary sequences and cause post-transcriptional gene silencing (Snead and Rossi, 2010; Carthew and Sontheimer, 2009). Short interfering RNAs (siRNAs), a subset of small RNAs recognized by the RNAi pathway, are selectively loaded into the Argonaute 2 (Ago2) protein (Yoda et al., 2010), an endoribonuclease (Liu et al., 2004), and enact gene silencing through binding and cleaving complementary mRNAs (Martinez et al., 2002; Elbashir et al., 2001c). The ability to design and deliver exogenous siRNAs for targeted protein knockdown offers an attractive therapeutic strategy (Angart et al., 2013).

RNAi is initiated when the pathway proteins recognize RNA duplexes with the characteristic siRNA structure, $\sim$19 internal base pairs and 2 nt $3'$ overhangs (Elbashir et al., 2001c; Zamore et al., 2000). Either strand from the duplex can be processed and loaded into Ago2 to form the mature RNA-induced silencing complex (RISC) (Rivas et al., 2005). The quantity of intended RISC formed and its activity determine the efficacy of an siRNA in silencing its target. While generally only one siRNA strand (the guide strand) can tar-

get a specific mRNA, in systems where targets exist for both siRNA strands, both strands can silence (Yoda et al., 2010; Sakurai et al., 2011). Even in the absence of a specific target, loading and function of the unintended strand (the passenger strand) can result in significant non-specific, off-target effects (Jackson et al., 2003, 2006b). Thus, to achieve maximal activity and specificity of silencing, it is essential to design siRNAs for loading and activity of only the intended guide strand. To do so requires an understanding of how to control i) selection and loading of the guide strand and ii) the half-life and activity of RISC. Concomitantly, these same design rules should be applied to minimizing the loading and activity of the passenger strand, thereby achieving maximal functional asymmetry, which we define as the ratio of the silencing activities of the siRNA strands.

The concept of asymmetry has been essential to designing siRNAs since their initial structure (double-stranded) and their functional form (single-stranded) were determined (Schwarz et al., 2003). Factors that differentiate between the two siRNA strands have been investigated, with the objective of determining how the RNAi pathway proteins identify the guide and passenger strand of each siRNA. The first feature discovered for predicting siRNA asymmetry was relative terminal hybridization stability ($\Delta\Delta$G) (Figure 1.5A) (Schwarz et al., 2003), which captures the relative stability of the duplex structures at the siRNA termini. $\Delta\Delta$G was shown to predict the functional asymmetry of siRNAs (Schwarz et al., 2003; Lu and Mathews, 2008) and their asymmetric binding by proteins of the RNAi pathway (Tomari et al., 2004; Noland et al., 2011). In our prior work, we classified siRNA asymmetries according to the $5'$ terminal nucleotide (TN) on each strand (TN Rank) (Figure 1.5) (Gredell et al., 2010; Malefyt et al., 2013; Walton et al., 2010), showing this feature to be more predictive of siRNA activity than $\Delta\Delta$G.

However, using activity as the readout minimizes the degree to which identified features can be linked to specific steps in the RNAi mechanism. Our goals in the present study were to understand whether TN Rank and $\Delta\Delta$G predict siRNA strand selection in addition to activity and to quantify the degree to which strand selection determines the final activity of an siRNA strand. Our results showed that both TN Rank and $\Delta\Delta$G predict asymmetric strand loading, albeit less accurately than predicting siRNA functional asymmetry. While the influence of TN on strand activity is principally post-loading, presumably through interactions between the guide strand and Ago2, hybridization stability, particularly at the $5'$ terminus, influences both strand selection and post-loading events (e.g., interactions with the mRNA target). The findings presented here increase our understanding of siRNA structure-function relationships and can be useful in the selection of active and specific siRNAs.

2.3 Results

2.3.1 TN Rank and $\Delta\Delta G_{3nn}$ Influence siRNA Functional Asymmetry in Cultured HeLa Cells

Our previous work described the importance of TN Rank and $\Delta\Delta G_{3nn}$ (Figure 1.5) in predicting siRNA activity (Malefyt et al., 2013). The set of siRNAs chosen to investigate TN Rank and $\Delta\Delta G_{3nn}$ was selected to possess favorable, neutral, and unfavorable rankings of the two features, with favorable referring to a characteristic that would be associated with higher silencing of the siRNA against its target mRNA, in this case PKR (Figure 2.1) (Malefyt et al., 2013). Here, we wished to determine if these parameters could also

predict siRNA functional asymmetry, as the most highly active and specific siRNAs will have functional asymmetries highly biased in favor of the intended guide strand.

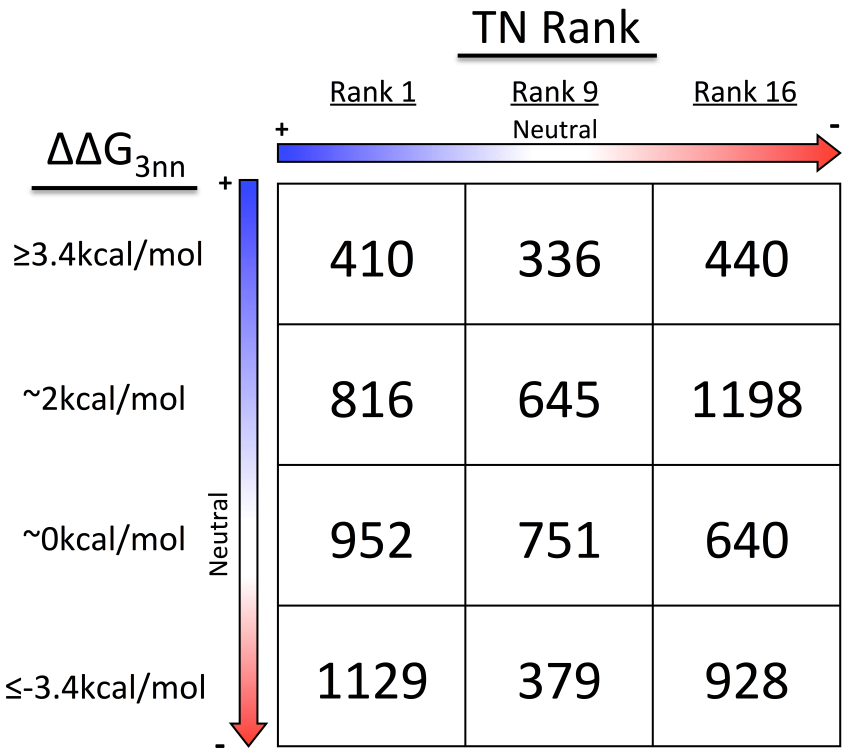


Figure 2.1: Characteristics of PKR-targeting siRNAs. Feature predicted to favor greater antisense (+) strand activity (Blue), feature predicted to favor greater sense (-) strand activity (Red). siRNAs are named by the $5'$ position on the PKR mRNA targeted by the siRNA (+) strand. siRNA sequences are provided in Table A.1.

We initially measured the activity of each siRNA strand using a luciferase reporter assay where the full-length coding sequence of PKR (hereafter the (+) target, targeted by the (+) strand of each siRNA) or the full-length complement sequence of PKR (hereafter the (-) target, targeted by the (-) strand of each siRNA) was cloned as a fusion product downstream of the *Renilla* luciferase gene (Figure 2.2). siRNA strand IC_{50} values were determined by co-transfecting HeLa cells with one of the plasmids and one of the PKR-targeting siRNAs. Functional asymmetry was calculated as the ratio of IC_{50} values for

each pair of complementary strands (Figure 2.3A). Two siRNA strands, PKR816(-) and PKR928(+), did not display any measurable activity (Figure 2.4); for these siRNAs functional asymmetry was estimated assuming an IC_{50} of 10 nM, the highest concentration tested, for the inactive strand. Examining our results for both TN Rank and $\Delta\Delta G_{3nn}$, we find that, as we found with predicting strand activity, TN Rank is a stronger predictor of functional asymmetry than $\Delta\Delta G_{3nn}$ (Figures 2.3B and 2.3C), with the information provided by $\Delta\Delta G_{3nn}$ improving the correlation but being insufficient to provide a significant correlation alone (Figures 2.3C and 2.3D). The importance of the $\Delta\Delta G_{3nn}$ is most evident for sequences with intermediate TN Rank, where positive $\Delta\Delta G_{3nn}$ sequences favor (+) strand activities while negative $\Delta\Delta G_{3nn}$ sequences favor (-) strand activities (Figure 2.3A).

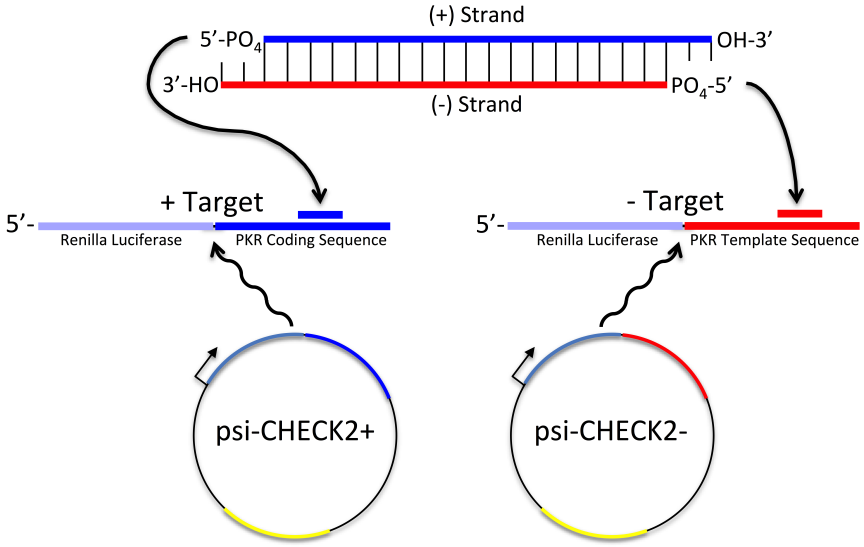
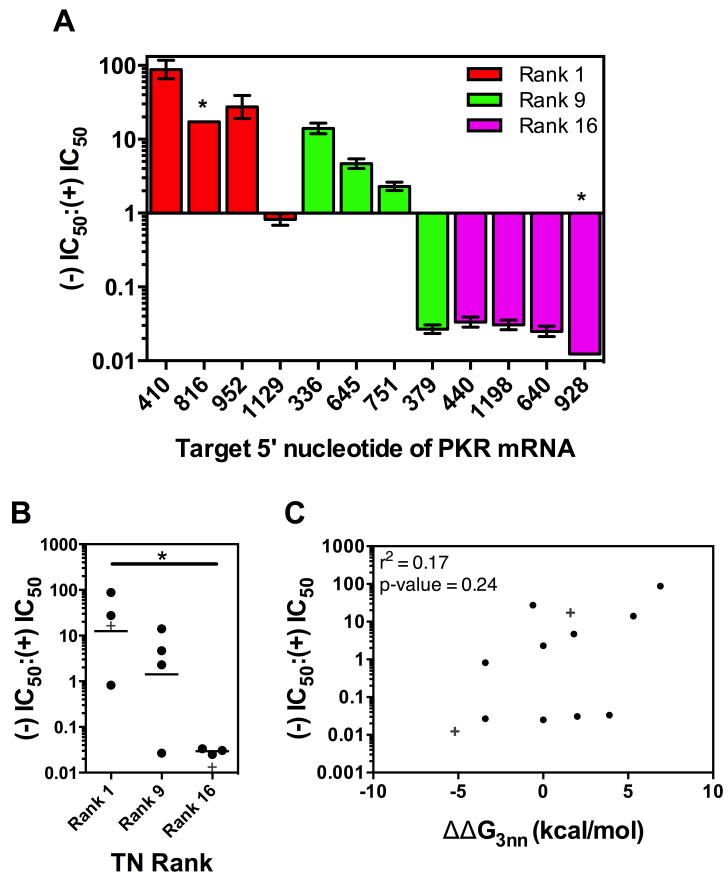


Figure 2.2: Schematic of siRNA Luciferase Assay. Luciferase reporter constructs were designed with either the PKR coding sequence, (+) target (Blue), or PKR template sequence, (-) target (Red), cloned downstream of the *Renilla* luciferase gene. Reporter constructs were transfected independently with each PKR targeting siRNA to determine the activity of each siRNA strand.



D

Variables (X)	Coefficients	Std Error	p-value	r^2	p-value
Rank 1	1	---	---		
Rank 9	-0.94	0.43	7.0×10^{-2}	0.89	2.7×10^{-3}
Rank 16	-2.84	0.46	8.2×10^{-4}		
$\Delta\Delta G_{3nn}$	0.21	0.055	8.3×10^{-3}		

$Y = (\text{Rank 1})(X_1) + (\text{Rank 9})(X_2) + (\text{Rank 16})(X_3) + (\Delta\Delta G_{3nn})(X_4) + C$

Figure 2.3: siRNA Functional Asymmetry. (A) Relative siRNA functional asymmetry sorted left to right by TN Rank and then by decreasing $\Delta\Delta G_{3nn}$. Shown are the ratios of siRNA (-) to (+) strand IC_{50} values determined from curve fitting to 4 biological replicates. Error bars are $\pm$ 1 SD. (*) represents values where one siRNA strand IC_{50} was not measurable and assumed to be 10 nM (the highest concentration tested) for calculating functional asymmetry. (B) Distribution of siRNA functional asymmetry with respect to TN Rank. Horizontal lines represent the mean activity within the TN Rank; overhead bar marked with (*) represents a statistically different pairwise comparison ($p < 0.05$) by one-way ANOVA with Tukey's *post hoc* analysis. (C) Correlation of functional asymmetry with $\Delta\Delta G_{3nn}$. (B, C) siRNAs PKR816 and PKR928 had one estimated IC_{50} and were marked with (+) and not included in any statistical analysis. (D) Multiple linear regression of siRNA functional asymmetry with TN Rank and $\Delta\Delta G_{3nn}$, where Y is the functional asymmetry ((-) IC_{50}:(+) IC_{50}) and X_i are the variables. Coefficients were normalized to Rank 1.

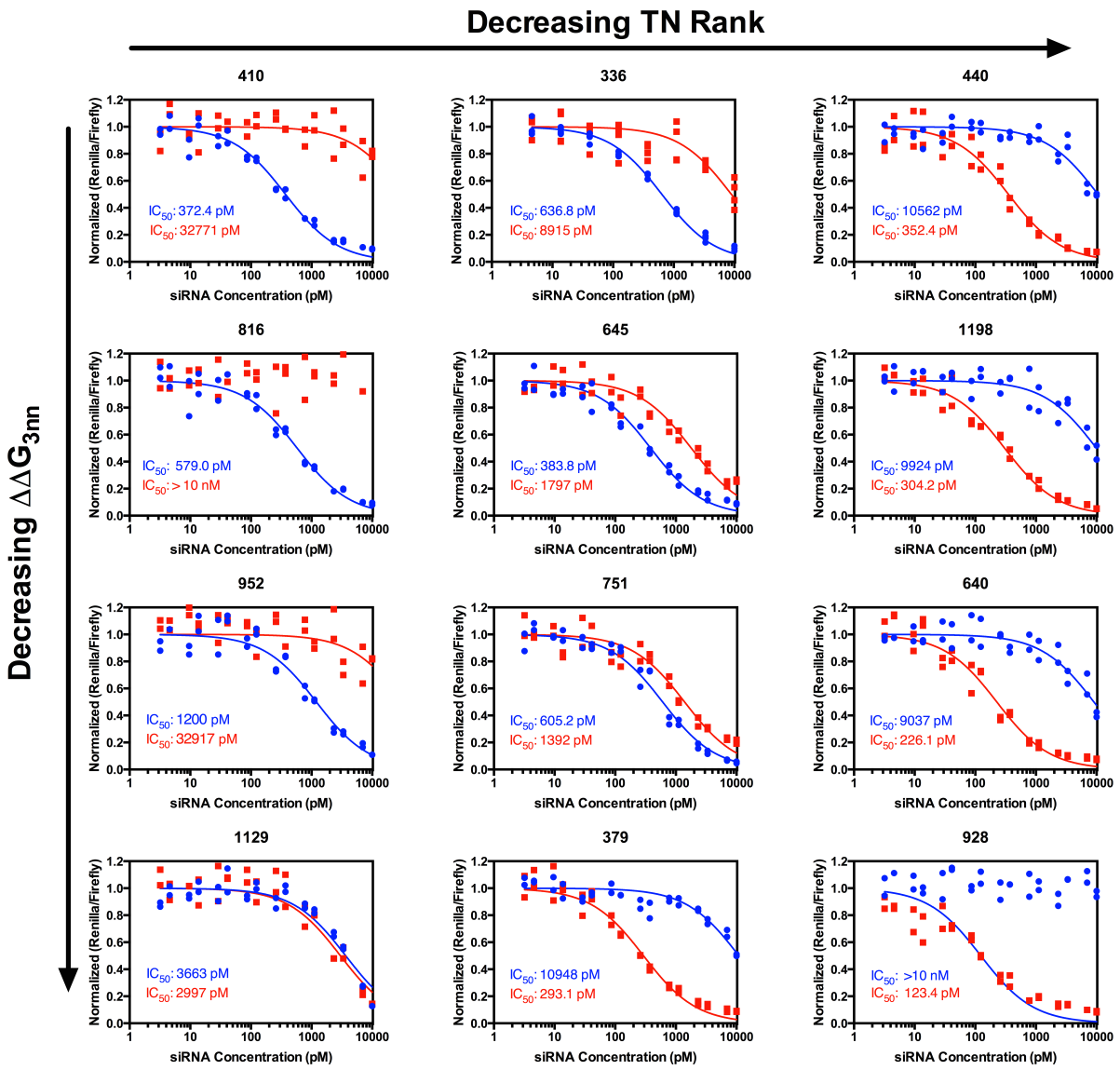
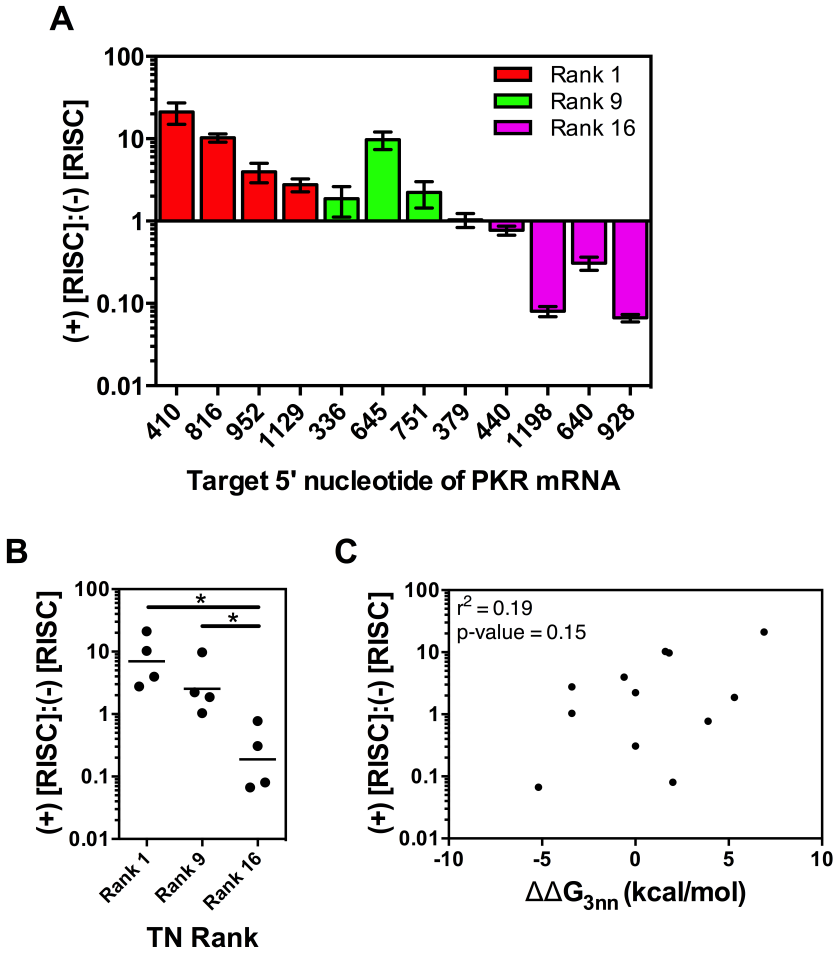


Figure 2.4: IC_{50} Curves for PKR-targeting siRNAs. (+) siRNA strand is shown in blue and the (-) siRNA strand in red.

2.3.2 Asymmetric siRNA Strand Loading

To better understand the steps of the mechanism that, on aggregate, result in functional asymmetry, we measured siRNA strand loading into Ago2 by transfecting HeLa cells with our siRNA of interest for 24 h, immunoprecipitating Ago2, and then measuring its asso-

ciated RNA by stem loop RT-qPCR. We observed that TN Rank correlated with siRNA asymmetric strand loading (Figures 2.5A and 2.5B). $\Delta\Delta G_{3nn}$ was not correlated with siRNA asymmetric strand loading (similar to its lack of correlation with functional asymmetry) (Figure 2.5C). In this case, however, it provided considerably less predictive information when combined with TN Rank (Figure 2.5D) than it did for relative strand activity, as shown by the relative differences in Akaike weights for the combined 2-factor models and their respective single factor models (i.e., a smaller difference in Akaike weights indicates a smaller gain in model information content; Table 2.1).



D

Variables (X)	Coefficients	Std Error	p-value	r^2	p-value
Rank 1	1	---	---		
Rank 9	-0.98	0.56	1.2×10^{-1}	0.87	7.8×10^{-4}
Rank 16	-3.44	0.56	2.8×10^{-4}		
$\Delta\Delta G_{3nn}$	0.174	0.066	3.1×10^{-2}		

$Y = (\text{Rank 1})(X_1) + (\text{Rank 9})(X_2) + (\text{Rank 16})(X_3) + (\Delta\Delta G_{3nn})(X_4) + C$

Figure 2.5: siRNA Asymmetric Strand Loading. (A) Relative siRNA strand loading into Ago2 sorted left to right by TN Rank and then by decreasing $\Delta\Delta G_{3nn}$. Shown are the ratios of siRNA (+) strand to (-) strand loading. N = 3, error bars are ± 1 SD. (B) Distribution of siRNA asymmetric loading by TN Rank. Horizontal lines represent the mean activity within the TN Rank; overhead bars marked with (*) represent statistically different pairwise comparisons ($p < 0.05$) by one-way ANOVA with Tukey's *post hoc* analysis. (C) Correlation of siRNA asymmetric strand loading with $\Delta\Delta G_{3nn}$. (D) Multiple linear regression of asymmetric strand loading with TN Rank and $\Delta\Delta G_{3nn}$, where Y is the relative strand loading ((+) [RISC]:(-) [RISC]) and X_i are the variables. Coefficients were normalized to Rank 1.

Table 2.1: Akaike Weights for Linear Regression Models. ΔAICc and Akaike weight comparisons of 1- and 2- factor linear regressions using TN Rank and $\Delta\Delta G_{3nn}$.

	(+) IC_{50}:(-) IC_{50}		(+) [RISC]:(-) [RISC]	
Model	ΔAICc	Akaike Weight	ΔAICc	Akaike Weight
TN Rank	14.9	0.001	1.1	0.365
$\Delta\Delta G_{3nn}$	16.8	0	10.4	0.003
Combined	0	0.999	0	0.632

For a number of siRNAs, we observed differences in the relative loadings and relative activities. PKR1129, which exhibited symmetrical strand activity ($p = 0.24$, extra sum-of-squares F test), had asymmetric strand loading ($p = 0.02$, two-tailed t-test) in line with its TN Rank. Interestingly, it is the (-) strand of PKR1129 that has an unexpectedly high activity compared to the other sequences with a $5'$ G and similar loading (PKR410(-) and PKR952(-); Figures 2.6A and 2.6B). This is presumably because this sequence has a $\Delta\Delta G_{3nn}$ strongly in favor of the (-) siRNA strand, which may partially offset its unfavorable $5'$ nucleotide. Additionally, sequences PKR336, PKR379, and PKR440 displayed symmetrical strand loadings and asymmetric activities (Figures 2.3A and 2.5A). As with PKR1129, these siRNAs have large $\Delta\Delta G_{3nn}$ values (either highly positive or negative), suggesting that $\Delta\Delta G_{3nn}$ may significantly influence siRNA function only after exceeding a threshold magnitude.

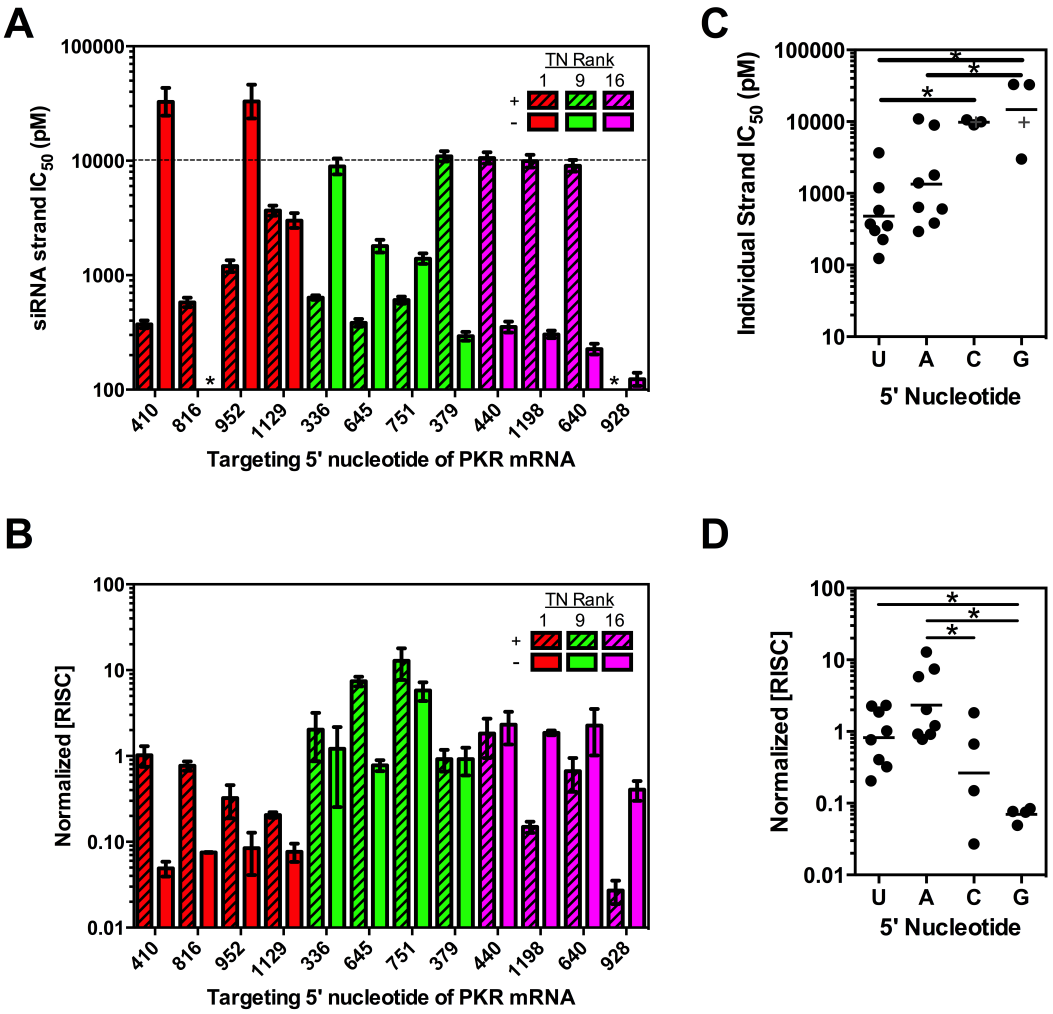


Figure 2.6: Individual siRNA Strand Activities and Loading. (A) Individual siRNA strand IC_{50} values (pM); N = 4; error bars are ± 1 SD. Dotted line represents maximum siRNA concentration assayed, values above the line are extrapolated from experimental data. No activity was observed for siRNA strand PKR816(-) and PKR928(+); denoted with (*). (B) Individual siRNA strand loading normalized to an internal standard; N = 3; error bars are ± 1 SD. TN Ranks correspond to nucleotide pairs presented in Figure 1.5B. siRNA strand activity (C) and loading (D) grouped by TN. Horizontal lines represent the mean activity within the TN; overhead bars marked with (*) represent statistically different pairwise comparisons ($p < 0.05$) by one-way ANOVA with Tukey's *post hoc* analysis

A direct comparison of siRNA strand loading with its activity demonstrates that while siRNA loading and activity are correlated, multiple strands diverge significantly from the trend (Figure 2.7). These findings suggest some modulation of activity post-siRNA loading.

Furthermore, the change in Akaike weights between the individual factors (TN Rank and $\Delta\Delta G_{3nn}$) and the combined model indicates that TN Rank and $\Delta\Delta G_{3nn}$ are complementary predictors of functional asymmetry (Table 2.1). In comparison to activity, $\Delta\Delta G_{3nn}$ adds less complementary information to the prediction of asymmetric strand loading (Table 2.1), indicating that $\Delta\Delta G_{3nn}$ contains information more relevant to post-loading events. For clarity, as we discuss post-loading events, we will use RISC specific activity to refer to the enzymatic characteristics of a mature RISC that are independent of the amount of RISC generated.

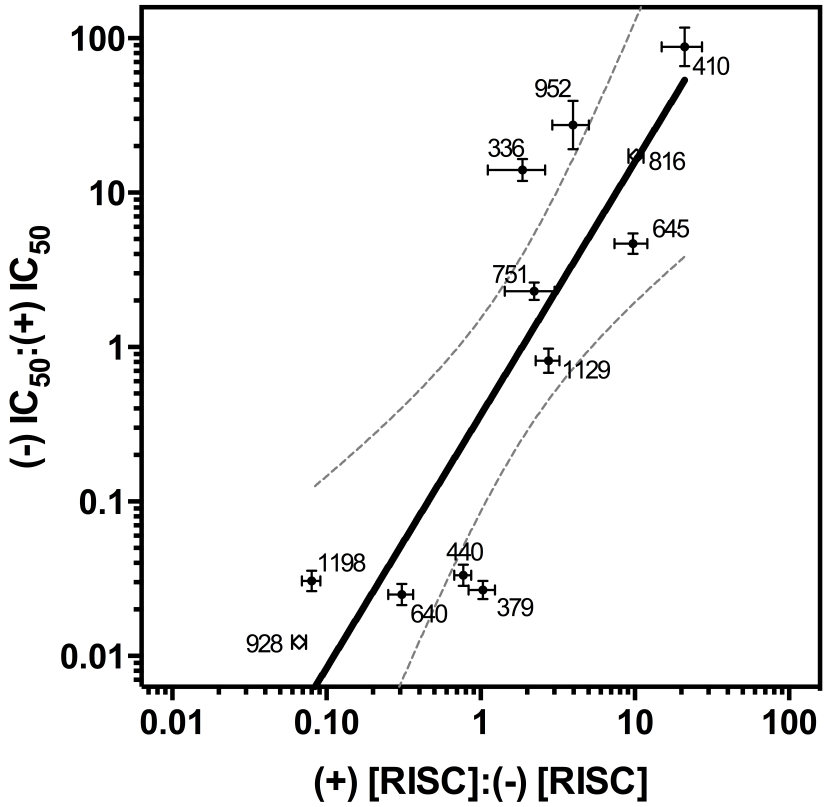


Figure 2.7: Relationship Between siRNA Functional Asymmetry and Asymmetric Loading. Functional asymmetry values are the same as Figure 2.3A and asymmetric loading values are the same as Figure 2.5A. Solid black line is a semi-log fit to the data. Dotted grey lines are the 95% confidence intervals on the fit. Diamonds (PKR816 and PKR928) represent points for which one of the IC_{50} values could not be fit and was estimated to be 10 nM, the highest concentration tested.

2.3.3 RISC Specific Activity

Because RISC specific activity is a feature of the siRNA after the separation of the two siRNA strands, we analyzed the behavior of the individual siRNA strands independent of their complements (Figure 2.8). Likewise, in lieu of TN Rank and $\Delta\Delta G_{3nn}$, we examined TN (U, A, C, G) and the energies of all nearest-neighbor pairs ($\Delta G_{(x)-(y)nt}$) along the length of the duplex (Figure 2.9). While similar, the nucleotide preferences for loading

(A>U>C>G) (Figure 2.6D) are different from those for activity (U>A>C>G) (Figure 2.6C), indicating that TN impacts RISC specific activity, as others have previously observed (De et al., 2013). However, TN alone does not account for all of the observed variation in activities; for instance, siRNAs with 5′ adenines with similar loading varied substantially in their activities (Figure 2.8).

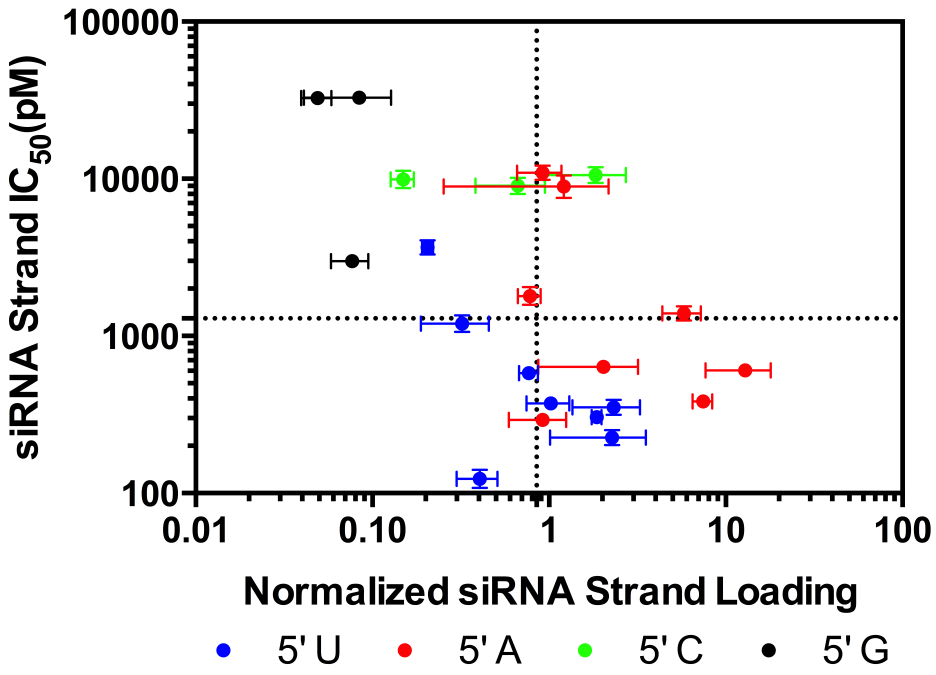


Figure 2.8: Correlation of siRNA Strand Activity with Ago2 Loading. X values are fitted IC_{50} values; N = 4; error bars are ± 1 SD. Y values are the normalized average siRNA loading; N = 3; error bars are ± 1 SD. Dotted lines represent the median of each data set.

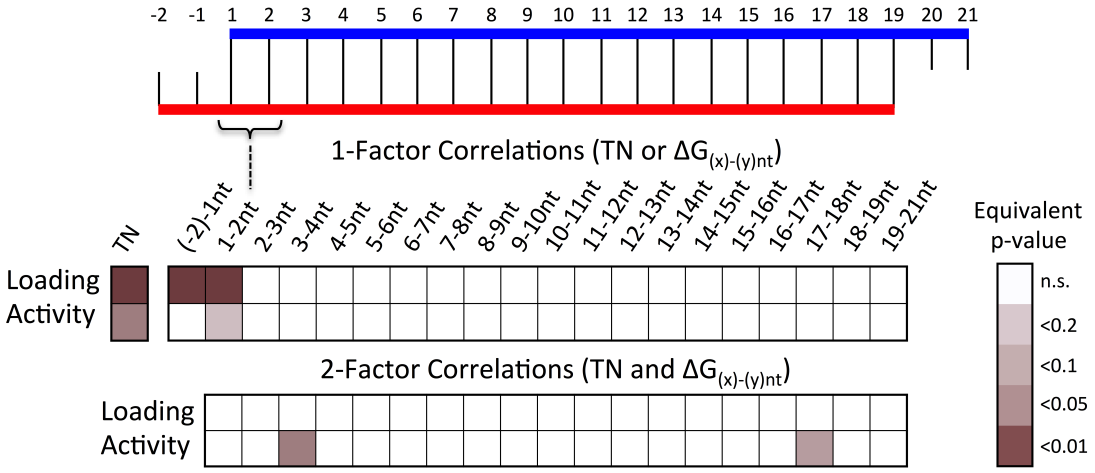


Figure 2.9: Evaluation of siRNA Loading and Activity Predictors. 1-factor and 2-factor correlations of siRNA loading and activity with TN and $\Delta G_{(x)-(y)nt}$ along the length of the duplex. 2-factor correlations use TN as one variable and a $\Delta G_{(x)-(y)nt}$ value as the other. Significance in these correlations indicates that the variable predicts loading or activity independent of TN. A Bonferroni correction was applied to correct for multiple comparisons.

To identify regions of the siRNA duplex important for controlling RISC specific activity, we correlated loading and activity with the TN and ΔG's along the length of the duplex, looking for regions in the duplex important for siRNA activity but not important for RISC loading (Figure 2.9). We found that only the nearest-neighbor parameters surrounding the TN, $\Delta G_{(-2)-1nt}$ and ΔG_{1-2nt}, correlated with activity (Figure 2.9). To decouple the effects of the TN sequence from ΔG, we used multiple linear regression and re-evaluated the correlation between ΔG and both siRNA loading and activity. We found that ΔG_{3-4nt} and $\Delta G_{17-18nt}$ were only predictive of siRNA activity not loading (Figure 2.9). This indicates that RISC specific activity is improved for sequences with weaker (less negative) ΔG_{3-4nt} and stronger (more negative) $\Delta G_{17-18nt}$. These ΔG values are included in the calculation of $\Delta\Delta G_{3nn}$ and likely contribute to its utility in predicting siRNA activity.

2.4 Discussion

Designing siRNAs for maximal function requires the double-stranded structure to efficiently enter the RNAi pathway such that it can be easily processed to RISC loaded with the correct single-strand. It is therefore critical to understand how the RNAi pathway proteins sense specific features of siRNA duplexes that differentiate the two siRNA strands. siRNA functional asymmetry was first characterized in *Drosophila* lysates and attributed to the difference in $\Delta\Delta G_{4nn}$ (Schwarz et al., 2003). Many Argonaute proteins contain an additional structure that preferentially binds $5'$ terminal nucleotides (Frank et al., 2010). Our previous work sought to explain how these two terminal features contribute to siRNA activity (Walton et al., 2010; Malefyt et al., 2013). Here, we showed how TN Rank and $\Delta\Delta G_{3nn}$ contribute to siRNA functional asymmetry and asymmetric strand loading. Furthermore, we have shown that, in cultured HeLa cells, functional asymmetry is modulated both before and after siRNA loading, with TN Rank and $\Delta\Delta G_{3nn}$ exerting influence at different stages of the mechanism.

Functional asymmetry and asymmetric strand loading were not perfectly correlated, indicating additional modulation of siRNA activity post-siRNA loading (Figure 2.7). Our top-down analysis identified free energies at two positions, ΔG_{3-4nt} and $\Delta G_{17-18nt}$, to be predictive of siRNA activity but not siRNA loading (Figure 2.9). Mature RISC splits siRNA-target mRNA interactions into 5 domains, the $5'$ anchor (Nucleotide: 1), the seed region (Nucleotides: 2-8), the central region (Nucleotides: 9-12), the $3'$ supplemental region (Nucleotides: 13-16), and the $3'$ tail (Nucleotides: 17-21) (Wee et al., 2013). The preference for weaker thermodynamic stability through nucleotides 1-4 of the siRNAs is well established (Schwarz et al., 2003; Reynolds et al., 2004; Amarzguioui and Prydz, 2004;

Hibio et al., 2012); our findings suggest that weaker hybridization in this region leads to greater RISC specific activity (Figure 2.9), in agreement with findings that high stability in the seed region decreases RISC turnover (Salomon et al., 2015). RISC-mRNA interactions beyond the $3'$ supplemental region have been shown to be dispensable for target cleavage or in some cases even deleterious (De et al., 2013; Wee et al., 2013). We contend that the preference for a stronger $\Delta G_{17-18nt}$ is also important in maintaining a stable interaction between RISC and its target, similar to base pairs in the $3'$ supplemental region (Wee et al., 2013).

After delivery of the siRNA to the cell cytoplasm, efficient guide strand loading requires nucleotide-independent phosphorylation of the siRNA $5'$ terminal nucleotides (Weitzer and Martinez, 2007), outcompeting native miRNAs for binding to Ago2 (Koller et al., 2006; Tanudji et al., 2010), conformational changes within Ago2 facilitated by HSP90 and potentially other proteins (Iwasaki et al., 2010; Sakurai et al., 2011; Bernard et al., 2015; Willkomm et al., 2016; Ye et al., 2011), and cleavage and removal of the passenger strand (Matranga et al., 2005; Leuschner et al., 2006). We observed a large variation in the total amount of Ago2 loading even among siRNA strands with the same terminal nucleotide (Figure 2.6B). In analyzing other characteristics for their ability to predict loading, we found $\Delta G_{(-2)-1nt}$ and ΔG_{1-2nt} of the $5'$ end of the guide strand to be the most predictive features of siRNA strand loading (Figure 2.9). Interestingly, there was no correlation between siRNA strand loading and the ΔG values near the $3'$ terminus (i.e., Nucleotides 17-21), suggesting that the interactions of Ago2 are primarily with the $5'$ ends of siRNAs. Thus, parameters based on relative strand properties (e.g., TN Rank and $\Delta\Delta G$) are less useful in predicting absolute siRNA strand loading and activity than in predicting relative

strand activities and loadings.

The established binding preferences for the Ago2 nucleotide specificity loop (U>A>C>G) match the order of siRNA activities in our data (Figure 2.6C) but differ from the loading preferences we observed (A>U>C>G) (Figure 2.6D). This suggests that $5'$ TN effects RISC specific activity, presumably through differences in RISC half-life (De et al., 2013), similar to what was observed in *Drosophila* Ago1 (Kawamata et al., 2011). The correlation of siRNA loading with the $5'$ terminal $\Delta G_{(-2)-1nt}$ and ΔG_{1-2nt} indicates that siRNA loading is controlled in part by the $5'$ terminal duplex energy (Figure 2.9), though TN remains a significant predictor of siRNA strand loading because of its collinearity with the $\Delta G_{(-2)-1nt}$ and ΔG_{1-2nt} terms. Recently, sensing of siRNA strand thermodynamics was proposed to occur through Ago2 MID domain interactions with the phosphate backbone of the first four nucleotides of the siRNA guide strand (Suzuki et al., 2015), sensing the accessibility and single-strand character of weakly base paired termini (Suzuki et al., 2015). Collectively, our findings agree with this model but suggest that the first and second nucleotide of the siRNA guide strand are the most significant in determining total strand loading (Figure 2.9). siRNA design algorithms should thus account for relative siRNA features (TN Rank and $\Delta\Delta G$) to ensure correct strand selection and absolute strand features (TN and ΔG) to ensure high strand loading and RISC specific activity.

While our experiments and analysis could not cover the entire siRNA feature space, the number of siRNAs tested was large enough to identify regions of the siRNA that explain some of the differences between activity and loading. $5'$ C and $5'$ G sequences are underrepresented in our individual siRNA strands; thus, a more comprehensive set of siRNAs could potentially have provided more insight into how features beyond the $5'$

TN influence siRNA behavior both pre- and post- loading. Nonetheless, the parameters identified indicate that it is possible to design siRNAs with the goal of high loading, high activity, or both.

2.5 Conclusion

Taken together, our findings suggest that siRNA loading and activity are partially independent, allowing for an additional degree of freedom in siRNA selection and design. The behavior of siRNA strands is strongly dependent on the $5'$ nucleotide, but the duplex hybridization energies have a second order effect on both siRNA loading and RISC specific activity. Specifically, $\Delta G_{(-2)-1nt}$ and ΔG_{1-2nt} are predictive of siRNA loading, and ΔG_{3-4nt} and $\Delta G_{17-18nt}$ are predictive of RISC specific activity. Lastly, we have shown that siRNA functional asymmetry is a result of many competing factors that ultimately control the activity of each siRNA strand. A more accurate understanding of the interplay among these factors will lead to better siRNAs.

Chapter 3

siRNA Asymmetry Sensing of PACT

3.1 Abstract

RNA interference (RNAi) is a post-transcriptional regulatory mechanism found in many eukaryotes that utilizes a small RNA, such as a miRNA, to direct silencing towards a complementary mRNA target. siRNAs are designed to enter the RNAi pathway immediately before the formation of the active RNA-induced silencing complex (RISC). The transition from siRNA to an active siRNA-programmed RISC requires the selection of one strand of the duplex and the removal of the other strand, allowing the selected strand to base pair with its target. The proteins involved in siRNA loading are not fully characterized, specifically, in regards to how proteins of the RNAi pathway contribute to the preferential selection of one siRNA strand. Previously, we investigated TRBP for its ability to act as an siRNA asymmetry sensor. Here we investigated PACT, a homologue of TRBP, for its ability to asymmetrically localize to the siRNA termini, a behavior indicative of siRNA asymmetry sensing. We found that like TRBP, PACT, also asymmetrically localizes to siRNA termini but that its pattern of localization to siRNA termini differed from TRBP for the set of siRNAs studied. PACT was found to localize to the siRNA terminus with greater hybridization stability, as defined by the $\Delta\Delta G_{4nn}$ model for siRNA asymmetry.

3.2 Introduction

Short interfering RNAs (siRNAs) provide a method of post-transcriptional, targeted gene knockdown that is of interest as a therapeutic strategy for the treatment of a variety of diseases (Bobbin and Rossi, 2016). Gene knockdown or silencing occurs through incorporation of an siRNA into the native eukaryotic RNA interference (RNAi) pathway, enabling the targeting and degradation of a mRNA sharing complementary to the siRNA sequence (Elbashir et al., 2001a; Zamore et al., 2000). Before the mature complex is formed, one of the strands of the siRNA duplex must be removed (Matranga et al., 2005). Selection of the desired strand is critical to ensure proper transcript targeting (Schwarz et al., 2003). The initial process of strand selection occurs through the asymmetric binding of the siRNA duplex by the proteins of the RNAi pathway, which can recognize specific features of the siRNA (Frank et al., 2010; Noland et al., 2011; Noland and Doudna, 2013; Gredell et al., 2010; Suzuki et al., 2015).

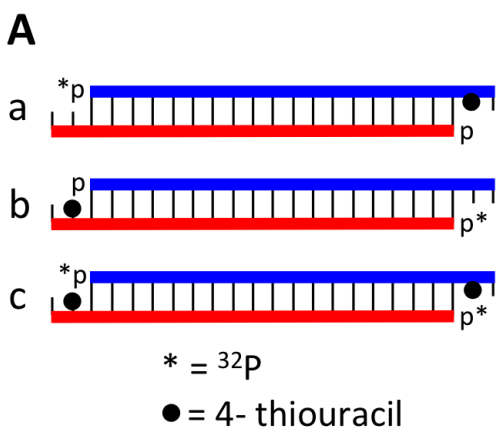
Loading of siRNAs into one of the Ago proteins was originally proposed to occur by positioning the siRNA into Dicer's helicase domain within the Dicer-TRBP complex (Wang et al., 2009; Ma et al., 2008; Noland et al., 2011), followed by the hand off of the mature siRNA to an Ago protein (Matranga et al., 2005; MacRae et al., 2008). siRNAs are then sorted to different Ago proteins based on different structural features (Dueck et al., 2012; Burroughs et al., 2011; Czech and Hannon, 2010), with the majority of siRNAs ending up in either Ago1 or Ago2 (Gu et al., 2011). Ago2 is the only Ago with an active nuclease domain capable of cleaving an mRNA target (Meister et al., 2004; Yoda et al., 2010), making it the only Ago protein that can efficiently silence with a small RNA that is highly complementary to its target, such as an siRNA.

Evidence has also suggested that siRNA loading into Ago2 can occur without the Dicer complex (Kok et al., 2007; Betancur and Tomari, 2012; Kim et al., 2014), suggesting a different mechanism of siRNA loading. The mechanism of siRNA loading remains an area of debate as different experimental models (i.e., loss of function models using knockout cell lines (Betancur and Tomari, 2012; Kim et al., 2014) or siRNAs (Takahashi et al., 2014; Kini and Walton, 2009; Chendrimada et al., 2005) and recombinant protein reconstitution models (Ye et al., 2011; Willkomm et al., 2016; Bernard et al., 2015; Iwasaki et al., 2010; Macrae et al., 2007)) produce conflicting results. To date, C3PO, TRBP, PACT, Dicer, and HSP90 have all been shown to individually constitute RISC activity with Ago2 *in vitro* (Ye et al., 2011; Willkomm et al., 2016; Bernard et al., 2015; Iwasaki et al., 2010). What remains unclear is how, and under what circumstances, each of these proteins may modulate siRNA function within the complex cellular environment. Of the proteins of the RNAi pathway, Dicer, TRBP, and Ago2 have been shown to asymmetrically interact with the termini of siRNA duplexes indicating that these proteins may influence siRNA (or miRNA) asymmetry (Noland et al., 2011; Gredell et al., 2010; Kini and Walton, 2009; Frank et al., 2010; Suzuki et al., 2015).

PACT is a protein that shares 44% homology with TRBP (Chendrimada et al., 2005). PACT, like TRBP, contains 2 dsRBDs and the dsRBD-like Medipal domain, through which, it interacts with Dicer, in a similar manner to TRBP (Haase et al., 2005; Laraki et al., 2008). TRBP has been shown to sense siRNA asymmetry by preferentially localizing to the more stably base-paired siRNA termini (Gredell et al., 2010). Based on the similarities between PACT and TRBP, we investigated the ability of PACT to sense siRNA asymmetry by assessing its ability to differentialy localize to siRNA termini. We were able to demonstrate

that PACT shows differential terminus localization, similar to TRBP.

In general, double-stranded RNA binding proteins (dsRBPs) do not sense differences in RNA duplex structure through the their double-stranded RNA binding domains. We have hypothesized that PACT, similar to TRBP, is capable of sensing differences in RNA structure allowing it to differentially localize to more stably base paired siRNA termini. Our method for determining terminus localization is with a 4-thiouracil photoreactive cross-linker incorporated at the 20^{th} position of the siRNA (Figure 3.1A). Given that the proposed method of RNA recognition is non-canonical, we investigated siRNA terminus localization with another generic dsRBP, NS3, to confirm that cross-linking patterns were a result of specific protein-RNA interactions and not a result of differences cross-linker efficiency. The criteria for the selection of a negative control dsRBP included the ability to bind to both dsRNA and ssRNA. NS3, a viral suppressor of RNAi from the *Rice stripe virus*, specifically binds to siRNA duplexes with high affinity and is also known to bind to ssRNA, although, with lower affinity (Shen et al., 2010). Our results demonstrate the efficacy of the photoreactive 4-thiouracil cross-linker by showing that all of the siRNAs in this study cross-linked with equal efficiency to the *Rice stripe virus* NS3 protein.



B

```
    Asymmetric siRNA (pp-luc)
AS 5′-  UCGAAGUAUUCCGCGUACGUG-3′
SS 3′-AUAGCUUCAUAAGGCGCAUGC  -5′

      Symmetric siRNA (sod1)
AS 5′-  GUCACAUUGCCCAAGUCUCUU-3′
SS 3′-UUCAGUGUAACGGGUUCAGAG  -5′

    Unknown siRNA (EGFP#274)
AS 5′-  UGCGCUCCUGGACGUAGCCUU-3′
SS 3′-UUACGCGAGGACCUGCAUCGG  -5′
```

Figure 3.1: Configuration of siRNAs for Cross-linking Experiments. (A) Placement of photoreactive cross-linkers (4-thiouracil) and radiolabels (^{32}P) to determine dsRBP terminus localization (configurations a and b). Configuration c should be the average of the cross-linking observed for configurations a and b. (B) siRNAs used to determine dsRBP terminal localization. pp-luc and sod1 siRNAs are highly functionally asymmetric and symmetric, respectively. The functional asymmetry of EGFP#274 is unknown.

3.3 Results

3.3.1 Functional Characterization of Recombinantly Expressed dsRBPs

We expressed His-tagged NS3 from the *Rice stripe virus* in Rosetta 2 (DE3) cells, purified it to homogeneity (assessed by SDS-PAGE gel, Figure 3.2A), and verified immunoreactivity with an anti-6x His antibody (Figure 3.2B). The apparent size of NS3 by both SDS-PAGE (Figure 3.2A) and western blotting (Figure 3.2B) was near the theoretical size of 24.6 kDa. Functionality of NS3 was confirmed by assessing its ability to bind ssRNA and siRNA (Figures 3.2C and 3.2D) via electrophoretic mobility shift assay (EMSA) with increasing quantities of NS3. NS3 did bind to siRNAs with high affinity (low nanomolar range),

however we observed a saturation in the quantity of complex formed with ∼60% binding of labeled siRNA (Figure 3.2D). NS3-siRNA complex formation saturated with increasing protein concentration suggesting that either, complexes are disturbed during electrophoresis or that the siRNA is in a conformation that discourages NS3 binding. In either case, the quantity of NS3-siRNA complexes formed, provided sufficient signal to perform cross-linking experiments. NS3-ssRNA complexes were not well resolved by EMSA with $< 1\%$ of ssRNA forming complexes with 1 μM of NS3 (Figure 3.2C).

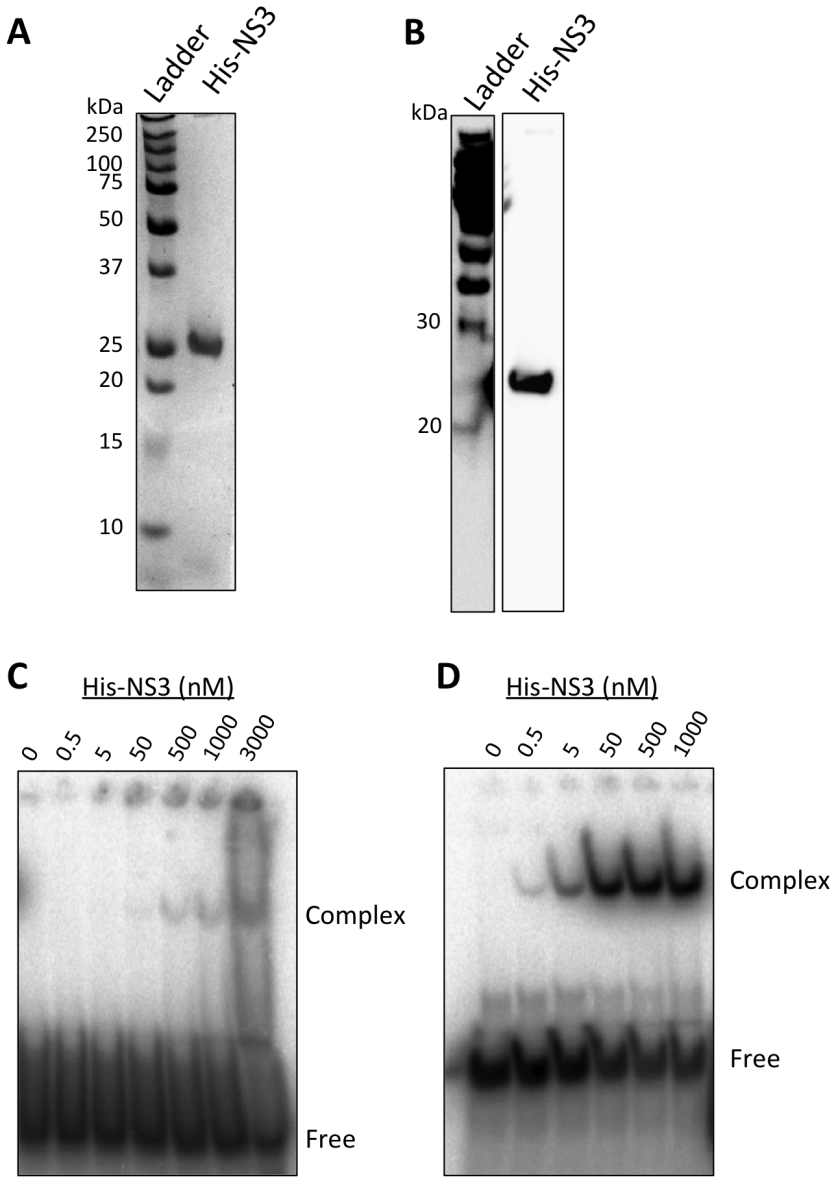


Figure 3.2: Functional Characterization of NS3. (A) SDS-PAGE analysis of recombinant His-NS3 stained with GelCode Blue total protein stain, theoretical size: 24.57 kDa. (B) Western blot of recombinant NS3 with anti-6x His antibody. (C) EMSA of NS3 binding to 5′-^{32}P labeled pp-luc AS strand resolved with native PAGE. (D) EMSA of NS3 with 5′-^{32}P labeled pp-luc siRNA resolved with native PAGE.

To study the asymmetric binding preferences of PACT with siRNAs, we recombinantly expressed PACT with the maltose binding protein (MBP) expression tag in *E. Coli* BL21(DE3)-RIPL cells. The MBP-tag was used to increase the solubility of PACT and has

been found not to interact with siRNAs in *in vitro* binding assays (Gredell et al., 2010). Recombinant protein was purified to ∼95% purity (assessed by SDS-PAGE gel, Figure 3.3A). Protein identity was confirmed with an anti-MBP antibody and had an observed size near the theoretical protein size of 79.8 kDa (Figure 3.3B). Functionality of MBP-PACT was confirmed by verifying its ability to bind ssRNA and siRNA via EMSA (Figure 3.3C and 3.3D, respectively). MBP-PACT affinity for siRNAs was characterized with increasing protein concentrations and a limiting siRNA concentration (∼1 nM siRNA) (Figure 3.3D). siRNA binding saturated at ∼50% with 1 μM MBP-PACT (Figure 3.3D). As with NS3, we suspect that lack of a complete shift of labeled siRNA develops from complex instability during electrophoresis or that siRNA is in a conformation that precludes binding by PACT. siRNA cross-linking signals with PACT were strong enough to differentiate siRNA terminus localization.

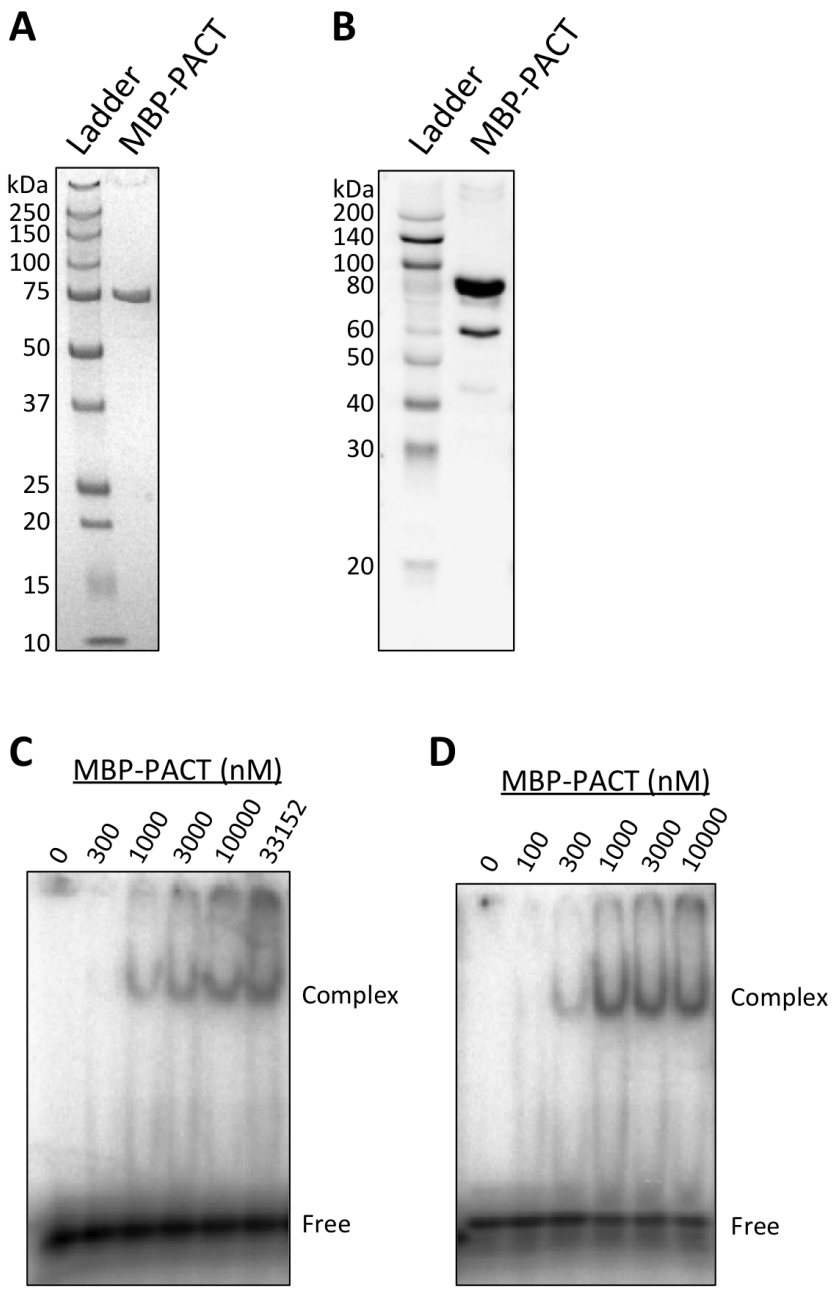


Figure 3.3: Functional Characterization of MBP-PACT. (A) SDS-PAGE analysis of recombinant MBP-PACT stained with GelCode Blue total protein stain, theoretical size: 79.8 kDa. (B) Western blot of recombinant MBP-PACT with anti-MBP antibody. (C) EMSA of MBP-PACT binding to $5'$-^{32}P labeled pp-luc AS strand resolved with native PAGE. (D) EMSA of MBP-PACT with $5'$-^{32}P labeled pp-luc siRNA resolved with native PAGE.

3.3.2 NS3 Does Not Asymmetric Bind of siRNAs

To control for the possibility that the measured asymmetric binding of siRNAs by dsRBPs results from a bias in the cross-linking efficiency of the siRNA strands, siRNA binding and cross-linking experiments were performed with NS3. Reactions were performed with 50 nM of protein, at which concentration approximately 50% of the siRNA is complexed with NS3 (Figures 3.4A and 3.4B). The fraction of cross-linked NS3 protein with siRNA configurations (a) and (b), was not significantly different for the asymmetric ($p = 0.61$), symmetric ($p = 0.70$), or Unknown ($p = 0.72$) siRNAs. These cross-linking patterns are different from those published with the TRBP protein, which showed a significant difference in the cross-linking pattern between the (a) and (b) configurations for the asymmetric and unknown siRNAs ($p < 1\text{x}10^{-4}$) (Gredell et al., 2010).

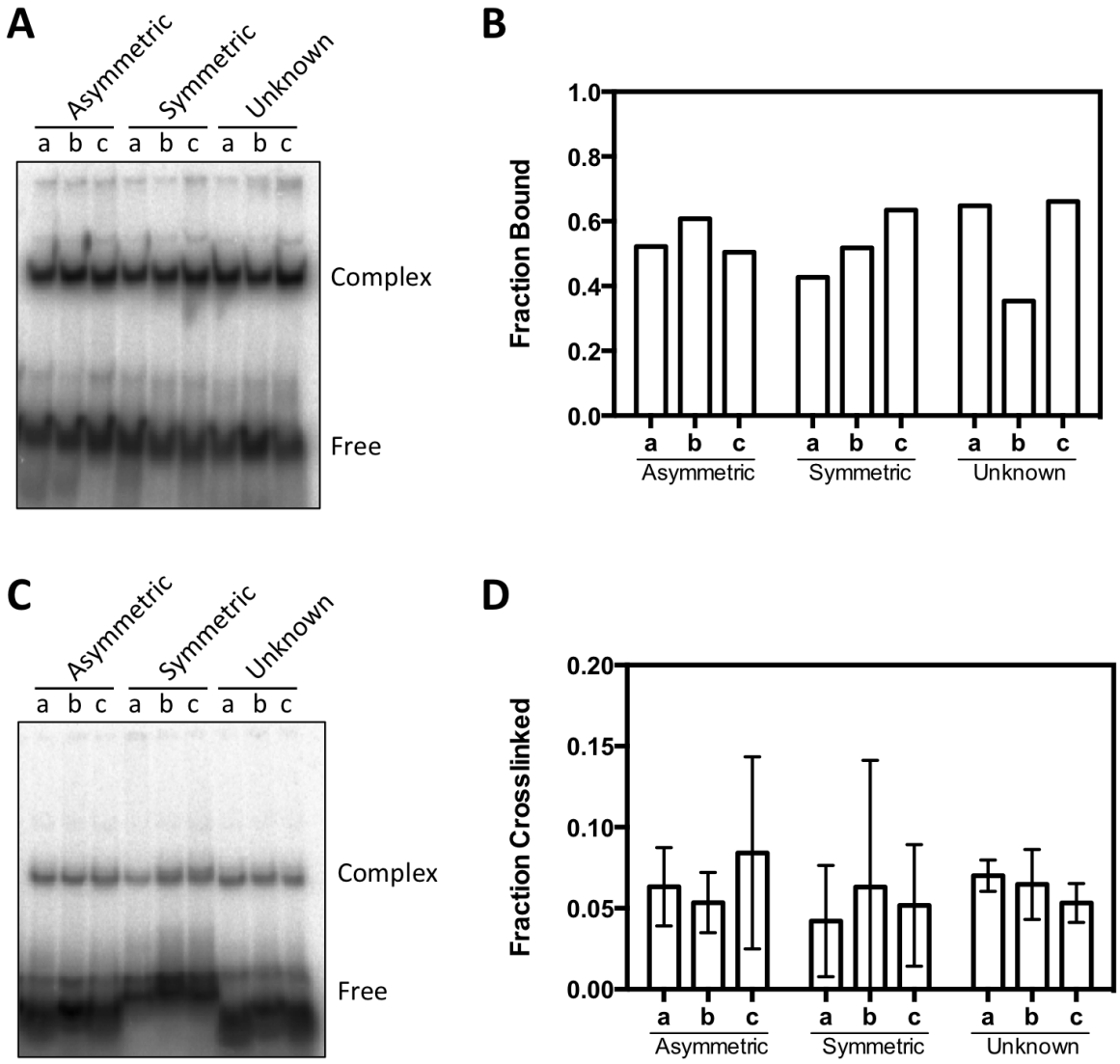


Figure 3.4: Symmetrical Cross-linking of NS3 to siRNA. (A) Native gel shift analysis and (B) quantification of NS3 binding to siRNAs with each cross-linker configuration in Figure 3.1A. (C) Denaturing gel electrophoresis of cross-linked siRNAs with NS3 and (D) Quantification of the NS3 siRNA cross-linking. Fraction cross-linked are mean ± 1 SD; N=3. Comparison of a and b configurations for each siRNA by one-way ANOVA with Tukey's *post hoc* analysis showed no significant differences.

NS3 binding to ssRNA was relatively low for the concentrations of protein tested (Figure 3.5A), however we were able to demonstrate > 10% cross-linking to all siRNA strands with 1 μM NS3 (Figure 3.5C). Similar disruption of TRBP-ssRNA complexes during native PAGE was observed under identical conditions (Gredell et al., 2010). Native binding of

NS3 to ssRNA varied with the different RNA sequences; however, the quantity of ssRNA that cross-linked to NS3 was not significantly different between the AS and SS for the sequences tested (asymmetric $p = 0.31$, symmetric $p = 0.28$, and EGFP#274 $p = 0.087$, two-tailed t-test). Furthermore, cross-linking was similar among all strands, except for the unknown SS sequence which was cross-linked to NS3 modestly lower than the other ssRNAs (Figure 3.5C and 3.5D). The ssRNA cross-linking patterns with NS3 vary substantially from the cross-linking patterns observed with MBP-TRBP (Figure 3.5D), which had significant differences in cross-linking between multiple ssRNA (Gredell et al., 2010). Collectively both the ssRNA and siRNA cross-linking results support the hypothesis that cross-linking efficiency is similar for the sequences tested, and that any differences in cross-linking are due to specific localization of the protein on the siRNA structure.

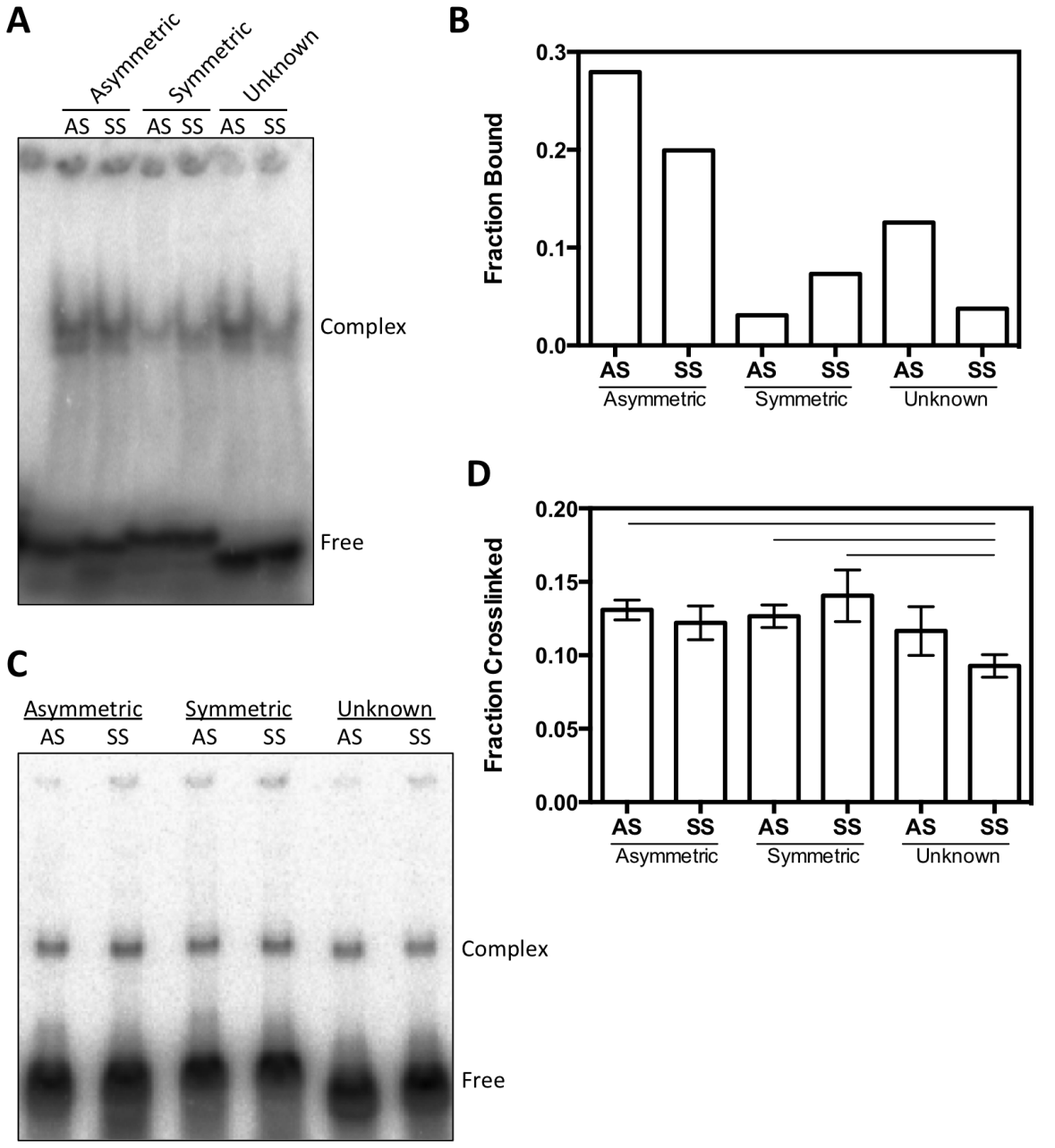


Figure 3.5: Symmetric Cross-linking of NS3 to ssRNA. (A) Native gel shift analysis and (B) quantification of His-NS3 binding to each ssRNA in Figure 3.1B. (C) Denaturing gel electrophoresis of cross-linked siRNAs with His-NS3 and (D) Quantification of the His-NS3 siRNA cross-linking to ssRNA, each with a 5′-^{32}P label and 4-thiouracil photoreactive cross-linker subsitituted at the 20^{th} nucleotide. Fraction cross-linked are mean ± 1 SD; N=3. Differences in fraction cross-linked for each ssRNA were by tested by one-way ANOVA with Tukey's *post hoc* analysis. Overhead bars represent values which are significantly different ($p < 0.05$).

3.3.3 PACT Asymmetrically Localizes to siRNA Termini

Based on the high degree of homology between TRBP and PACT (Chendrimada et al., 2005), we hypothesized that PACT would possess similar asymmetric terminal localization preferences despite its lower affinity for siRNAs. MBP-PACT binding and cross-linking experiments were performed at a concentration of 3 μM, where fractional siRNA binding was 40% (Figure 3.6A and 3.6B). MBP-PACT cross-linked preferentially to the more stable terminus of the functionally asymmetric siRNA pp-luc ($p = 0.065$, two-tailed t-test), and cross-linked equally to both termini of the functionally asymmetric siRNAs ($p = 0.66$, two-tailed t-test) and siRNA with unknown functional asymmetry (EGFP#274; $p = 0.53$, two-tailed t-test) (Figures 3.6C and 3.6D, Table 3.1). MBP-PACT cross-linking patterns were similar to TRBP and R2D2, the *Drosophila* homolog of TRBP, for the functionally asymmetric and symmetric siRNAs (Table 3.1). However, the symmetric cross-linking pattern observed for EGFP#274 with MBP-PACT differed from the asymmetric cross-linking pattern previously observed with MBP-TRBP (Gredell et al., 2010).

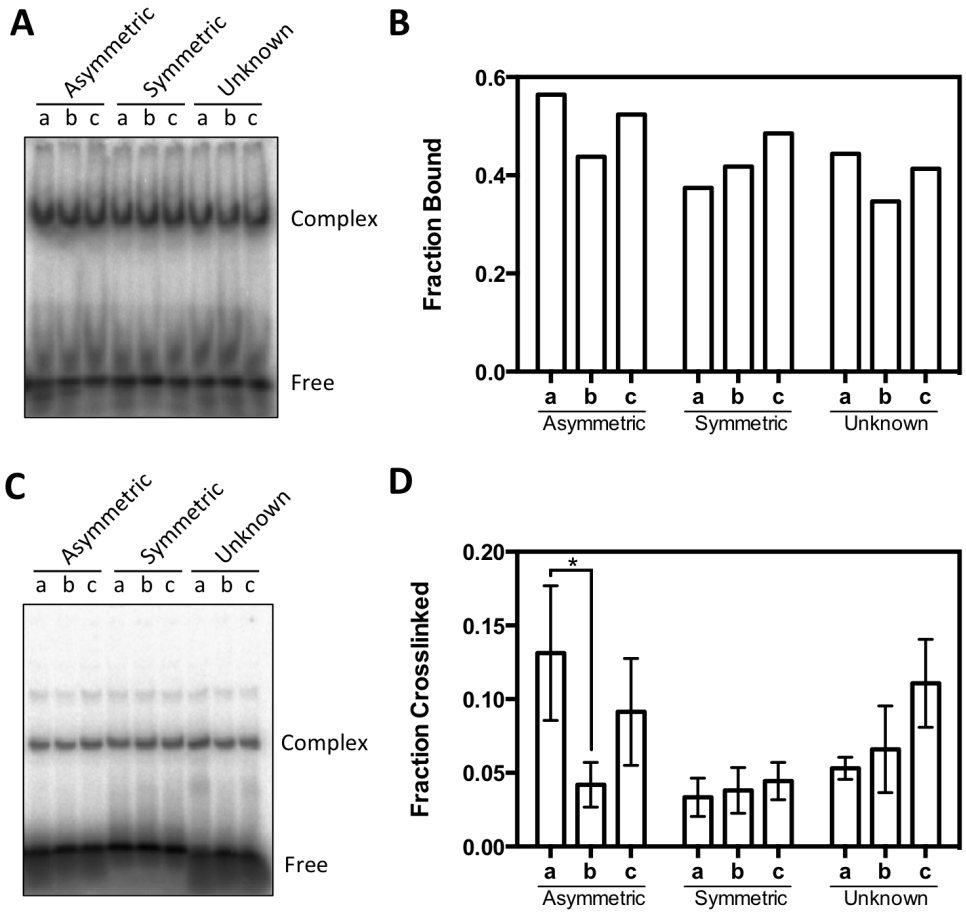


Figure 3.6: Asymmetrical Binding of MBP-PACT to siRNA. (A) Native gel shift analysis and (B) quantification of MBP-PACT binding to siRNAs with each cross-linker configuration in Figure 3.1A. (C) Denaturing gel electrophoresis of cross-linked siRNAs with MBP-PACT and (D) Quantification of the His-NS3 siRNA cross-linking. Fraction cross-linked are mean $\pm$ 1 SD; N=3. Comparison of a and b configurations for each siRNA by one-way ANOVA with Tukey's *post hoc*, (*) represents a $p = 0.065$, no other comparisons of a and b configurations are statistically significant.

Table 3.1: PACT Asymmetric siRNA Binding Differs From TRBP. Asymmetric cross-linking patterns for dsRBPs, R2D2, Drosophila homolog of TRBP, (Tomari et al., 2004), TRBP (Gredell et al., 2010), and PACT. (+) Represents higher cross-linking with the siRNA antisense strand, neutral means that cross-linking between the siRNA termini was the same, and ND is not determined.

	R2D2	TRBP	PACT
Asymmetric	+	+	+
Symmetric	Neutral	Neutral	Neutral
EGFP#274	ND	+	Neutral

3.3.4 PACT Dimerizes with ssRNA and siRNA

TRBP-siRNA binding reactions produced two complexes on a native gel (Figure 3.7A, lane 2). The first complex has a 1:1 stoichiometric ratio of siRNA:MBP-TRBP and the second, a 1:2 stoichiometric ratio (Gredell et al., 2010). Interestingly, PACT only forms one complex with both siRNA and ssRNA, and runs at a similar size to the dimeric MBP-TRBP (Figure 3.7A, lanes 3 and 4). The complex formed does not change with increasing protein concentration (Figures 3.3C and 3.3D), which is in agreement with the findings that PACT natively exists as a dimer (Takahashi et al., 2013).

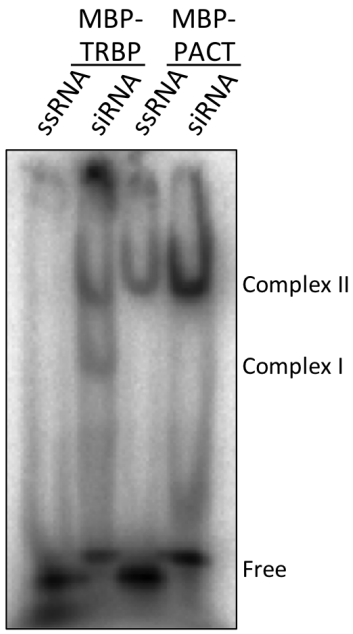


Figure 3.7: Different Binding Modes of TRBP and PACT. (A) EMSA of MBP-TRBP (Lanes 1 and 2) and MBP-PACT (Lanes 3 and 4) with 5′-^{32}P labeled pp-luc AS ssRNA (Lanes 1 and 3) and 5′-^{32}P labeled pp-luc siRNA (Lanes 2 and 4). Protein concentrations for each binding reaction are as follows, Lane 1: 1250 nM MBP-TRBP, Lane 2: 350 nM MBP-TRBP, Lane 3: 3000 nM, Lane 4: 10000 nM.

3.4 Discussion

Small RNA asymmetry develops from the binding preferences of the RNAi proteins for specific features of the small RNAs. The localization or orientation of the RNAi proteins when interacting with small RNAs can then drive small RNA strand selection. The cumulative effect of asymmetric interactions leads to the asymmetric loading of one small RNA strand and exclusion of the other. We were able to demonstrate that PACT differentially localizes to siRNA termini with some siRNAs but not others, similar to but still unique from the function of TRBP (Figure 3.6) (Gredell et al., 2010). Furthermore, we found that NS3, a protein with no known preferences for specific dsRNA structures or nucleotides, did not

asymmetrically cross-link to siRNA termini. Thus, differences in cross-linking between the different termini of the siRNA with PACT and TRBP are a function of the protein-RNA interaction and are not a result of differences in cross-linking efficiency.

The asymmetric localization of both TRBP and PACT with functionally asymmetric duplexes indicates that these proteins are capable of directing the asymmetric binding of siRNA duplexes and influencing siRNA asymmetric strand loading. While not explored here, it is possible that the difference in oligomeric state between TRBP and PACT, impacted how each protein interacts with siRNA termini (Figure 3.7). The full roles of TRBP and PACT in RNAi are not understood and it is possible that either TRBP and PACT influence siRNA loading through their involvement with the Dicer complex (Lee et al., 2013), through direct interaction with Ago2 (Willkomm et al., 2016), or not at all (Kim et al., 2014; Wilson et al., 2015). Nonetheless, our results elucidate differences in TRBP and PACT binding preferences, demonstrate the possibility of PACT to act as an siRNA asymmetry sensor, and informs how PACT may interact with other dsRNA structures.

Two models have been proposed to predict the localization of the R2D2 and TRBP to siRNA termini, $\Delta\Delta G_{4nn}$ (Tomari et al., 2004) and TN Rank (Gredell et al., 2010). TRBP, R2D2, a *Drosophila* of TRBP, and PACT binding patterns were all asymmetric with pp-luc and symmetric with sod1 siRNAs (Table 3.1); consistent with both models for siRNA terminus localization and the known functional asymmetry of these siRNAs (Table 3.2). The asymmetric binding patterns were not similar between TRBP and PACT with EGFP#274 (Table 3.1), nor were the predicted binding pattern consistent between the TN Rank and $\Delta\Delta G_{4nn}$ models (Table 3.2). The $\Delta\Delta G_{4nn}$ of EGFP#274 is -0.1 kcal/mol, a relatively small difference in hybridization stability. If the $\Delta\Delta G_{4nn}$ of EGFP#274 can be

interpreted as symmetrical, then the $\Delta\Delta G_{4nn}$ model predicted the asymmetric localization of PACT for all three siRNAs, making $\Delta\Delta G_{4nn}$ a better model for the asymmetric binding pattern of PACT, compared to TN Rank. However, the validation of these models is very limited and many more sequences are needed before these models could be used with any utility in predicting terminus localization of TRBP or PACT with siRNAs (See Chapter 4, Future Directions: TRBP and PACT).

Table 3.2: Predicited Asymmetric Localization of R2D2, TRBP, and PACT. Prediction of asymmetric binding pattern on the Asymmetric, Symmetric, and Unknown siRNAs using the TN Rank and $\Delta\Delta G_{4nn}$ models. (+) Represents higher cross-linking with the siRNA antisense strand, neutral means that cross-linking between the siRNA termini was the same, and ND is not determined.

	$\Delta\Delta G_{4nn}$(kcal/mol)	TN Rank
Asymmetric	0.5	+
Symmetric	0	Neutral
EGFP#274	-0.1	+

3.5 Conclusions

Here we demonstrated that the protein PACT, an RNAi pathway protein hypothesized to aid in the loading of small RNAs, is capable of asymmetrically localizing to siRNA termini. PACT binds to siRNAs but at a much lower affinity than its homolog, TRBP. PACT also differs from TRBP in its asymmetrical binding pattern with one of the siRNAs tested, and unlike TRBP, aligned well with a $\Delta\Delta G_{4nn}$ model to predict termini localization. These results indicate that PACT is capable of sensing characteristics of the siRNA duplex, suggesting that PACT can serve as an asymmetry sensor within the RNAi pathway.

Chapter 4

Conclusions and Future Directions

4.1 Conclusions

The purpose of these studies was to better understand the siRNA characteristics and protein-RNA interactions responsible for making one siRNA strand of a duplex more active than the other. The results from this work provide both details of the RNAi mechanism and informs siRNA design.

In order to better understand the relative importance of two siRNA duplex features that are predictive of siRNA activity, TN Rank and $\Delta\Delta G_{3nn}$, we investigated how each characteristic predicts siRNA strand loading and activity with set of siRNAs that sampled both features. We found that TN Rank alone was able to predict siRNA functional asymmetry, whereas $\Delta\Delta G_{3nn}$ predicts siRNA functional asymmetry only when applied as a second order effect to TN Rank. A similar trend was observed with asymmetric strand loading, but the predictive ability of $\Delta\Delta G_{3nn}$ was much weaker compared to functional asymmetry. Upon further investigation, we found that individual siRNA strand loading was better predicted by the $5'$ terminal duplex energies ($\Delta G_{(-2)-1nt}$ and ΔG_{1-2nt}). TN Rank was predictive of siRNA activity because of its collinearity with the $5'$ terminal duplex energies. Post-siRNA loading, low hybridization stability in the primary seed region was found to correlate with greater RISC specific activity, possibly indicating that a weak

target interaction to encourages RISC turnover. Greater hybridization stability near the $3'$ termini ($\Delta G_{17-18nt}$) was also found to correlate with RISC specific activity. Because this region of the guide strand has been shown, with recombinant mouse Ago2, to be part of the $3'$ supplement (Wee et al., 2013), a region of the guide strand not involved in RISC-target interactions, our results suggest that the cellular environment may alter the RISC interactions at the $3'$ end of the siRNA. These results refine our understanding of how the high level parameters, TN Rank and $\Delta\Delta G$, impact siRNA function at different stages in the RNAi pathway.

The molecular mechanisms underlying siRNA asymmetric strand selection are not fully known. Here we characterized PACT, a homolog to the protein, TRBP, and demonstrated that PACT, like TRBP, asymmetrically interacts with siRNA duplexes although not in an identical manner to TRBP. We propose that the difference in duplex localization is because PACT natively exists as a dimer, whereas TRBP exists natively as a monomer, altering the way in which these proteins interact with siRNAs but still retaining structural preferences for more stably base paired termini. Asymmetric localization was shown using a photoreactive 4-thiouracil cross-linker substituted in the $3'$ overhang of siRNA duplex; the mechanism of cross-linking was rigorously validated to confirm unbiased cross-linking. These results provide general insight into how these proteins differentially interact with RNA sequences.

4.2 Future Directions

As we begin to understand more about the RNAi mechanism, siRNA activity predictions have improved. Outlined below are future directions of the work presented here.

4.2.1 Parsing with Larger Datasets

Our analysis in Chapter 2 investigated how the TN and ΔG parameters influence siRNA strand selection and activity using 2-factor linear regression modeling. A 2-factor model was necessary to control for the effect of TN. We were limited in the number of variables that we could simultaneously analyze by the size of our dataset and the coverage of different siRNA features. As a result we were unable to look at interactions between TN and $\Delta G_{(x)-(y)nt)}$. In particular, we were curious if the TN could influence the significance of a correlation between ΔG and activity at a given location in the siRNA duplex.

We performed a preliminary analysis using a dataset containing the activity data of 2431 siRNAs (Huesken et al., 2005), hereafter referred to as the Novartis dataset. We began our analysis utilizing the same 1-factor (TN or $\Delta G_{(x)-(y)nt)}$) and 2-factor (TN or $\Delta G_{(x)-(y)nt)}$) linear regression models (Figure 4.1A) used in Chapter 2 to compare our results to those of the Novartis dataset. Because of the size of the dataset, many of the ΔGs within the siRNA duplex were found to be significant in both the 1-factor and 2-factor models (Figure 4.1A). To provide a comparison between our dataset and the Novartis dataset, we decided to only evaluate the top two parameters, because we were only capable of identifying two parameters in our activity dataset. The top two most significant parameters identified using 1-factor model (TN and ΔG_{1-2nt}) matched the parameters identified in the 1-factor activity model from our dataset (Figure 2.9). However, the results from the Novartis 2-factor model varied from our 2-factor model (Figure 2.9), with the Novartis 2-factor model retaining a strong correlation between siRNA activity and $5'$ terminus duplex energies ($\Delta G_{(-2)-(1)nt}$ and ΔG_{1-2nt}). These discrepancies likely develop from the design of each dataset, with our dataset having an unequal representation of $5'$ TNs, potentially biasing the importance

of a specific features, and the Novartis dataset using a single high concentration of siRNA, limiting the dynamic range of siRNA activities.

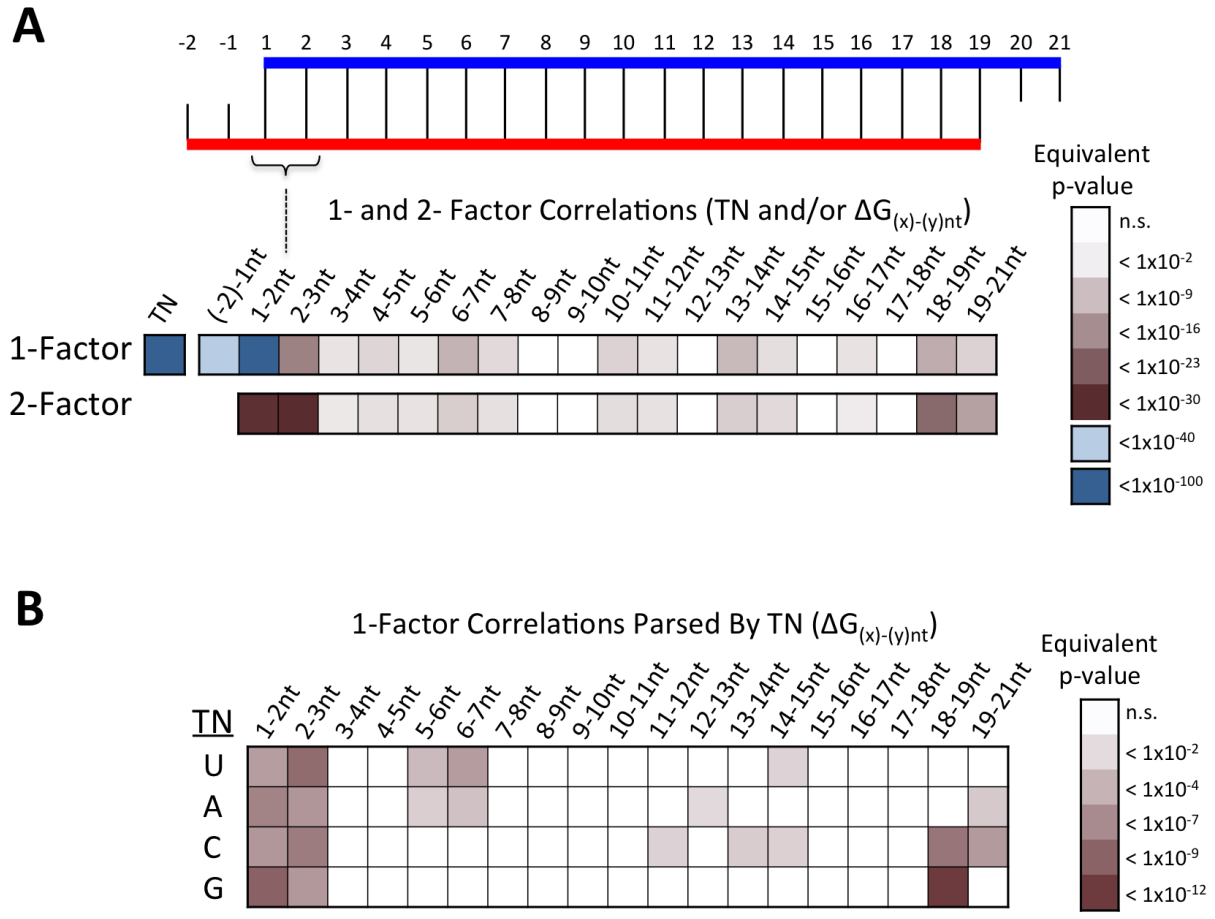


Figure 4.1: TN can Alter the Role of ΔG in Predicting siRNA Activity. (A) 1-factor and 2-factor correlations of siRNA activity data from the Novartis dataset with TN and $\Delta G_{(x)-(y)nt}$ along the length of the duplex. 2-factor correlations use TN as one variable and a $\Delta G_{(x)-(y)nt}$ value as the other. Significance in these correlations indicates that the variable predicts activity independent of TN. A Bonferroni correction was applied to correct for multiple comparisons.

To look for interactions between TN and $\Delta G_{(x)-(y)nt}$, we parsed the Novartis dataset by TN (Figure 4.1B). After parsing the data by TN, we again looked for correlations between activity and $\Delta G_{(x)-(y)nt}$, using a 1-factor linear regression model. We observed a preference for low hybridization stability at the $5'$ terminus (ΔG_{1-2} and ΔG_{2-3}) for all TN.

The $\Delta G_{(-2)-1}$ parameter was not unique after parsing the data by TN and could not be incorporated into the linear model. We also observed a preference for greater hybridization stability at the $3'$ terminus of siRNAs with $5'$ adenine, cytidine, and guanine TN, but not uridine nucleotides. The correlation for siRNA sequences with $5'$ adenine TN was much weaker than the correlations at the $3'$ termini for siRNA sequences with a $5'$ cytidine or guanine TN. These results suggest a differential dependence on $3'$ terminal duplex energies. A preference for weaker hybridization stability at nucleotides 5-6 and 6-7 was also observed for siRNA sequences with uridine and adenine TN. The significant ΔGs in the Novartis dataset do not perfectly align with the results presented in Chapter 2, however, because of differences in dataset design it is difficult to make direct comparisons. These results do, however, indicate that the TN can alter how duplex stability effects siRNA activity. Greater consideration should be given to understanding the interactions among siRNA activity predictors. Future work in this area should explore the conditional nature of the $3'$ duplex energies with different TN, looking at why this characteristic is important.

4.2.2 TRBP and PACT

Both the work presented in Chapter 3 and previous work suggest that TRBP and PACT have preferences for certain features of the siRNA duplex (Gredell et al., 2010), allowing them to differentially localize to siRNA termini. Future work can be split into two areas, i) understanding the RNA structural preferences of TRBP and PACT and ii) investigating TRBP and PACT function in siRNA loading using an tissue culture model.

In our prior work, only three sequences were used to evaluate the terminus localization of TRBP and PACT. To understand how TRBP and PACT can drive siRNA asymmetric

strand loading, many more sequences will need to be evaluated. The current experimental design is not amenable to screen more than a few siRNA sequences, so in order to map RNA structural preferences of TRBP and PACT a more robust experimental strategy will be required.

Some efforts have been made to map the structural preferences of TRBP through DNA substitutions, which perturb TRBP binding (Takahashi et al., 2014; Gredell et al., 2010), and through the introduction of mismatches within duplex structures (Acevedo et al., 2015). These studies have found that TRBP binding is perturbed by internal bulges and mismatches (Acevedo et al., 2015), and that TRBP most strongly associates with siRNA termini (Takahashi et al., 2014; Kok et al., 2011). While helpful, these studies were not able to look at a sufficiently complex pool of sequences to fully map TRBP or PACT interactions.

TRBP and PACT structural preferences could be mapped with a selection experiment using a pool of unique RNA sequences and evaluating the enrichment of RNA sequences. Deep-sequencing allows the parallel quantification of RNA sequences and would be ideally suited for the application. The minimum length duplex that TRBP can bind to is $\sim$15 nt (Koh et al., 2013); using a complex pool of 15 nt RNA fragments would limit the number of possible motifs TRBP could recognize in an individual sequence. The minimum length duplex that PACT can bind to is not known and would need to be determined before experimentation. Further characterization of the individual domains of TRBP and PACT would help to understand how TRBP and PACT recognize different duplex structures and could also be performed using a similar enrichment assay.

To the second point, many studies of TRBP and PACT, including those performed

in Chapter 3, were performed *in vitro*. No experiment has definitively demonstrated the function of TRBP or PACT in siRNA loading in cells. Likewise, the only evidence indicating the importance of TRBP in siRNA mediated silencing is the observation of a decrease in silencing when TRBP was knocked down with an siRNA (Haase et al., 2005; Chendrimada et al., 2005; Takahashi et al., 2014). More rigorous experiments looking at the functional aspects of TRBP and PACT will be required to understand the role of these proteins in RNAi. The proposed roles of TRBP and PACT as proteins involved in the loading of siRNAs implies that their interactions with siRNAs are transient, therefore, the quantities of siRNA interacting with either of these proteins at a single point in time are low and hard to detect. As a result, functional assays are required to investigate TRBP and PACT in live cells. A HeLa cell line with both TRBP and PACT knocked out has recently been developed (Kim et al., 2014). Using this cell line, siRNA functional asymmetry and siRNA loading could be investigated with and without rescuing of TRBP or PACT protein expression using the methods developed in Chapter 2. Further, the role of the $5'$ ΔG, as was shown in Chapter 2 to predict siRNA strand loading, can be investigated.

APPENDICES

Appendix A

Materials and Methods for Ch. 2

Cell Culture and Transfection

HeLa cells were maintained in DMEM High Glucose (Life Technologies cat# 11965-092) supplemented with 10% FBS (Life Technologies cat# 16000-044) and 1% Pen-Strep (Life Technologies cat# 15240062) and incubated at 37°C and 5% CO_2 in a humidified incubator. Before transfection, cells were plated at 15,000 cells/well in 96-well plates or 350,000 cells/well in 6-well plates in media without Pen-Strep for 24 h. Forward transfections were performed with Lipofectamine 2000 (Life Technologies cat#11668019) at a final concentration of 2.33 μg/mL, based on final well volume. Lipoplexes were prepared per the manufacturer's instructions in Opti-MEM (Life Technologies cat# 31985-070); siRNA and plasmid concentrations are specified for each experiment.

Nucleic Acids

Protein Kinase R (PKR)-targeting siRNAs were ordered from Dharmacon as duplexes, designed with a 19 bp antisense strand targeting the PKR gene and $3'$ uridine dinucleotide overhangs. siRNA sequences are listed in Table A.1. siRNA nearest-neighbor energy parameters were obtained from the DINAMelt Web Server (Markham and Zuker, 2005, 2008)

and used to calculate relative terminal hybridization stabilities (Figure 1.5A). All DNA primers for this study were ordered from Integrated DNA Technologies.

Table A.1: PKR-targeting siRNA Sequences. siRNA sequences targeting the PKR gene with their corresponding TN Rank and $\Delta\Delta G_{3nn}$. (+) siRNA strand is in blue and (-) siRNA strand is in red. Algorithm rank is the predicted relative activity compared to all siRNA sequences targeting the PKR gene when taking into consideration their TN Rank and $\Delta\Delta G_{3nn}$ (Malefyt et al., 2013). Some algorithm rankings and $\Delta\Delta G_{3nn}$ values differ from our prior work due to updated nearest-neighbor parameters and refinements to the algorithm. The last sequence is the siRNA used as an internal standard for RT-qPCR experiments.

5' Target Position	siRNA Sequence	TN Rank	$\Delta\Delta G_{3nn}$ (kcal/mol)	Algorithm Rank
410	UAAUGAAAUCCUUCUGGCCUU UUAUUACUUUAGGAAGACCGG	U:G (1)	6.9	1
816	UUUAAAAUCCAUGCCAAACUU UUAAAUUUUAGGUACGGUUUG		1.6	137
952	UUGCCAAUGCUUUUACUUCUU UUAACGGUUACGAAAAUGAAG		-0.6	427
1129	UCCCUUUAUCACAGAAUUCUU UUAGGGAAAUAGUGUCUUAAG		-3.4	930
336	AAUUCUAUUGAUAAGGCCUUU UUUUAAGAUAACUAUUCCGGA	A:A (9)	5.3	335
645	AUCUGCUGAGAAGUCACCUUU UUUAGACGACUCUUCAGUGGA		1.8	974
751	AUCUGGGUGCCAAAGAUCUUU UUUAGACCCACGGUUUCUAGA		0	1192
379	AUGCACACUGUUCAUAAUUUU UUUACGUGUGACAAGUAUUAA		-3.4	1662
440	CUAUAUUCUUUCUGUCCCAUU UUGAUAUAAGAAAGACAGGGU	C:U (16)	3.9	1764
1198	CAAAGAGUUCCAAAGCCAAUU UUGUUUCUCAAGGUUUCGGUU		2	1960
640	CUGAGAAGUCACCUUCAGAUU UUGACUCUUCAGUGGAAGUCU		0	2140
928	CCGCCUUCUCGUUAUUAUAUU UUGGCGGAAGAGCAAUAAUAU		-5.2	2346
Internal Standard	UGUAGUUGCCGUCGUCCUUUU UUACAUCAACGGCAGCAGGAA	---	---	---

For luciferase experiments, the PKR gene was PCR amplified from pET28a-PKR (Bevilacqua and Cech, 1996) (kindly provided by Dr. Philip C. Bevilacqua) and subcloned into the psiCHECK2-AS34a vector (Navarro et al., 2009) (Addgene # 37099) in the forward

(psiCHECK2-PKR+) or reverse (psiCHECK2-PKR-) direction using Clontech In-fusion cloning (Figure 2.2). Plasmids were verified by Sanger sequencing. Primers used for cloning and sequencing are list in Table A.2.

Table A.2: Cloning and Sequencing Oligos. The following primers were used for the cloning of the PKR gene into the psiCHECK2 vector, in both the forward and reverse directions, and for sequencing verification.

Name	5' --> 3' Primer Sequence
(+) PKR FWD	TAGGCGATCGCTCGAGATGGCTGGTGATCTTTCAGCAGGTTTC
(+) PKR RVR	TTGCGGCCAGCGGCCGCCTAACATGTGTGTCGTTCATTTTTCTCTGGG
(-) PKR FWD	TAGGCGATCGCTCGAGCTAACATGTGTGTCGTTCATTTTTCTCTGGG
(-) PKR RVR	TTGCGGCCAGCGGCCGCATGGCTGGTGATCTTTCAGCAGGTTTC
psiCHECK2 FWD Sequencing	CGGGCTTTAAAAAGAAAGATA
psiCHECK2 RVR Sequencing	CCAAATGGTGTTACTTTGAGA

Primers for stem-loop RT-qPCR were designed and prepared according to published specifications (Varkonyi-Gasic and Hellens, 2011), with a constant hairpin region and a 6 nt overlap with the amplification target. All sequences used are listed in Supplementary Table A.3. qPCR was performed with a sequence-specific forward primer and a universal reverse primer. All primers were verified by spiking in known quantities of siRNA and verifying amplification efficiency. Standard curves were generated for each siRNA sequence to account for variation in amplification efficiency.

Table A.3: Stem-loop Primer Sequences. Primer design follows the protocol outlined by Chen C. et al. (Chen et al., 2005), with a standard stem-loop sequence and a 3′ tail with a 6 nt overlap specific to the 3′ end of each siRNA strand. qPCR was performed with a universal reverse primer and a siRNA strand specific forward primer. G/C nucletotides were added to the 5′ end of the forward primer to adjust T_m to ∼60°C.

Stem-loop Primer:

```
  G GTATT C
A             GCACCAGAGCCAACNNNNNN
G
  C CTGGG A CGTGGTCTCGGTTG-5'
```

Reverse Primer:

5′-CTGGTGCAGGGTCCGAGGTA

siRNA	+/-	Stem-loop Overlap	Forward Primer
410	+	aaggcc	GCGCGCGTAATGAAATCCTTCTGG
	-	aataat	CGCGGGCCAGAAGGATTTC
816	+	aagttt	CGCGCGTTTAAAATCCATGCCAA
	-	aattta	GCGCGGTTTGGCATGGATTT
952	+	aagaag	GCGCGTTGCCAATGCTTTTACT
	-	aattgc	GCGCGCGAAGTAAAAGCATTGG
1129	+	aagaat	CGCGCGTCCCTTTATCACAGAAT
	-	aatccc	CGCGCGCGAATTCTGTGATAAAGG
336	+	aaaggc	GCGCGCGCAATTCTATTGATAAGGC
	-	aaaatt	GCGCGCAGGCCTTATCAATAGA
645	+	aaaggt	GCGCGATCTGCTGAGAAGTCAC
	-	aaatct	GCGCGAGGTGACTTCTCAGC
751	+	aaagat	CGCGATCTGGGTGCCAAAGA
	-	aaatct	GCGCAGATCTTTGGCACCC
379	+	aaaatt	GCGCGCATGCACACTGTTCATAA
	-	aaatgc	GCGCGCAATTATGAACAGTGTGC
440	+	aatggg	GCGCGCGCTATATTCTTTCTGTCC
	-	aactat	GCGCGCTGGGACAGAAAGAAT
1198	+	aattgg	GCGCGCAAAGAGTTCCAAAGC
	-	aacaaa	GCGCGTTGGCTTTGGAACTC
640	+	aatctg	GCGCGCTGAGAAGTCACCTTC
	-	aactga	GCGCGCTCTGAAGGTGACTTC
928	+	aatata	GCGCCCGCCTTCTCGTTATT
	-	aaccgc	GCGCGCGTATAATAACGAGAAGGC
304		aatgta	CGCAAGGACGACGGCAAC

Dual-Luciferase Assay

HeLa cells were seeded in a 96-well plate at 15,000 cells/well in 100 μL of media 24 h before transfection. After 24 h, the media was changed, and both the PKR and non-targeting siRNAs were diluted in Opti-MEM and co-transfected with 40 ng/well of either the (+) or (-) psiCHECK2-AS34a plasmid, using 0.35 μL/well of Lipofectamine 2000 for a total transfection volume of 50 μL. Serial dilutions of PKR siRNAs ranging from 3.2-10,000 pM were transfected in duplicate, and transfections were performed twice for each siRNA. The PKR siRNA was diluted in a non-targeting siRNA to maintain total siRNA concentration at 10 nM per well. 24 h post-transfection, the media was aspirated, replaced with 79 μL of Dulbecco's PBS (Invitrogen cat# 14040133), and lysed using 79 μL of Dual-Glo Luciferase Reagent (Promega cat# E2940). After incubating at room temperature for 15 min on a rocker, 150 μL of the solution was transferred to a solid white 96-well plate, and firefly activity was measured using a Synergy H4 microplate reader (Biotek). 75 μL of Dual-Glo Stop&Glo Reagent (Promega cat# E2940) was added, and, after a 10 min incubation period, *Renilla* activity was measured. Relative siRNA activity was determined by subtracting background, dividing *Renilla* signal by the firefly luciferase signal within a well, and then normalizing the ratio to the ratio from a well treated with a non-targeting siRNA. Values of 0 and 1 indicate complete silencing and no silencing, respectively. IC_{50} values were determined by fitting data to the following equation,

$$Y = \frac{1}{1 + 10^{X - Log(IC_{50})}}$$

using least-squares regression, where X is the concentration of targeting siRNA and Y is the relative siRNA activity.

Stem-loop RT-qPCR

HeLa cells were plated in a 6-well plate and transfected after 24 h with 10 nM PKR siRNA and 10 nM of a non-targeting siRNA (internal standard). siRNA-Ago2 co-immunoprecipitation was similar to the procedure described (Beitzinger and Meister, 2011), with several exceptions. Immunoprecipitations were scaled down to 400 μg of cell lysate; cell lysate protein concentrations were determined via Bradford assay (Bio-Rad cat# 5000201) with a BSA standard. Immunoprecipitation of Ago2 was performed with 3.375 μg of Ago2 Antibody (Sigma-Aldrich, Clone 11A9 cat# SAB4200085-200UL) pre-bound to 25 μL of Protein G magnetic beads (NEB cat# S1430S). Proteinase K Digestion Buffer was also supplemented with $\sim$400 μg/mL tRNA (Roche cat# 10109541001), to act as a carrier, and Ago2-bound RNAs purified using Direct-zol RNA purification kit according to the manufacturer's protocol (Zymo Research cat# R2051).

The protocol for RT-qPCR quantification was developed based on published methods (Varkonyi-Gasic and Hellens, 2011; Chen et al., 2005; Kramer, 2001). Stem-loop primers were folded by heating to 95°C for 10 min, ramping to 75°C over 10 min, then holding at 75°C, 68°C, 65°C, 62°C, and 60°C for 30 min each, before ramping to 4°C over 2 h. RNA samples were heated to 65°C for 5 min and snap cooled on ice for 5 min immediately before reverse transcription. Pulsed reverse transcription reactions were assembled with 1x First-Strand Buffer, 10 mM DTT, 0.25 mM dNTP mix, 20 U Superscript III (Life Technologies cat# 18080044), 0.2 U of SUPERaseIn (Ambion cat# AM2696), and 1 nM

stem-loop primer in 20 μL with 1 μL of purified RNA and cycled as follows: 16°C for 30 min followed by 60 cycles of 30°C for 30 sec, 42°C for 30 sec, and 50°C for 1 sec, reverse transcriptase was then heat inactivated at 85°C for 5 min. qPCRs were assembled with 300 nM of an siRNA specific forward primer, 300 nM universal reverse primer, 1x IQ SYBR Green Supermix (Bio-Rad cat# 170882), and 1.8 μL of cDNA in a 25 μL reaction. qPCR cycling was performed as follows: 95°C for 10 min followed by 40 cycles of 95°C for 10 s and 60°C for 10 s on a MyiQ Thermocycler (Bio-Rad). (+) siRNA strand, (-) siRNA strand, and internal standard siRNA strand were reverse transcribed and quantified separately (Table A.3). Amplification specificity was verified by performing stem-loop RT-qPCR on samples that were transfected with an siRNA that was not complementary to the stem-loop primers. Control PCRs were performed on all samples following reverse transcription in the absence of reverse transcriptase. No amplification that interfered with quantification was observed in any PCR control.

Statistical Methods

Linear regression analysis was performed in python with the pandas and statsmodels packages, using an ordinary least squares regression. Briefly, 1- and 2- factor models were built with either, siRNA IC_{50} values or siRNA loading into Ago2 as the response variable, and correlated with TN Rank and $\Delta\Delta G_{3nn}$ (double-stranded analyses) or with TN and ΔG (single-stranded analyses). p values for 1-factor models were used to test if the slope was non-zero. 2-factor models were built using either i) TN Rank with $\Delta\Delta G_{3nn}$ or ii) TN and ΔG at each position along the duplex using p values to test if the slope of the $\Delta\Delta G_{3nn}$ or ΔG parameter was non-zero. A Bonferroni correction was applied to p values to correct for

multiple comparisons. All other statistical analyses were performed using Graphpad Prism 6, with the exception of ΔAICc, which was calculated using RStudio with the AICcmodavg package. ΔAICc and Akaike weights were used to compare linear regression models. Akaike weights provide relative probabilities that one model better fits the data. ΔAICc and Akaike weights contain a correction for models with different numbers of variables, allowing direct comparison among models. When an additional variable is added to a model and the Akaike weight favors the new model, that variable is providing new, explanatory information.

Appendix B

Materials and Methods for Ch. 3

Recombinant Expression of MBP-PACT

Cloning

The gene encoding the PACT protein was subcloned into the pMBP-6His-TEV-MMS2 vector, a gift from Cynthia Wolberger (Addgene # 25465), from the pET15b-PACT plasmid (Kindly provided by Dr. Ganes Sen), using In-fusion cloning (Clontech cat# 639684). Briefly, the PACT gene was PCR amplified using Advantage HD Polymerase mix (Clontech cat# 639241) according to the manufacture's protocol, with a T_m of 57°C and an elongation time of 1 min. The following primers were used for PCR amplification: 5′-GTACTTCCAGGGATCCATGTCCCAGAGCAGGCAC and 5′-GGCCAGTGCCAAGCTTTTACTTTCTTTCTGCTATTATCTTTAAATACTG. The pMBP-6His-TEV-MMS2 plasmid was sequentially digested with HindIII and then BamHI (NEB cat# R0104S and R0136S, respectively). HindIII was heat inactivated before BamHI digestion. Both the plasmid and PCR fragment were cleaned up using a PCR Cleanup kit (Qiagen cat# 28104), recombined by In-fusion cloning, and then transformed into Top10 cells (Invitrogen cat# C404003). Clones were screened by colony PCR and verified by Sanger sequencing with the following primers: 5′-AAGACGCGCAGACTAATTC and 5′-

GGCCTCTTCGCTATTACG.

Expression and Purification

Expression of PACT was preformed similarly to the procedure used to express TRBP (Gredell et al., 2010; Laraki et al., 2008). An overnight culture of the BL21(DE3)-RIPL cells (Agilent cat# 230280) with the pMBP-6His-TEV-PACT plasmid was diluted to OD_{600} 0.05 and cultured in LB broth until the cultured reached an OD_{600} of 0.6 at which time cultures were induced with 300 mM IPTG for 2 h. Cells were collected by centrifugation at 4000 rpm for 10 min and pellets were frozen until purification. Cell pellets were suspended in column buffer (20 mM Tris-HCl pH 7.4, 200 mM NaCl, 1 mM EDTA) and supplemented with 1 mg/mL lysozyme, 250 U of DNase I and 1x Complete protease inhibitor (Roche cat# 11873580001). Cells were incubated for 30 min with gentle rocking followed by sonication. The cell lysate was then cleared by centrifugation at 15,000g for 25 min at 4°C. The cell lysate was passed through a 0.22 μM filter and loaded onto an AKTA FPLC with a MBPTrap HP column (GE Healthcare cat# 28-9187-79). The column was washed with 10 column volumes (CV) of column buffer and step eluted with 5 CV of column buffer supplemented with 10 mM Maltose. Fractions were analyzed by SDS-PAGE and frozen at -80°C. Protein quantification was done with Quick Start Bradford Dye (BioRad cat# 5000205) with BSA standards and purity assessed by SDS-PAGE gel stained with GelCode Blue (ThermoFisher cat# 24590).

Recombinant Expression of His- and MBP-NS3

Cloning of MBP-NS3

The pET28a-NS3 vector was a gift from Dr. Keqiong Ye. The vector was amplified with Top10 cells and used for protein expression without any modifications to the vector. An MBP-tagged version of NS3 was generated by In-fusion cloning into the pMBP-6His-TEV-MMS2 vector. The vector was digested and cleaned up similarly to the generation of the pMBP-6His-TEV-PACT vector. PCR amplification was performed with Advantage HD polymerase master mix with a pre-amplification step. The following primers were used for PCR amplification, 5′-GTACTTCCAGGGATCCATGGGCAACGTGTTCACATCGTC and 5′-GGCCAGTGCCAAGCTTTTACAGCACAGCTGGAGAGCTGC. Pre-amplification was performed by running 10 cycles: 98°C for 10 sec, 64-59°C for 15 sec, and 72°C for 45 sec. The second step in the amplification procedure was ramped down from 64°C to 59°C over the course of the 10 cycles. Amplification of the target was then continued by running 25 cycles as follows: 98°C for 10 sec, 59°C for 15 sec, and 72°C for 45 sec. The PCR product was DPNI treated and cleaned up before In-fusion homologous recombination with the pMBP-6His-TEV-MMS2 vector. Constructs were transformed in to Top10 Competent cells and clones verified by Sanger sequencing.

Expression and Purification

Both a His-tagged and MBP-tagged construct were expressed in *E. coli* Rosetta 2(DE3) cells (Novagen cat# 71397-3). Expression and purification was performed with similar conditions to those used by Shen et al. (2010). An overnight culture of expressing *E. coli*

Rosetta 2(DE3) expressing MBP-NS3 was used to inoculate cultures to an OD_{600} of 0.01 and grown to an OD of ∼0.4 at which time the incubator was turned down from 37°C to 16°C. Once the cultures reached an OD_{600} of ∼0.6 cultures were induced with 400 μM of IPTG and cultured overnight. Cells were collected by centrifugation and frozen at -80°C.

Purification of His-NS3 required the use of a HisTrap HP column (GE Healthcare cat# 17-5247-01), followed by a Heparin HP column to remove small contaminating RNAs. Briefly, the cell pellet containing His-tagged NS3 was suspended in H300 Buffer (20 mM HEPES-KOH pH 7.6, 300 mM KCl, 5% Glycerol, and 25 mM Imidazole) supplemented with 1x Complete Protease Inhibitors, 1 mM EDTA, and 5 μg/mL DNAse I (Roche). Cells were incubated with 1 mg/mL Lysozyme for 10 min and then sonicated. Cell supernatant was collected by centrifugation at 15,000 rpm and then passed through a 0.22 μM filter. The clarified cell lysate was then passed over a HisTrap HP column using the AKTA FPLC system. The column was then washed with 10 CV of the same H300 buffer but with 50 mM Imidazole. A linear gradient going from 50 mM Imidazole to 500 mM Imidazole over 5 CV was used to elute His-tagged NS3 from the column. Fractions containing the most protein were pooled and loaded onto a HiTrap Heparin HP (GE Healthcare cat# 17-0406-01) column equilibriated in Buffer H (20 mM HEPES-KOH pH 7.6, 5% Glycerol) with 100 mM KCl. The column was washed with 5 CV of Buffer H with 300 mM KCl and His-NS3 eluted with a linear gradient between 300 mM and 1000 mM KCl over 3CV. Fractions were analyzed by SDS-PAGE and the fractions containing the His-NS3 were supplemented with DTT to a final concentration of 5 mM and then flash frozen and stored at -80°C.

Nucleic Acids and 5′-^{32}P Labeling

siRNAs used for cross-linking assays (Figure 2.1) are the same as those used Gredell et al. (2010). siRNAs were synthesized as single strands with and without the substituted 4-thiouracil base by Dharmacon. siRNA strands were either Hot labeled with ATP-[$\gamma -^{32} P$] (Perkin Elmer cat# BLU502A) or Cold with ATP. Labeling reactions were setup with 3 pmol of RNA and 10 pmol ATP using T4 Polynucleotide Kinase (NEB cat# M0201) in a 25 μL reaction. Complementary siRNA strands mixed and supplemented with 1x STE (10 mM Tris-HCl pH 7.4, 100 mM NaCl, and 1 mM EDTA), then heated to 90°C for 3 min, and cooled to 37°C for 60 min. Excess label was removed using G-25 sephadex spin columns (Roche cat# 11273990001).

Gel Shift Assay

Native gel shift assays were used to confirm functionality of the MBP-PACT and His-NS3 constructs as well as to determine their affinity for both siRNAs and ssRNAs. Briefly, 30,000 cpm ($\sim$1 nM) of 5′ labeled siRNA or ssRNA was incubated with an increasing protein in a 10 μL reaction. MBP-PACT binding reactions were supplemented with 20 mM HEPES pH 7.6, 40 mM KCl, 1.5 mM $MgCl_2$, 0.1% CA630, and 10 U of SUPERaseIn (Ambion cat# AM2696). His-NS3 binding reactions were supplemented with Hepes-KOH pH 7.6, 100 mM KCl, 2 mM $MgCl_2$, 1 mM DTT, and 0.01% CA630. Binding reactions were incubated for 30 min at room temperature after which 2 μL of 5x loading was added to each reaction, and then loaded on a native 1x TBE polyacrylamide gel. After electrophoresis, the gel was transferred to Whatman filter paper and dried under vacuum at 80°C for 60

min. The dried gel was exposed to a storage phosphor screen for ∼16 h and imaged on an Amersham Storm 860 scanner.

Binding and Cross-linking Assay

Experimental Design

TRBP localization was measured, previously, with the use of a substituted photoreactive uracil base (4-thiouracil) at the 20^{th} position of the siRNA (Gredell et al., 2010). Asymmetric siRNA terminus localization was demonstrated by alternating the location of the 4-thiouracil cross-linker between the two termini and comparing the fractions of TRBP cross-linked for each configuration (Figure 3.1A). The fraction of an siRNA strand that cross-linked to TRBP was visualized with a ^{32}P-label, enzymatically incorporated on the $5'$ terminus of the siRNA strand containing the 4-thiouracil base. The complexed and free siRNA strands were resolved using denaturing SDS-PAGE, allowing the separation of the two RNA strands, and visualized with a ^{32}P-label on the $5'$ terminus of the siRNA via Phosphor imaging.

Protein-siRNA binding reactions were incubated for 30 min, followed by cross-linking of the 4-thiouracil base to nearby amino acids by exposure to 312 nm light (Meisenheimer and Koch, 1997; Sontheimer, 1994). The 4-thiouracil cross-linker forms only short-range cross-links, providing the resolution needed to distinguish differences in cross-linking between the two-siRNA termini (Gredell et al., 2010; Pellino, 2007). Excitation at wavelengths > 300 nm limits non-specific RNA cross-linking, which occurs at shorter wavelengths (250-270 nm) (Meisenheimer and Koch, 1997). The individual siRNA strands are indistinguishable

by SDS-PAGE, thus, requiring cross-linking reactions to be performed separately for each siRNA strand (Figure 3.1A). Differences in terminus localization were correlated with siRNAs with known asymmetric (pp-luc) and symmetric (sod1) functional asymmetries, as well as, a third sequence with unknown asymmetry (EGFP#274; Figure 3.1B) (Tomari et al., 2004; Gredell et al., 2010).

Two different strategies are used to manipulate siRNA $\Delta\Delta$G for study of its effect on siRNA asymmetric binding (Tomari et al., 2004; Noland et al., 2011; Gredell et al., 2010; Sakurai et al., 2011), asymmetric strand selection (Suzuki et al., 2015) (Chapter 2), and functional asymmetry (Schwarz et al., 2003; Sakurai et al., 2011). The first method is to introduce mismatches between the $5'$ nucleotide of one strand and the 19^{th} nucleotide of its complement (Noland et al., 2011; Schwarz et al., 2003; Tomari et al., 2004; Gredell et al., 2010). The second method is to change the siRNA sequence and model the change in duplex thermodynamics with nearest-neighbor parameters (Gredell et al., 2010; Tomari et al., 2004; Schwarz et al., 2003; Suzuki et al., 2015). While terminal mismatches can successfully alter the asymmetric cross-linking pattern between TRBP and some duplexes, not all siRNA cross-linking patterns can be manipulated with terminal mismatches (Gredell et al., 2010), as they are in functional siRNA mediated silencing assays (Schwarz et al., 2003). Alternatively, by using siRNAs with varying sequence, the siRNA structures are geometrically more consistent but there is the concern that the local RNA sequence can change the efficiency of 4-thiouracil cross-linking (Noland et al., 2011). For this reason, we used a generic dsRBP, NS3, to demonstrate that the efficiency of cross-linking does not change with local RNA sequence.

Experimental Setup

Similar to the gel shift assay, 10 μL protein binding reactions with 30,000 cpm ($\sim$1 nM) of siRNA (with 4-thiourcail base substitution, Figure 2.1A) was incubated 50 nM of His-NS3 or 1 μM of MBP-PACT for 30 min at room temperature, allowing binding reactions to reach equilibrium. Binding reactions were the placed on an ice cold aluminum block, covered with a petri dish to block wavelengths $<$ 300 nM and exposed to UV light (312 nM) with a Transilluminator (Fisher Scientific) for 10 min. Reactions were then mixed with 5 μL 3x SDS Loading Dye (NEB cat# B7703S) and heated to 95°C for 5 min before being resolved by SDS-PAGE analysis. SDS-PAGE gels were resolved similar to native gels.

Appendix C

Expression of Bioactive Brain-Derived Neurotrophic Factor (BDNF) in *Brevibacillus choshinensis*

This work was produced as part of a collaboration and contains a significant amount of work but does not fit thematically within the rest of my dissertation and is thus being included in the appendix.

Abstract

Background :

Brain-derived neurotrophic factor (BDNF) is a member of the neurotrophin family critical for neuronal cell survival and differentiation, making characterization of its function essential for understanding nerve cell biology and for the development of therapeutics for neurological disorders and spinal cord injuries. Neurotrophins like BDNF fold with a characteristic cystine-knot conformation, complicating expression of soluble, bioactive recombinant protein in most traditional microbial expression systems.

Results :

Here we investigated the production of BDNF using *Brevibacillus choshinensis*. We evaluated the effects of media type (2SY and TM), temperature (25°C and 30°C), and culture time (48-120 h) on the production of bioactive BDNF. Protein production was higher in TM media, though recoveries by Ni^{2+} chromatography were comparable to 2SY media. Greater bioactivity (per unit mass) was observed for BDNF produced in 2SY cultures at extended times (96 h at 30°C or $>$ 72 h at 25°C), which was comparable to bioactivity of commercially-available BDNF.

Conclusion :

This study provides confirmation that B. choshinensis is capable of producing biologically active BDNF. The choice of culture conditions impacted the production rate and purification of bioactive BDNF, and further optimization is possible. Of the conditions tested, the condition that led to the greatest quantity of biologically active protein in the shortest culture time was in 2SY media at 25°C for 72 h resulting in 264 $\pm$ 82 μg/L of bioactive BDNF.

Introduction

Brain-derived neurotrophic factor (BDNF) plays a direct role in the regulation of multiple processes involved in neuronal cell growth, differentiation and survival (Numakawa et al., 2010; Leibrock et al., 1989; Binder and Scharfman, 2004; Ibáñez and Simi, 2012). BDNF is natively expressed as proBDNF, which readily dimerizes intracellularly (Heymach and

Shooter, 1995; Jungbluth et al., 1994) and matures upon proteolytic cleavage of the pro peptide (Lee et al., 2001; Mowla et al., 2001; Lu, 2003). BDNF signaling is mediated by two receptors, tropomyosin-related kinase B (TrkB) and p75 (Soppet et al., 1991; Teng et al., 2005; Numakawa et al., 2010; Nagahara and Tuszynski, 2011). The TrkB receptor recognizes the mature dimerized form of BDNF and enhances neuronal growth and survival (Numakawa et al., 2010; Lu et al., 2005a), while the p75 receptor recognizes proBDNF causing neuronal cell death (Teng et al., 2005; Lu et al., 2005a; Numakawa et al., 2010; Ibáñez and Simi, 2012). Disruption in BDNF signaling has been implicated in a number of neurodegenerative diseases, including Alzheimer's, Huntington's, and Parkinson's (Zuccato and Cattaneo, 2009; Nagahara and Tuszynski, 2011; Lu et al., 2005a). BDNF, and other members of the neurotrophin family, are also necessary to direct neuronal regeneration post-injury (Menei et al., 1998; Lu et al., 2005b). For these reasons, studies of BDNF function are critical for understanding neuronal biology and for the development of new therapeutic strategies for neurodegenerative diseases (Nagahara and Tuszynski, 2011) and spinal cord injuries (Lynam et al., 2015). However, studies of BDNF function are limited by the cost of the protein from commercial sources.

BDNF is a cystine-knot protein (Binder and Scharfman, 2004), containing a series of non-sequential disulfide bonds that make it difficult to express properly folded protein that retains the ability to homodimerize (Numakawa et al., 2010). Attempts to express BDNF in a variety of hosts have been limited by aggregate formation and low bioactivity (*E. coli*, *B. subtilis*, and *S. cerevisiae*) (Fukuzono et al., 1995; Takeshita et al., 1996; Hoshino et al., 2002; Park and Shimizu, 1996; Burns et al., 2014, 2016) or low yields (Sf21, RK_{13}, CHO, and HEK cells) (Meyer et al., 1994, 1992; Burton, 1993; Knusel et al., 1991), motivating

the need for a better host for expressing BDNF at useful quantities with reasonable labor and material costs.

Brevibacillus choshinensis, a gram-positive bacterium, was originally isolated from soil and found to secrete large quantities of protein with low extracellular protease activity (Takagi et al., 1989a). Because *B. choshinensis* is gram-positive, secreted protein can be purified without concern for contaminating endotoxins (Ilk et al., 2011). *B. choshinensis* has been used to successfully express a number of mammalian proteins, including the growth factors VEGF and NGF, which are, like BDNF, cystine-knot proteins (Sun and Davies, 1995). In this work, we have expressed the mature BDNF sequence using *B. choshinensis*, preliminarily investigating the effects of temperature, media composition, and culture time on bioactive protein production. We have found that culture conditions greatly impacted the production and purification of bioactive BDNF. This study demonstrates that *B. choshinensis* is capable of producing useful quantities of bioactive BDNF in laboratory scale cultures.

Methods

Plasmid Construction and Expression Strain

BDNF cDNA encoding the mature sequence (amino acids 129-247) was purchased from Bioclon (San Diego, CA) and subcloned with a C-terminal 6x His-tag into the pNCMO2 *B. choshinensis* expression vector (Clontech cat# HB112) by In-fusion cloning (Clontech cat# 638909). Briefly, the BDNF gene was PCR amplified with the following primers, 5′-TCCCATGGCTTTCGCTCACTCTGACCCGGCTCGC and

5′-TACCGAATTCCTCGACAGCCGGATCTCAGTGGTGG, containing 16 bp and 15 bp 5′ overlaps with the pNCMO2 plasmid, respectively. An additional nucleotide was included in the forward primer to maintain the correct translational frame. The pNCMO2 vector was digested with PstI-HF and XhoI restriction enzymes overnight (NEB cat# R3140 and R0146, respectively). Both the PCR product and plasmid were cleaned up using a PCR cleanup kit (Qiagen cat# 28104), recombined with In-fusion cloning, and transformed into JM109 cells (Promega cat# L1001). Colonies were screened by colony PCR and verified by Sanger sequencing with the following primers, 5′-CGGGCTTTAAAAAGAAAGATA and 5′-CCAAATGGTGTTACTTTGAGA.

B. choshinensis competent cells were purchased from Clontech (cat# HB116). The strain used in this study has been genetically modified to remove any remaining proteolytic genes and the genes responsible for sporulation (Hanagata and Nishijyo, 2010). Transformation of verified plasmids was performed by the Tris-PEG method, per the manufacturer's instructions (Clontech). Single colonies were obtained after transformation, and protein expression was verified by western blotting (details below).

B. choshinensis Culture Conditions

Protein was expressed in 50 mL shake flasks in either TM Media (1% glucose, 1% polypeptone, 0.5% meat extract, 0.2% yeast extract, 0.001% $FeS0_4$*$7H_2O$, 0.001% $MnSO_4$*$4H_2O$, 0.0001% $ZnSO_4$*$7H_2O$, 50 μg/mL neomycin) or 2SY Media (2% glucose, 4% soytone, 0.5% yeast extract, 0.015% $CaCl_2$*$7H_2O$, 50 μg/mL neomycin), per the manufacturer's recommendations (Clontech). Cultures were inoculated with a single colony from transformed *B. choshinensis* and grown overnight at 30°C and 200 RPM in TM media. 50 mL cultures

were inoculated to an OD_{600} of 0.01 and grown either at 25°C or 30°C and 200 RPM. When needed, cultures were diluted in PBS to ensure accuracy of the OD_{600} measurements. Samples were removed or cultures collected at specified time points ranging up to 96 h at 30°C and 120 h at 25°C.

Protein Purification

Cells were separated from the BDNF-containing supernatants by centrifugation at 8000g for 10 min. Supernatants were supplemented with 10x MOPS buffer (250 mM MOPS pH 7.4, 5 M NaCl, and 40 mM imidazole) to a 1x final concentration to control pH, ionic concentration, and non-specific binding to Ni^{2+} Sepharose 6 Fast Flow (referred to as Ni^{2+} sepharose; GE Healthcare cat# 17-5318-06). Supernatants were then centrifuged at 16,000g for 20 min at 4°C to remove any remaining insoluble material. 35 mL of the recovered supernatants were then incubated with 200 μL of Ni^{2+} sepharose for 1 h with end-over-end rotation and recovered by centrifugation at 1000g for 5min. The Ni^{2+} sepharose beads were washed 3 times in 1x MOPS Buffer and eluted 3 times in 1x MOPS buffer with 300 mM imidazole, using centrifugation at 1000g for 5 min to separate the Ni^{2+} sepharose from the wash or elution buffer. All three elutions were combined and filter sterilized with a 0.22 μm PVDF low binding filter (Millipore cat# SLGV013SL).

SDS-PAGE Analysis, Western Blotting, and BDNF Quantification

Samples for SDS-PAGE and western blotting were first supplemented with 3x SDS Loading Dye (Cell Signaling cat# 7722) and 30x 1.25 M DTT, and then heated to 95°C for 5 min before being resolved on a 4-20% SDS-PAGE gel. Gel staining was performed using

GelCode Blue (Thermo Fisher cat# 24590) according to the manufacturer's instructions. Separate gels were run for western blot analysis. Gel-separated proteins were transferred onto a 0.2 μm nitrocellulose membrane (Bio-Rad cat# 161-0112) at 90 V for 70 min. BDNF was detected with either an anti-BDNF(N20) (Santa Cruz cat# sc-546), anti-BDNF(H117) (Santa Cruz cat# sc-20981), or HRP conjugate anti-6x His (Cell Signaling cat# 9991) antibody. Detection with the anti-BDNF(N20) antibody was performed as follows, membranes were blocked for 1 h at RT in 5% Milk-TBST, incubated with primary antibody at 1:2500 dilution overnight at 4°C in 5% Milk-TBST, washed 3 times for 15 min in TBST, incubated in secondary HRP-linked anti-Rabbit antibody at a 1:1000 dilution in TBST at RT for 2 h, washed 3 times for 15 min in TBST, and then resolved with SuperSignal West Femto Maximum Sensitivity Substrate (ThermoFisher cat# 34094). Detection with anti-BDNF(H117) antibody was performed with similar conditions, except 5% BSA was used instead of 5% Milk, primary antibody dilution was at 1:200, and secondary antibody incubation was in 5% BSA-TBST. A biotinylated ladder (Cell Signaling cat# 7727) was used for size identification on western blots; HRP-linked anti-biotin antibody was added at 1:5000 dilution with the secondary antibody. For detection with the anti-6x His antibody, membranes were blocked with 5% BSA-TBST (0.1% Tween-20) for 3 h, incubated with primary antibody at 1:5000 dilution overnight at 4°C in 5% BSA-TBST (0.1% Tween-20), washed 3 times for 15 min, washed 3 times for 15 min and resolved with the same chemiluminescent substrate. A Tween-20 concentration of 0.05% was used unless otherwise specified.

BDNF concentrations were determined by western blotting with the anti-6x His antibody and quantified relative to a set of dilutons of a BDNF standard. The BDNF standard was purified to homogeneity from TM culture supernatant under denaturing and reducing

conditions and quantified by Bradford Assay in comparison to a BSA standard (Figure C.1). Due to the denaturing conditions required for purification, the BDNF standard retained no bioactivity and thus could only be used as a western blot standard.

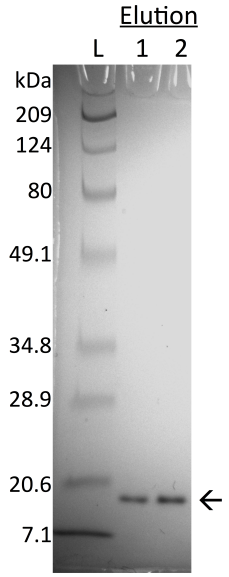


Figure C.1: Purity of BDNF Standard. A BDNF standard was purified to homogeneity from *B. choshinensis* cultures using Ni^{2+} sepharose and quantified by Bradford assay using a BSA standard. Purification required denaturing and reducing conditions, which disrupted the BDNF structure and precluded bioactivity testing. Protein concentrations were 59 μg/mL and 100 μg/mL for elutions 1 and 2, respectively.

Proliferation Assay for BDNF Bioactivity

BDNF bioactivity was assayed via cell proliferation in mouse fibroblast 3T3 cells overexpressing the TrkB receptor (3T3-TrkB; generously provided by Jeffrey Sakamoto) (McCarty and Feinstein, 1998). For the bioactivity assay, black, 96-well, clear bottom plates (Costar cat# 3904) were coated with 50 μL of 50 μg/mL poly-L lysine (Sigma, cat# P9155) for 1 h, washed 3 times with water, air-dried, and then coated with 50 μL of 2.5 μg/mL fibronectin (Sigma cat# F1141) overnight. Plates were washed with DMEM (Life Tech cat# 11965092) before cell plating. Cells were seeded at 2000 cells/well in defined media

(3:1 DMEM 11965:Hams F12 (Life Tech cat# 11965092 and 11765047, respectively), 15 mM HEPES, 4 μM $MgCl_2$, 3 mM L-histidine, 10 M ethanolamine, 1x ITS+1 (Sigma cat# I2521), 2 μM hydrocortisone, 150 μg/mL G418 (ThermoFisher cat# 10131035) for 3 h. To normalize for cell number, 10 μL ($1/10^{th}$ well volume) of AlamarBlue (ThermoFisher cat# DAL1285) was also included in the media. AlamarBlue fluorescence was measured on a Biotek H4 plate reader (ex: 570 nM, em: 590 nM, filter: 9.0) after 3 h. Initial fluorescence readings were used to normalize for differences in cell plating. Media was then changed to BDNF-containing defined media, adding 150 μL of media and exchanging 100 μL every 24 h for 4 days. Dilutions of each protein preparation were tested starting with 1% v/v purified protein and serially diluted 1:3 to 4.6x10^{-4}% v/v. For consistency in the media composition and to control for effects of the elution buffer, a 1% v/v composition of elution buffer was used in all tests, equivalent to the concentration of buffer in the most concentrated protein samples. Commercial BDNF served as a positive control for protein activity (Affymetrix cat# 14-8366-62). This protein did not possess a His-tag and therefore was not used as a standard in the western blotting studies described above. A mock protein preparation of *B. choshinensis* carrying the empty pNCMO2 vector under each culture condition was used as a negative control. After 4 days, AlamarBlue fluorescence was measured again, as before, and fold-difference in proliferation was calculated after background subtraction using the following equation:

$$Fold - Difference = \frac{F_{Final,Sample}/F_{Initial,Sample}}{F_{Final,Control}/F_{Initial,Control}}.$$

where F is the relative fluorescence intensity. AlamarBlue fluorescence was verified to increase linearly with cell number (Figure C.2). Images of cell proliferation were obtained

by plating 20,000 3T3-TrkB cells/well in a 12-well plate. Plate coating was performed as described above with volumes of poly-L lysine and fibronectin adjusted to 400 μL/well. The assay was performed as described but with the omission of AlamarBlue readings. Images were taken prior to each media change.

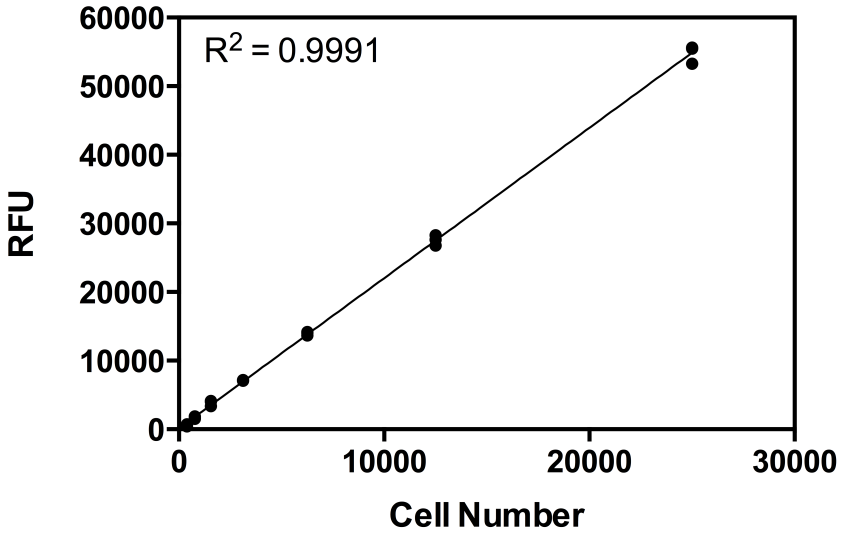


Figure C.2: Standard Curve of AlamarBlue Fluorescence. 3T3-TrkB cells were plated in black-walled, clear-bottomed plates for 3 h. AlamarBlue was then added and incubated for 1 h. AlamarBlue fluorescence was then measured on a Biotek H4 plate reader Ex: 570 nm/Em: 590 nm; N=3.

Statistical Analysis

2-tailed t-test and 2-way ANOVA analyses were performed using Graphpad Prism 6. BDNF bioactivities were compared by equivalence test using 90% confidence intervals to compare prepared BDNF to commercial BDNF. These analyses were done using 1-way ANOVA with a Dunnett *post hoc* test in Graphpad Prism 6.

Image Capture and Alignment

Images were captured using a Spot RT Color camera (Diagnostic Instrument, Inc.) on a Leica Microsystems DMIL inverted microscope. Images were visually aligned to marks on the bottom of the plate each day and alignment was adjusted using ImageJ with the align slices in stack plugin. Images were then cropped to 800x1000 using the TransformJ plugin (Meijering et al., 2001) and contrast normalized through a Stack Contrast Adjustment plugin (Capek et al., 2006). Images were arranged using the tiff package in R.

Results and Discussion

Secretion of Soluble BDNF from *B. Choshinensis*

His-tagged BDNF was constitutively expressed under the P2 promoter in *B. choshinensis* using the pNCMO2 vector with an N-terminal secretion tag. Two types of growth media are suggested for expression in *B. choshinensis*, TM and 2SY. We monitored *B. choshinensis* growth and BDNF expression in both media types over 72 h at 30°C and 96 h at 25°C (Figure C.3A, C.3C, C.3E, C.3G). Supernatant and cellular fractions were separated by centrifugation and analyzed by western blotting (Figures C.3B, C.3D, C.3F, and C.3H), confirming the expression of BDNF, the presence of a His-tag, and the presence of soluble BDNF in the culture supernatants.

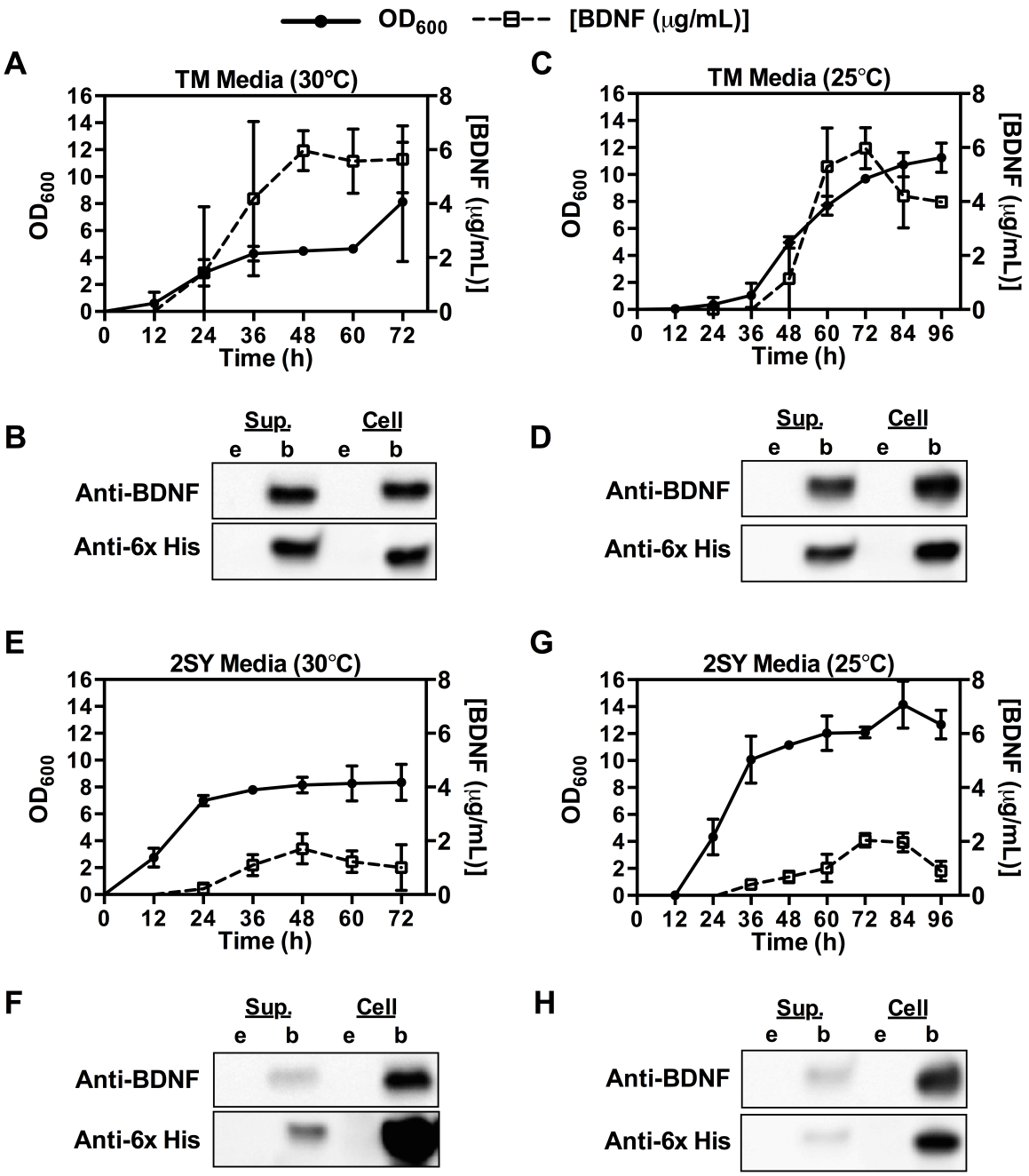


Figure C.3: *B. choshinensis* **Growth and BDNF Expression.** Growth of *B. choshinensis* and evaluation of BDNF secretion using two different media types, TM (A, C) and 2SY (E, G) each at two different temperatures, 25°C (C, G) and 30°C (A, E). Growth was assayed using measurements of OD_{600} (solid curves). Protein expression was measured by western blotting (dashed curves). N=3; error bars are ± 1 SD. (B, D, F, H) Anti-BDNF (top panel) and anti-6x His (bottom panel) antibodies against samples from *B. choshinensis* cultures transformed with an empty vector (e) or the BDNF vector (b). Results are shown for the supernatant (Sup.) and cell pellet (Cell) fractions.

Protein recovered in both the supernatant and cellular fractions appeared at the size of mature BDNF, indicating removal of the *B. choshinensis* secretion tag (with tag: 17.7 kDa and without tag: 14.5 kDa; Figure C.4). Accumulation of BDNF without the secretion tag in the cellular fraction indicates that BDNF is successfully secreted through the plasma membrane but then either remains associated with the cell wall or precipitates out of solution; in either case, it remains associated with the cellular fraction during centrifugation. The secretion tag is derived from the middle wall protein of *B. choshinensis* (Mizukami et al., 2010; Yamagata et al., 1987) and mediates secretion of as much as 1.5 g/L of recombinant, disulfide-bonded protein (Takagi et al., 1989b). Thus, we conclude that the absence of detectable protein with secretion tag is due to the robustness of the secretion system (i.e., minimal intracellular accumulation of BDNF). BDNF associated with the pellet was found to be largely insoluble (Figure C.5) and not further purified or tested for bioactivity. That said, media type affected the distribution of BDNF between the supernatant and cell fractions, with a greater fraction of BDNF remaining in the supernatant in the TM cultures relative to the 2SY cultures (TM - 45 $\pm$ 4% vs 2SY - 13 $\pm$ 2%; $p < 0.05$, 2-tailed t-test; N = 2; Figure C.3).

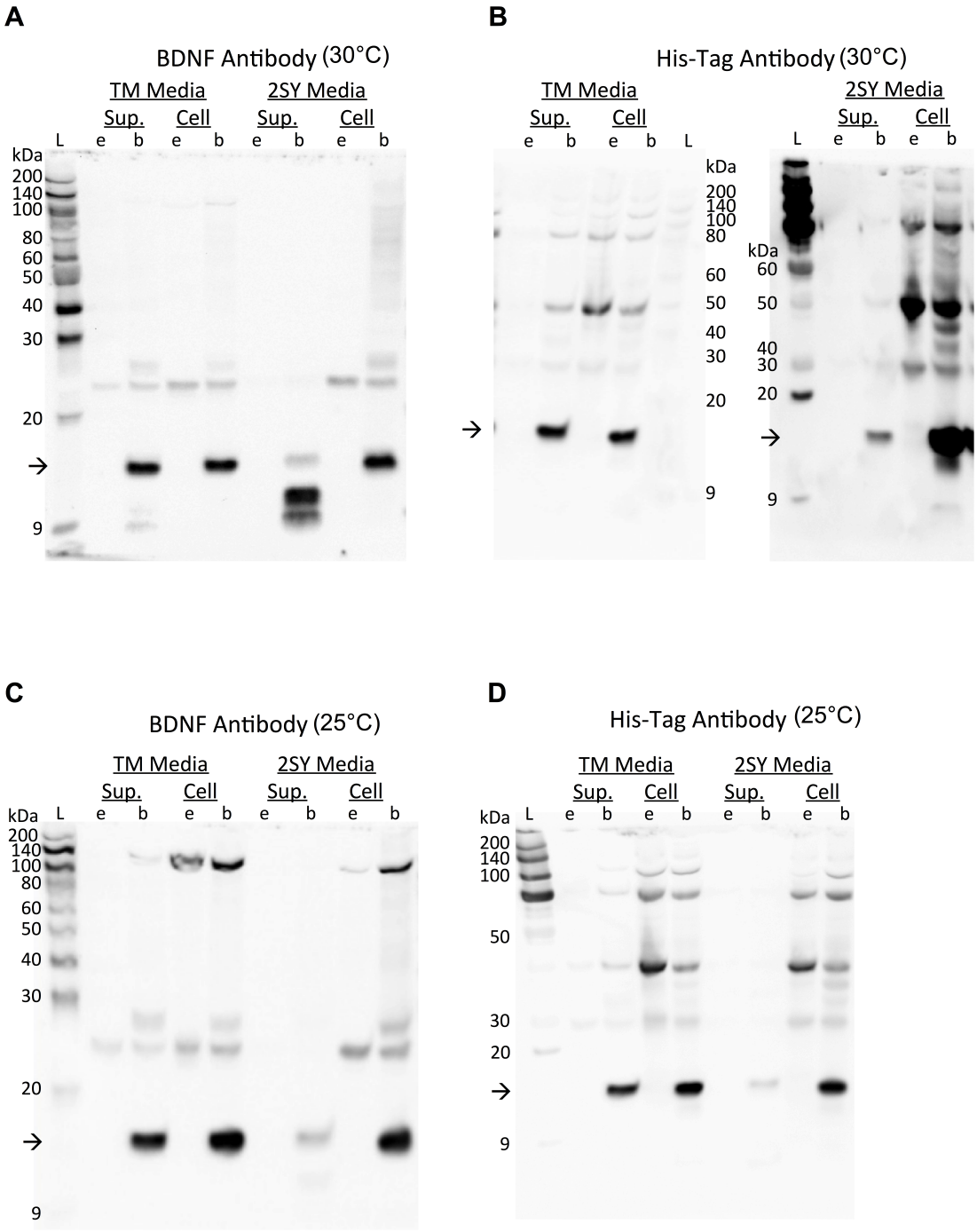


Figure C.4: BDNF Size Indicates Loss of Secretion Tag. Representative western blot of culture supernatants and cell pellet fractions taken from *B. choshinensis* cultures expressing either an empty pNCMO2 vector (e) or BDNF expressing pNCMO2-BDNF vector (b) at 30°C (A, B) or 25°C (C, D) after 48 h and 72 h of culture, respectively. Samples were probed with an anti-BDNF(N20) antibody (A, C) and an anti-6x His antibody (B, D). Arrow indicates band specific to BDNF (theoretical size = 15 kDa). Lane L - molecular weight ladder (kDa).

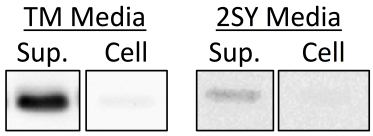


Figure C.5: BDNF in Cellular Fraction is Largely Insoluble. Cell pellets from 30°C TM and 2SY cultures were suspended in 1x BugBuster (EMD Millipore cat# 70584-3) cell lysis reagent and centrifuged to remove insoluble material. Equivalent volumes of cell supernatant and the solubilized fraction from the cell pellet were loaded in each well.

Purification of BDNF by Ni^{2+} IMAC for Bioactivity Testing

Cultures were grown in triplicate, and BDNF was purified from culture supernatants by immobilized metal affinity chromatography (IMAC) using Ni^{2+} sepharose. With this one-step protocol, we were able to purify sufficient quantities to test BDNF bioactivity, but the low stringency method resulted in variable protein recovery (Figure C.6A, Table C.1) and purity (Figure C.7). Higher yields were obtained for lower initial protein loadings indicating saturation of the Ni^{2+} sepharose; however, the maximum quantity of BDNF recovered (18 μg) was well below the theoretical binding capacity of the Ni^{2+} sepharose (8 mg). Nonetheless, some BDNF did not bind to the Ni^{2+} sepharose (indicated by presence of BDNF flow through, Figure C.6B) or bound and did not elute with imidazole (Figure C.62B). Both of these difficulties indicate the presence of aggregates that can preclude His-tag accessibility and precipitate on the Ni^{2+} sepharose, potentially fouling the Ni^{2+} sepharose and limiting its binding capacity in the process.

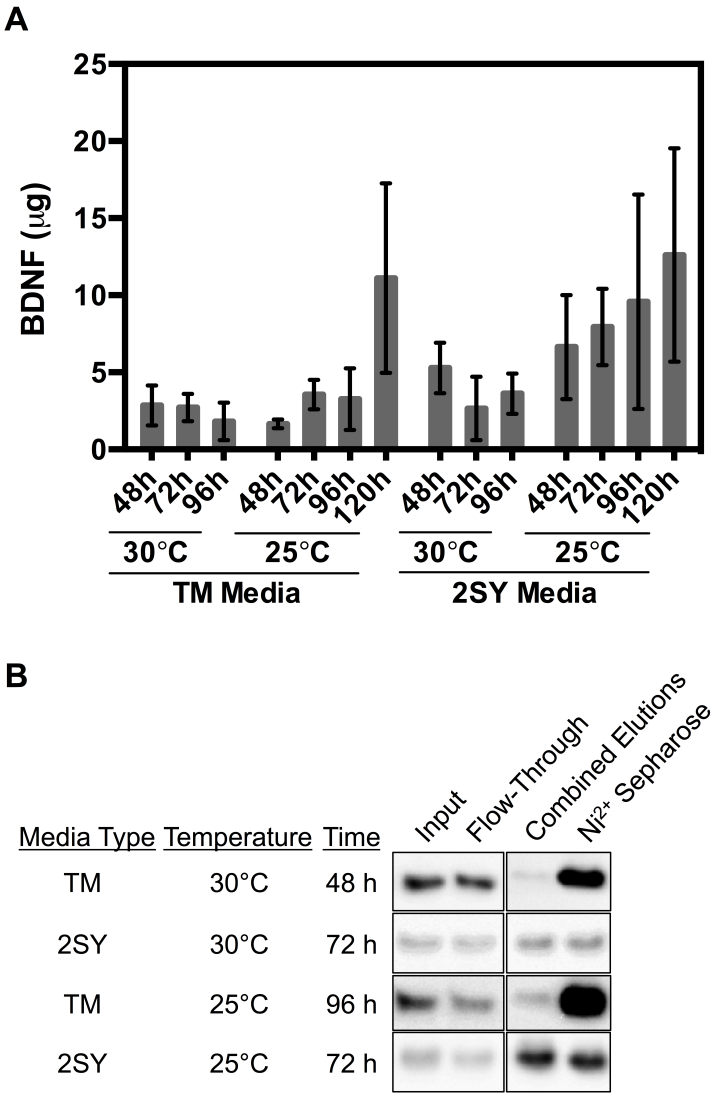


Figure C.6: Purification of BDNF. (A) Quantity of BDNF recovered by IMAC purification determined by western blotting with a standard curve. N=3; error bars represent ± 1 SD. (B) Representative western blots of samples taken during the purification process including: input (35 mL), flow-through (35 mL), combined elutions (600 μL), and Ni^{2+} sepharose suspended in elution buffer (600 μL); volumes are the total volumes at each step in the process. 6.67 μL were loaded per lane. The Ni^{2+} sepharose suspended in elution buffer was analyzed to detect BDNF that remained associated with the Ni^{2+} sepharose after elution. Samples for each culture condition (each row) were run on a single western. Images show the purification fractions for samples from the earliest time at which maximal protein bioactivity was measured for the given condition (see Figure C.8).

Table C.1: Protein Recovered at Stages of the BDNF Purification Process. Quantity of BDNF from culture supernatants (Protein Loaded) and following protein purification (Protein Purified) was quantified by western blotting. N=3 (except for (a) where N=2); error bars represent ± 1 SD.

Media Type	Temperature (°C)	Time (h)	Protein Loaded (ng)	Protein Purified (ng)	Yield (%)
TM	30	48	721. ± 54.	2.84 ± 1.2	0.38 ± 0.148
		72	194. ± 10.	2.71 ± 0.8	1.38 ± 0.399
		96	201. ± 197	1.81 ± 1.2	1.11 ± 0.455
	25	48	172. ± 243	1.65 ± 0.2	3.86 ± 3.383
		72	178. ± 35.	3.56 ± 0.9	2.07 ± 0.797
		96	149. ± 23.	3.25 ± 2.0	2.11 ± 1.044
		120	430. ± 265	11.1 ± 6.1	2.79 ± 0.557
2SY	30	48	60.5 ± 26.	5.29 ± 1.6	11.6 ± 9.882
		72	69.5 ± 30.	2.65 ± 2.0	3.99 ± 3.773
		96	73.3 ± 32.	3.61 ± 1.3	5.97 ± 3.844
	25	48	12.7 ± 2.2[a]	6.63 ± 3.3	66.1 ± 37.21
		72	47.9 ± 12.	7.94 ± 2.4	16.8 ± 3.685
		96	157. ± 52.	9.58 ± 6.9	5.58 ± 2.260
		120	137. ± 17.	12.6 ± 6.9	9.47 ± 5.749

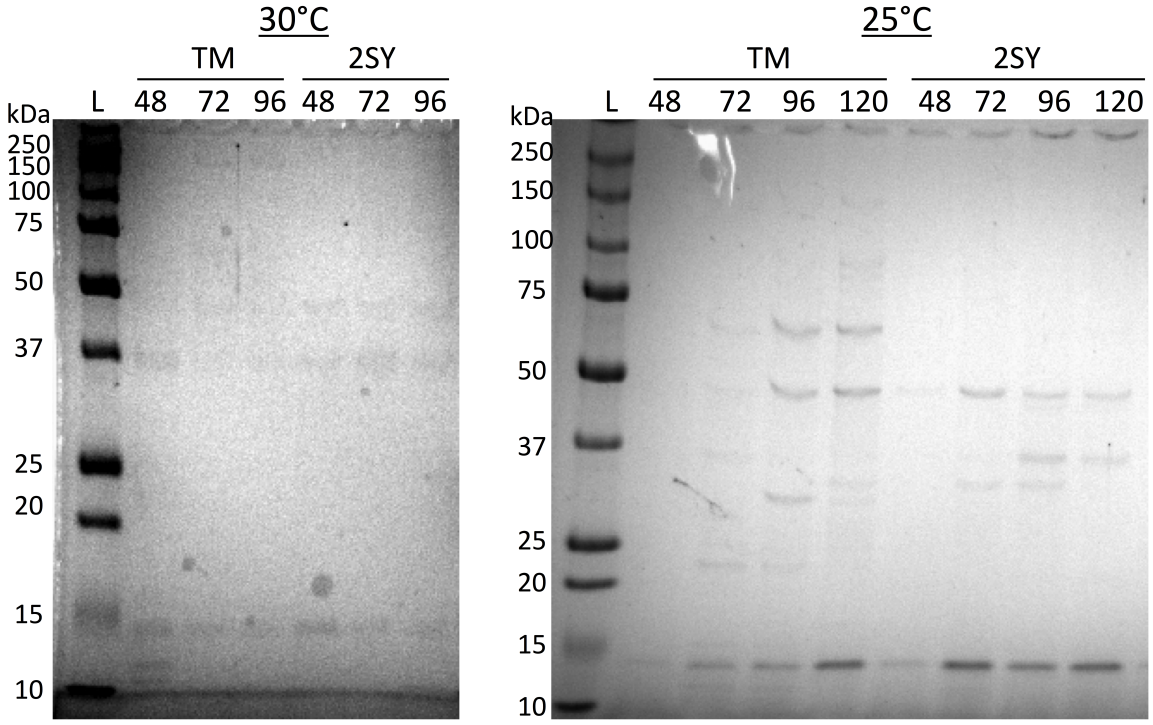


Figure C.7: BDNF Purity After IMAC Chromatography. BDNF purified by IMAC chromatography was analyzed on a 4-20% denaturing SDS-PAGE gel and stained with GelCode Blue total protein stain. Each lane is denoted with the media type, culture temperature, and time of culture growth (h). Lane L - molecular weight ladder (kDa). By image analysis, BDNF purity after this step is estimated to be on average 50%.

Recombinant BDNF is Bioactive

BDNF bioactivity was measured via a proliferation assay, using mouse NIH3T3 fibroblasts, which do not naturally express the TrkB receptor, and were instead engineered to respond to BDNF through overexpression of TrkB (McCarty and Feinstein, 1998). Activation of TrkB signaling with BDNF causes a mitogenic response in the 3T3-TrkB cells (McCarty and Feinstein, 1998). In the absence of serum and BDNF, 3T3-TrkB cells do not proliferate but remain generally healthy and metabolically active (Figure C.8C, (-) Control). Cell proliferation in the presence or absence of BDNF in serum-free media was monitored over the course of 4 days with the media changed daily. After 4 days, AlamarBlue was used

to assess cell number, and proliferation was determined by the increase in cell number of BDNF-treated wells relative to control wells (Figure C.8A). Because BDNF samples were not dialyzed after elution from the Ni^{2+} sepharose, the imidazole concentration present at 1% v/v of purified BDNF (3 mM) was made up in cultures exposed to diluted BDNF samples. A single EC_{50} curve was fit to bioactivity data of BDNF protein prepared from three replicate *B. choshinensis* cultures for each culture condition and time (Figure C.9).

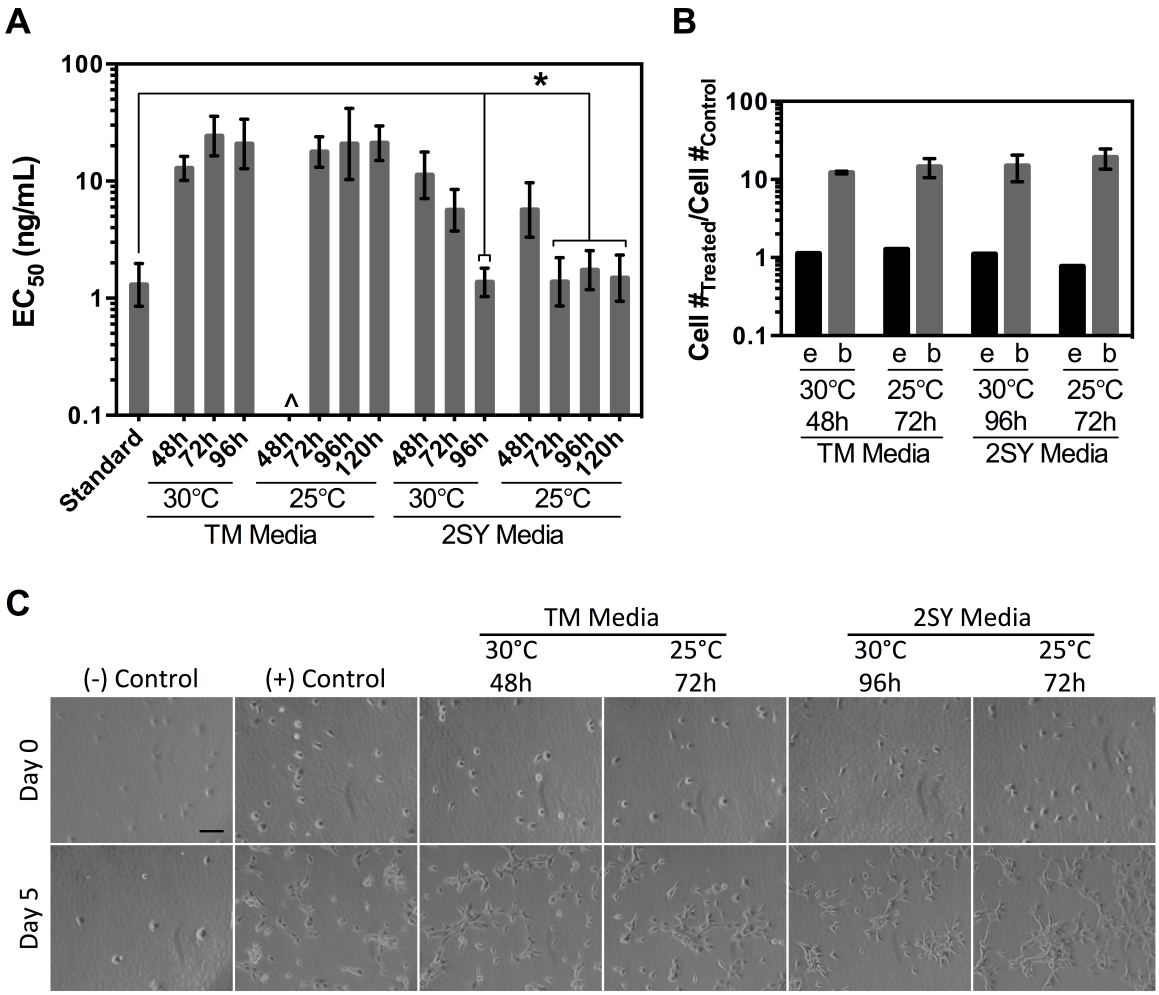


Figure C.8: Bioactivity of BDNF. BDNF specific activity was measured by proliferation assay. (A) EC_{50} values measured for BDNF prepared from three replicate *B. choshinensis* cultures for each expression condition and fit with a single dose-response curve; (^) EC_{50} was not reached at the highest concentration tested (1% v/v); error bars indicate ± 1 SD of the dose-response curve fit; (*) bioactivity is comparable to commercially-available BDNF by equivalence test using 90% confidence intervals (Figure C.10). (B) Comparison of bioactivity of protein purified from cultures transfected with an empty vector (e) or the BDNF vector (b). (C) Purified protein samples were added at 1% v/v to serum-free defined medium. Recombinant BDNF at 100 ng/mL and wells without BDNF were used as (+) and (-) controls, respectively. Shown are representative images from days 1 and 5 of the same location within each well. Images are shown for cells exposed to protein purified from the earliest *B. choshinensis* cultures at which maximal protein bioactivity was measured for the given media and temperature conditions. Scale bar represents 100 μm. Images from all days are included in Figure C.11.

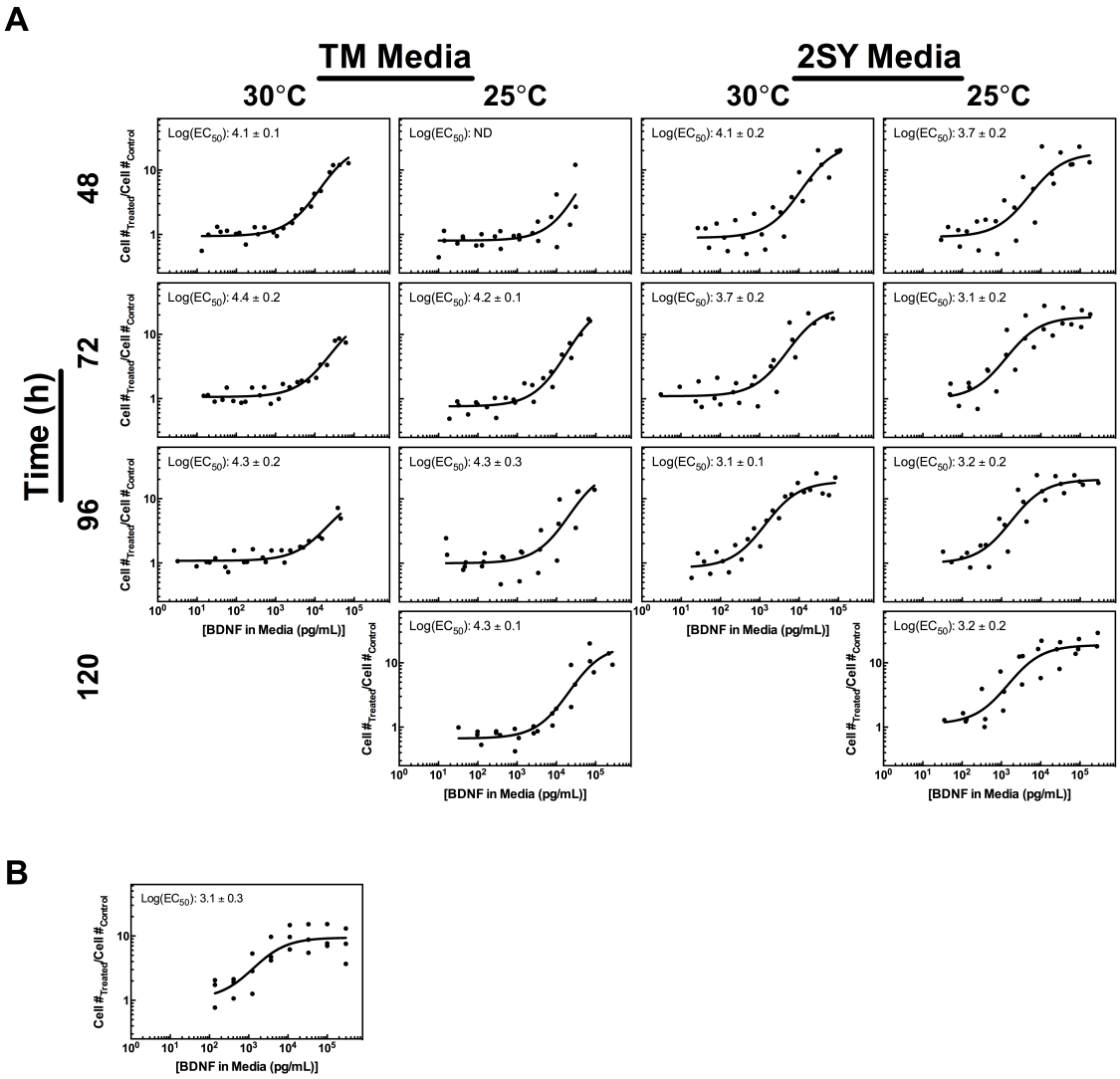


Figure C.9: BDNF EC_{50} Curves. (A) EC_{50} curves were fit to proliferation data from three replicate cultures of 3T3-TrkB cells, each treated with BDNF prepared in a separate *B. choshinensis* culture. BDNF protein concentrations were determined by western blotting (Table C.1). (B) Bioactivity of commercial BDNF fit to a single dose-response curve from three replicates. Error bars indicate ± 1 SD of the dose-response curve fit.

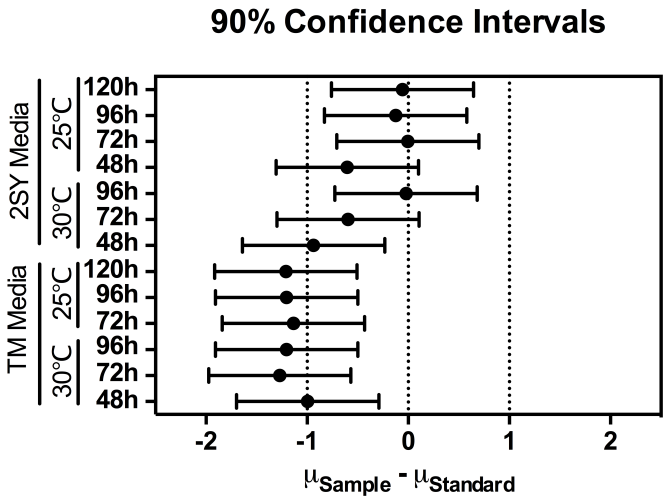


Figure C.10: Confidence Intervals Comparing BDNF Samples to Bioactive BDNF. BDNF bioactivity was compared to commercial BDNF using an equivalence test with 90% confidence intervals, calculated using 1-way ANOVA with Dunnett's *post hoc* analysis. Samples whose error bars are wholly contained between -1 and +1 are considered equivalent to commercially-available BDNF.

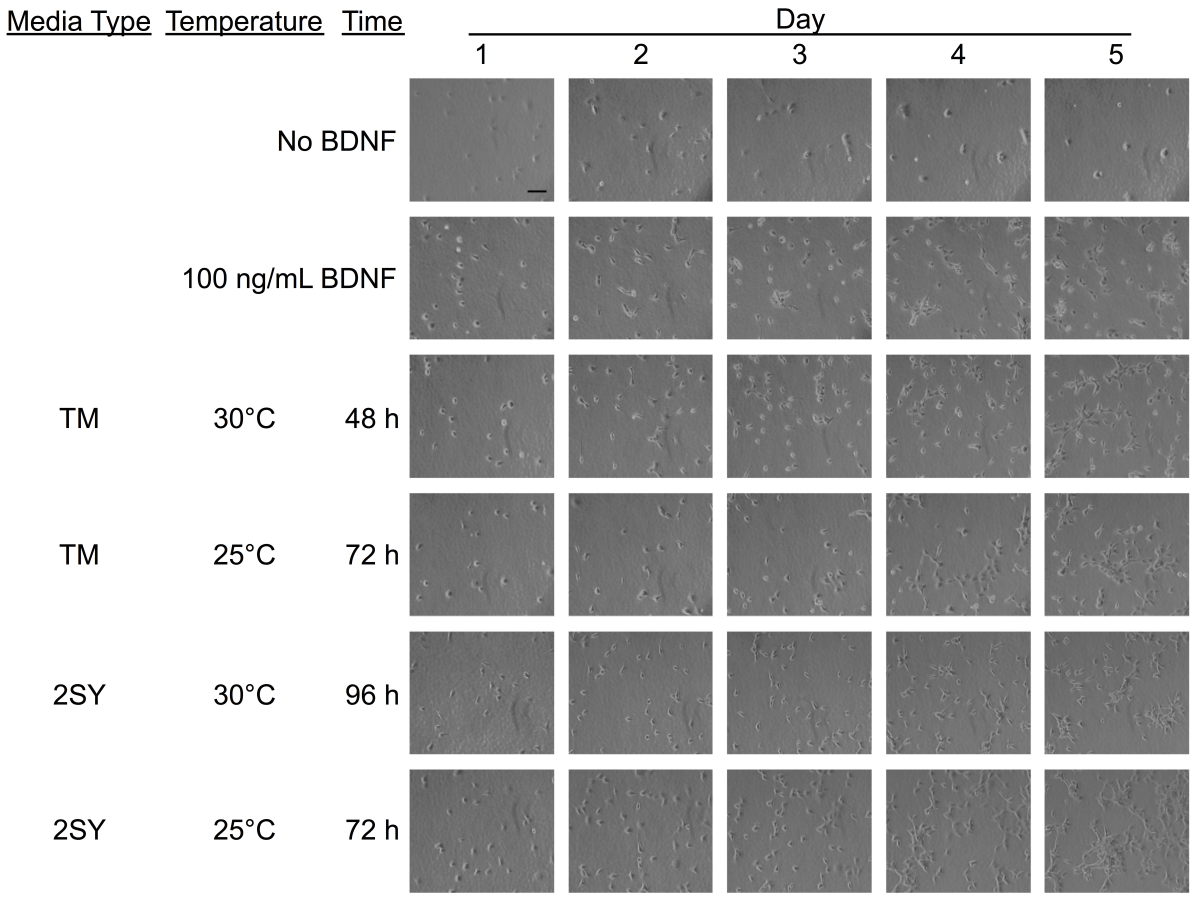


Figure C.11: BDNF Bioactivity Images. 3T3-TrkB cells were treated with BDNF three hours after plating (Day 1) and, subsequently, every 24 h thereafter (Days 2-5). Purified protein from each culture condition was added at 1% v/v to the serum-free defined medium. Recombinant BDNF at 100 ng/mL and no BDNF were used as positive and negative controls, respectively. Three images were taken in each well at each time point; only one set of images are shown. Images are shown for cells exposed to protein purified from the earliest *B. choshinensis* cultures at which maximal protein bioactivity was measured for the given media and temperature conditions.

Changes in bioactivity with culture time were investigated for each media type, culturing for a minimum of 48 h and up to 120 h at 25°C or 96 h at 30°C. We began with 48 h cultures as these were the earliest cultures for which BDNF levels were maximal under some conditions (Figure C.3). We took samples every 24 h until at least 3 different time-points were taken after BDNF levels had reached a maximum, resulting in the longer sampling period for cultures at 25°C relative to 30°C. BDNF bioactivity was compared to a bioactive,

commercial BDNF (Figure C.8A). Only protein obtained from cultures grown in 2SY media had comparable bioactivity to the commercial protein (Figure C.8A), indicating that, at least when using our current purification protocol, cultures grown in TM media produced a lower fraction of bioactive BDNF in the supernatants. Furthermore, in 2SY media we observed that EC_{50} values decreased with culture time (Figure C.8A). This suggests that, as with TM, not all of the protein produced in 2SY media is bioactive and that the fraction of bioactive protein present in the supernatant varies with time. We confirmed that bioactivity was due specifically to protein produced from the BDNF-containing vector, as shown by our empty vector controls (Figures C.8B and C.8C). Decreasing cell numbers with time in the empty-vector controls reflect some eventual cytotoxicity due to the absence of serum and BDNF.

Discussion

In this work, we demonstrated that BDNF purified from *B. choshinensis* cultures was biologically active. In testing a variety of culture conditions and times, we identified conditions that provided sufficient yields for subsequent bioactivity studies. Use of *B. choshinensis* as the host mitigated the principal difficulty in achieving expression of bioactive BDNF, its cystine-knot structure, which can lead to intermolecular disulfide linkages, incorrect intramolecular disulfide linkages, and aggregate formation (Binder and Scharfman, 2004; Burns et al., 2014; Meyer et al., 1994; Knusel et al., 1991). Nonetheless, despite our yields of bioactive protein, considerable amounts of BDNF produced still formed disulfide-bonded aggregates (Figure C.12). Additional optimization of expression conditions should focus on reducing the formation of aggregates, either during the expression process or after ac-

cumulation in the culture supernatant, which will likely have the concomitant benefits of improving recovery of BDNF and reducing the concentration of contaminants that co-purify with BDNF.

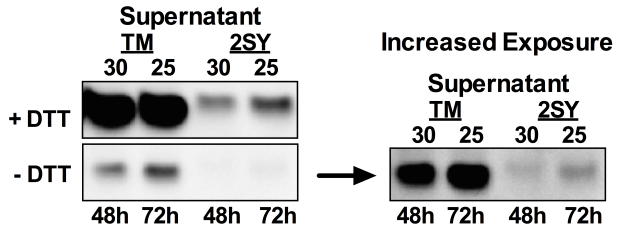


Figure C.12: Evidence of BDNF Aggregation. Representative BDNF western blot of the cell supernatant from each culture condition where samples were (+ DTT) or were not (- DTT) treated with DTT to evaluate the present of intermolecular disulfide bonds (aggregates). Shown are results from samples taken when protein concentrations reached a maximum for the given media and temperature conditions.

While aggregates were observed in all conditions, the quantity of non-aggregated BDNF in the supernatant was greater in TM cultures, suggesting that further optimization to minimize aggregation may be best focused on cultures using TM media. We hypothesize that the aggregation hindered accessibility of the His-tag and fouled the Ni^{2+} sepharose, both contributing to reduced yields of non-aggregated BDNF. However, total BDNF expression, including both the supernatant and cell pellet fractions, was comparable between both media types. This suggests that aggregates are less soluble in 2SY media than TM media (Figure C.3), resulting in a higher fraction of biologically active BDNF in 2SY supernatants.

A more efficient means of removing aggregates could be achieved by leveraging the stability of BDNF at pH values as low as 4 (Tanaka and Kumano, 2000); at these low pH conditions, aggregates should precipitate and can then be removed by centrifugation, similar to what has been done with EGF purification (Miyauchi et al., 1999). Non-aggregated BDNF could then be purified using a variety of other techniques, including IMAC, as we

performed here, or by cationic exchange, anionic exchange, or reversed-phase chromatographic methods, as have been studied elsewhere (Jungbluth et al., 1994; Burton, 1993; Rosenfeld and Benedek, 1993).

Reducing the culture temperature from 30°C to 25°C resulted in a greater fraction of biologically active BDNF at an earlier culture time in 2SY media (72 h vs 96 h, Figure C.8A). In other expression systems (i.e., *E. coli*), reductions in culture temperature are known to improve the expression of recombinant proteins, by reducing the rate of translation (Vasina and Baneyx, 1997) and the propensity for protein aggregation (Baldwin, 1986; Schellman, 1997). In 2SY media, the reduction in culture temperature also increased the stationary phase OD_{600} and decreased the rate of BDNF accumulation in the culture supernatant, indicating a marked reduction in the per cell expression of BDNF. While not explored explicitly here, the lower rate of BDNF expression (by reduction in temperature) does correspond with the ability to produce biologically active BDNF and could be explored more completely by use of a lower strength promoter (Onishi et al., 2013). Further reduction in temperature (20°C) resulted in significantly longer lag phase, especially in TM media, and was not explored further due to practical culture time limitations (Figure C.13). However, altering the culture conditions (i.e., beginning with a more concentrated inoculum) to reduce culture duration would allow investigation of expression at temperatures below 25°C and could be beneficial in the production and recovery of biologically active BDNF.

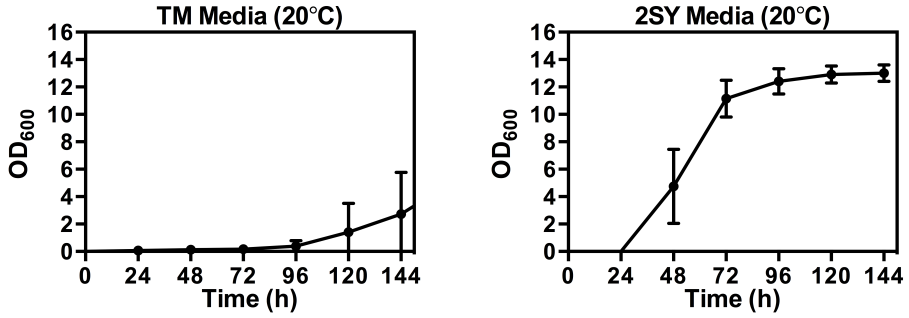


Figure C.13: *B. choshinensis* **Growth at 20°C.** *B. choshinensis* growth (OD_{600}) was examined at 20°C over 6 days in TM (A) and 2SY (B) media. N=3, error bars are ± 1 SD.

With the best conditions examined here (2SY media at 25°C for 72 h), we were able to purify 264 ± 82 μg/(L of culture) of biologically active BDNF. We consider these our best conditions as they resulted in: i) BDNF with the highest bioactivity, ii) the highest yield of BDNF, and iii) the minimum culture time. The quantity of recoverable BDNF did not significantly increase at times > 72 h in 2SY media at 25°C, but we did observe an increase in the variability of purified BDNF with time, with several cultures producing > 500 μg/L of biologically active BDNF. The yield of BDNF is greater than or of the same magnitude as other expression systems (Fukuzono et al., 1995; Meyer et al., 1994; Knusel et al., 1991; Rosenthal et al., 1991). In addition, the use of a microbial system offers both cost (< \$10/L media) and technical advantages (standard microbial culturing techniques) over higher complexity expression systems (Meyer et al., 1994; Knusel et al., 1991; Rosenthal et al., 1991) or methods that require refolding BDNF (Fukuzono et al., 1995). Recently, ∼1 mg/L expression of BDNF was demonstrated in yeast after the engineering of BDNF to facilitate protein folding (Burns et al., 2014, 2016); it is conceivable that using a modified BDNF sequence for expression in *B. choshinensis* could also significantly increase expression of biologically active BDNF in this host.

Conclusions

BDNF is a highly valuable protein, both for its potential clinical relevance and as a tool for researchers. The difficulties in producing bioactive BDNF protein significantly hinder its widespread use. Here, we have shown that *B. choshinensis* can produce 264 ± 82 μg/L of bioactive BDNF, when cultured at 25°C in 2SY media for 72 h. Similar to other expression systems, expression of BDNF in *B. choshinensis* produces aggregates that complicate the purification of BDNF by IMAC chromatography and form inactive protein. This work demonstrates the utility of the gram-positive host *B. choshinensis* to produce bioactive BDNF at useful scales and provides a starting point for additional optimization studies.

BIBLIOGRAPHY

BIBLIOGRAPHY

Acevedo, R., Orench-Rivera, N., Quarles, K. A., and Showalter, S. A. (2015). Helical defects in microRNA influence protein binding by TAR RNA binding protein. *PLoS One*, 10(1):e0116749.

Allerson, C. R., Sioufi, N., Jarres, R., Prakash, T. P., Naik, N., Berdeja, A., Wanders, L., Griffey, R. H., Swayze, E. E., and Bhat, B. (2005). Fully 2'-modified oligonucleotide duplexes with improved *in vitro* potency and stability compared to unmodified small interfering RNA. *Journal of Medicinal Chemistry*, 48(4):901–904.

Amarzguioui, M. (2003). Tolerance for mutations and chemical modifications in a siRNA. *Nucleic Acids Research*, 31(2):589–595.

Amarzguioui, M., Lundberg, P., Cantin, E., Hagstrom, J., Behlke, M. A., and Rossi, J. J. (2006). Rational design and *in vitro* and *in vivo* delivery of Dicer substrate siRNA. *Nature Protocols*, 1(2):508–517.

Amarzguioui, M. and Prydz, H. (2004). An algorithm for selection of functional siRNA sequences. *Biochemical and Biophysical Research Communications*, 316(4):1050–1058.

Ameres, S. L., Martinez, J., and Schroeder, R. (2007). Molecular basis for target RNA recognition and cleavage by human RISC. *Cell*, 130(1):101–112.

Ameres, S. L. and Zamore, P. D. (2013). Diversifying microRNA sequence and function. *Nature Reviews Molecular Cell Biology*, 14(8):475–488.

Angart, P., Vocelle, D., Chan, C., and Walton, S. P. (2013). Design of siRNA therapeutics from the molecular scale. *Pharmaceuticals*, 6(4):440–68.

Baldwin, R. L. (1986). Temperature dependence of the hydrophobic interaction in protein folding. *Proceedings of the National Academy of Sciences*, 83(21):8069–8072.

Bartel, D. (2009). MicroRNAs: Target recognition and regulatory functions. *Cell*, 136(2):215–233.

Beitzinger, M. and Meister, G. (2011). Experimental identification of microrna targets by immunoprecipitation of argonaute protein complexes. In Dalmay, T., editor, *MicroRNAs in Development*, volume 732, pages 153–67. Humana Press, Totowa, NJ.

Bernard, M. A., Wang, L., and Tachado, S. D. (2015). DICER-ARGONAUTE2 complex in continuous fluorogenic assays of RNA interference enzymes. *Plos One*, 10(3):e0120614.

Bernstein, E., Caudy, A. A., Hammond, S. M., and Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, 409(6818):363–366.

Betancur, J. G. and Tomari, Y. (2012). Dicer is dispensable for asymmetric RISC loading in mammals. *RNA*, 18(1):24–30.

Bevilacqua, P. C. and Cech, T. R. (1996). Minor-groove recognition of double-stranded RNA by the double-stranded RNA-binding domain from the RNA-activated protein kinase PKR. *Biochemistry*, 35(31):9983–9994.

Binder, D. K. and Scharfman, H. E. (2004). Brain-derived neurotrophic factor. *Growth Factors*, 22(3):123–131.

Birmingham, A., Anderson, E. M., Reynolds, A., Ilsley-Tyree, D., Leake, D., Fedorov, Y., Baskerville, S., Maksimova, E., Robinson, K., Karpilow, J., Marshall, W. S., and Khvorova, A. (2006). $3'$ UTR seed matches, but not overall identity, are associated with RNAi off-targets. *Nature Methods*, 3(3):199–204.

Bobbin, M. L. and Rossi, J. J. (2016). RNA interference (RNAi)-based therapeutics: Delivering on the promise? *Annual Review of Pharmacology and Toxicology*, 56(1):103–122.

Bohula, E. A., Salisbury, A. J., Sohail, M., Playford, M. P., Riedemann, J., Southern, E. M., and Macaulay, V. M. (2003). The efficacy of small interfering RNAs targeted to the type 1 insulin-like growth factor receptor (IGF1R) is influenced by secondary structure in the IGF1R transcript. *Journal of Biological Chemistry*, 278(18):15991–15997.

Braasch, D. A., Jensen, S., Liu, Y., Kaur, K., Arar, K., White, M. A., and Corey, D. R. (2003). RNA interference in mammalian cells by chemically-modified RNA. *Biochemistry*, 42(26):7967–7975.

Bramsen, J. B., Laursen, M. B., Damgaard, C. K., Lena, S. W., Babu, B. R., Wengel, J., and Kjems, J. (2007). Improved silencing properties using small internally segmented interfering RNAs. *Nucleic Acids Research*, 35(17):5886–5897.

Bramsen, J. B., Laursen, M. B., Nielsen, A. F., Hansen, T. B., Bus, C., Langkjaer, N., Babu, B. R., Hojland, T., Abramov, M., Van Aerschot, A., Odadzic, D., Smicius, R., Haas, J., Andree, C., Barman, J., Wenska, M., Srivastava, P., Zhou, C., Honcharenko, D., Hess, S., Muller, E., Bobkov, G. V., Mikhailov, S. N., Fava, E., Meyer, T. F., Chattopadhyaya, J., Zerial, M., Engels, J. W., Herdewijn, P., Wengel, J., and Kjems, J. (2009). A large-scale chemical modification screen identifies design rules to generate siRNAs with high activity, high stability and low toxicity. *Nucleic Acids Research*, 37(9):2867–2881.

Broderick, J. A., Salomon, W. E., Ryder, S. P., Aronin, N., and Zamore, P. D. (2011). Argonaute protein identity and pairing geometry determine cooperativity in mammalian RNA silencing. *RNA*, 17(10):1858–1869.

Brown, K. M., Chu, C. Y., and Rana, T. M. (2005). Target accessibility dictates the potency of human RISC. *Nature Structural & Molecular Biology*, 12(5):469–470.

Burns, M. L., Malott, T. M., Metcalf, K. J., Hackel, B. J., Chan, J. R., and Shusta, E. V. (2014). Directed evolution of brain-derived neurotrophic factor for improved folding and expression in *Saccharomyces cerevisiae*. *Applied Environmental Microbiology*, 80(18):5732–5742.

Burns, M. L., Malott, T. M., Metcalf, K. J., Puguh, A., Chan, J. R., and Shusta, E. V. (2016). Pro-region engineering for improved yeast display and secretion of brain derived neurotrophic factor. *Biotechnology Journal*, 11(3):425–36.

Burroughs, A. M., Ando, Y., de Hoon, M. J. L., Tomaru, Y., Suzuki, H., Hayashizaki, Y., and Daub, C. O. (2011). Deep-sequencing of human Argonaute-associated small RNAs provides insight into miRNA sorting and reveals Argonaute association with RNA fragments of diverse origin. *RNA Biology*, 8(1):158–177.

Burton, L. E. (1993). Production of recombinant neurotrophic factors. In A Bolton, G Baker, F. H., editor, *Neurotrophic Factors*, volume 25, pages 155–176. Springer Science.

Capek, M., Janácek, J., and Kubínová, L. (2006). Methods for compensation of the light attenuation with depth of images captured by a confocal microscope. *Microscopy Research and Technique*, 69(8):624–635.

Carthew, R. W. and Sontheimer, E. J. (2009). Origins and mechanisms of miRNAs and siRNAs. *Cell*, 136(4):642–655.

Cekaite, L., Furset, G., Hovig, E., and Sioud, M. (2007). Gene expression analysis in blood cells in response to unmodified and $2'$-modified siRNAs reveals TLR-dependent and independent effects. *Journal of Molecular Biology*, 365(1):90–108.

Chandradoss, S. D., Schirle, N. T., Szczepaniak, M., MacRae, I. J., and Joo, C. (2015). A dynamic search process underlies MicroRNA targeting. *Cell*, 162(1):96–107.

Chang, C. I., Yoo, J. W., Hong, S. W., Lee, S. E., Kang, H. S., Sun, X., Rogoff, H. A., Ban, C., Kim, S., Li, C. J., and Lee, D. (2009). Asymmetric shorter-duplex siRNA structures trigger efficient gene silencing with reduced nonspecific effects. *Molecular Therapy*, 17(4):725–732.

Cheloufi, S., Santos, C. O. D., Chong, M. M. W., and Hannon, G. J. (2010). A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature*, 465(7298):584–589.

Chen, C., Zhu, C., Huang, J., Zhao, X., Deng, R., Zhang, H., Dou, J., Chen, Q., Xu, M., Yuan, H., Wang, Y., and Yu, J. (2015). SUMOylation of TARBP2 regulates miRNA/siRNA efficiency. *Nature Communications*, 6:8899.

Chen, C. F., Ridzon, D. A., Broomer, A. J., Zhou, Z. H., Lee, D. H., Nguyen, J. T., Barbisin, M., Xu, N. L., Mahuvakar, V. R., Andersen, M. R., Lao, K. Q., Livak, K. J., and Guegler, K. J. (2005). Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Research*, 33(20):e179.

Chendrimada, T., Gregory, R., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., and Shiekhattar, R. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature*, 436(7051):740–744.

Chiu, Y. L. and Rana, T. M. (2003). siRNA function in RNAi: A chemical modification analysis. *RNA*, 9(9):1034–1048.

Chu, C. and Rana, T. M. (2008). Potent RNAi by short RNA triggers. *RNA*, 14(9):1714–1719.

Collingwood, M. A., Rose, S. D., Huang, L., Hillier, C., Amarzguioui, M., Wiiger, M. T., Soifer, H. S., Rossi, J. J., and Behlke, M. A. (2008). Chemical modification patterns compatible with high potency Dicer-substrate small interfering RNAs. *Oligonucleotides*, 18(2):187–200.

Czech, B. and Hannon, G. J. (2010). Small RNA sorting: Matchmaking for Argonautes. *Nature Reviews Genetics*, 12(1):19–31.

Daniels, S. M. and Gatignol, A. (2012). The multiple functions of TRBP, at the hub of cell responses to viruses, stress, and cancer. *Microbiology and Molecular Biology Reviews*, 76(3):652–66.

Daniels, S. M., Melendez-Pena, C. E., Scarborough, R. J., Daher, A., Christensen, H. S., El Far, M., Purcell, D. F. J., Laine, S., and Gatignol, A. (2009). Characterization of the TRBP domain required for dicer interaction and function in RNA interference. *BMC Molecular Biology*, 10(1):38.

Daviet, L., Erard, M., Dorin, D., Duarte, M., Vaquero, C., and Gatignol, A. (2000). Analysis of a binding difference between the two dsRNA-binding domains in TRBP reveals the modular function of a KR-helix motif. *European Journal of Biochemistry*, 267(8):2419–2431.

De, N., Young, L., Lau, P. W., Meisner, N. C., Morrissey, D. V., and MacRae, I. J. (2013). Highly complementary target RNAs promote release of guide RNAs from human Argonaute2. *Molecular Cell*, 50(3):344–355.

Didiano, D. and Hobert, O. (2006). Perfect seed pairing is not a generally reliable predictor for miRNA-target interactions. *Nature Structural & Molecular Biology*, 13(9):849–851.

Diebold, S. S., Kaisho, T., Hemmi, H., Akira, S., and Reis e Sousa, C. (2004). Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA. *Science*, 303(5663):1529–1531.

Diebold, S. S., Massacrier, C., Akira, S., Paturel, C., Morel, Y., and Reis e Sousa, C. (2006). Nucleic acid agonists for Toll-like receptor 7 are defined by the presence of uridine ribonucleotides. *European Journal of Immunology*, 36(12):3256–3267.

Doench, J. G., Petersen, C. P., and Sharp, P. A. (2003). siRNAs can function as miRNAs. *Genes & Development*, 17(4):438–442.

Doench, J. G. and Sharp, P. A. (2004). Specificity of microRNA target selection in translational repression. *Genes & Development*, 18(5):504–511.

Dua, P., Yoo, J. W., Kim, S., and Lee, D. K. (2011). Modified siRNA structure with a single nucleotide bulge overcomes conventional siRNA-mediated off-target silencing. *Molecular Therapy*, 19(9):1676–1687.

Dueck, A., Ziegler, C., Eichner, A., Berezikov, E., and Meister, G. (2012). microRNAs associated with the different human Argonaute proteins. *Nucleic Acids Research*, 40(19):9850–9862.

Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001a). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, 411(6836):494–498.

Elbashir, S. M., Lendeckel, W., and Tuschl, T. (2001b). RNA interference is mediated by 21-and 22-nucleotide RNAs. *Genes & Development*, 15(2):188–200.

Elbashir, S. M., Martinez, J., Patkaniowska, A., Lendeckel, W., and Tuschl, T. (2001c). Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. *The EMBO journal*, 20(23):6877–6888.

Fedorov, Y., Anderson, E. M., Birmingham, A., Reynolds, A., Karpilow, J., Robinson, K., Leake, D., Marshall, W. S., and Khvorova, A. (2006). Off-target effects by siRNA can induce toxic phenotype. *RNA*, 12(7):1188–1196.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis* elegans. *Nature*, 391(6669):806–811.

Forsbach, A., Nemorin, J.-G., Montino, C., Müller, C., Samulowitz, U., Vicari, A. P., Jurk, M., Mutwiri, G. K., Krieg, A. M., Lipford, G. B., and Vollmer, J. (2008). Identification of RNA sequence motifs stimulating sequence-specific TLR8-dependent immune responses. *Journal of Immunology*, 180(6):3729–3738.

Foster, D. J., Barros, S., Duncan, R., Shaikh, S., Cantley, W., Dell, A., Bulgakova, E., O'Shea, J., Taneja, N., Kuchimanchi, S., Sherrill, C. B., Akinc, A., Hinkle, G., White, A. C. S., Pang, B., Charisse, K., Meyers, R., Manoharan, M., and Elbashir, S. M. (2012). Comprehensive evaluation of canonical versus Dicer-substrate siRNA *in vitro* and *in vivo*. *RNA*, 18(3):557–568.

Frank, F., Sonenberg, N., and Nagar, B. (2010). Structural basis for 5′-nucleotide base-specific recognition of guide RNA by human AGO2. *Nature*, 465(7299):818–822.

Friedman, R. C., Farh, K. K., Burge, C. B., and Bartel, D. P. (2008). Most mammalian mRNAs are conserved targets of microRNAs. *Genome Research*, 19(1):92–105.

Fucini, R. V., Haringsma, H. J., Deng, P., Flanagan, W. M., and Willingham, A. T. (2012). Adenosine modification may be preferred for reducing siRNA immune stimulation. *Nucleic Acid Therapeutics*, 22(3):205–210.

Fukuzono, S., Fujimori, K., and Shimizu, N. (1995). Production of biologically active mature brain-derived neurotrophic factor in *Escherichia coli*. *Bioscience, Biotechnology, and Biochemistry*, 59(9):1727–1731.

Gantier, M. P. and Williams, B. R. G. (2007). The response of mammalian cells to double-stranded RNA. *Cytokine & Growth Factor Reviews*, 18(5-6):363–371.

Gong, W., Ren, Y., Xu, Q., Wang, Y., Lin, D., Zhou, H., and Li, T. (2006). Integrated siRNA design based on surveying of features associated with high RNAi effectiveness. *BMC Bioinformatics*, 7:516.

Goodchild, A., Nopper, N., King, A., Doan, T., Tanudji, M., Arndt, G. M., Poidinger, M., Rivory, L. P., and Passioura, T. (2009). Sequence determinants of innate immune activation by short interfering RNAs. *BMC Immunology*, 10(40).

Gredell, J. A., Dittmer, M. J., Wu, M., Chan, C., and Walton, S. P. (2010). Recognition of siRNA asymmetry by TAR RNA binding protein. *Biochemistry*, 49(14):3148–3155.

Gregory, R. I., ping Yan, K., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The microprocessor complex mediates the genesis of microRNAs. *Nature*, 432(7014):235–240.

Grimm, D., Streetz, K. L., Jopling, C. L., Storm, T. A., Pandey, K., Davis, C. R., Marion, P., Salazar, F., and Kay, M. A. (2006). Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature*, 441(7092):537–541.

Gu, S., Jin, L., Zhang, F., Huang, Y., Grimm, D., Rossi, J. J., and Kay, M. A. (2011). Thermodynamic stability of small hairpin RNAs highly influences the loading process of different mammalian Argonautes. *Proceedings of the National Academy of Sciences*, 108(22):9208–9213.

Ha, M. and Kim, N. (2014). Regulation of microRNA biogenesis. *Nature Reviews Molecular Cell Biology*, 15:509–524.

Haase, A., Jaskiewicz, L., Zhang, H., Laine, S., Sack, R., Gatignol, A., and Filipowicz, W. (2005). TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. *EMBO Reports*, 6(10):961–967.

Haley, B. and Zamore, P. D. (2004). Kinetic analysis of the RNAi enzyme complex. *Nature Structural & Molecular Biology*, 11(7):599–606.

Han, J., Lee, Y., Yeom, K., Kim, Y., Jin, H., and Kim, V. N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. *Genes & Development*, 18(24):3016–3027.

Hanagata, H. and Nishijyo, T. (2010). *Brevibacillus choshinensis* and process for prodcuing protein with use of the microbe as host. US Patent 7,655,452.

Haringsma, H. J., Li, J. J., Soriano, F., Kenski, D. M., Flanagan, W. M., and Willingham, A. T. (2012). mRNA knockdown by single strand RNA is improved by chemical modifications. *Nucleic Acids Research*, 40(9):4125–4136.

Haupenthal, J., Baehr, C., Kiermayer, S., Zeuzem, S., and Piiper, A. (2006). Inhibition of RNAse A family enzymes prevents degradation and loss of silencing activity of siRNAs in serum. *Biochemical Pharmacology*, 71(5):702–710.

Haussecker, D. (2012). The business of RNAi therapeutics in 2012. *Molecular Therapy — Nucleic Acids*, 1(2):e8.

Heil, F., Hemmi, H., Hochrein, H., Ampenberger, F., Kirschning, C., Akira, S., Lipford, G., Wagner, H., and Bauer, S. (2004). Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8. *Science*, 303(5663):1526–1529.

Heymach, J. V., J. and Shooter, E. M. (1995). The biosynthesis of neurotrophin heterodimers by transfected mammalian cells. *Journal of Biological Chemistry*, 270(20):12297–304.

Hibio, N., Hino, K., Shimizu, E., Nagata, Y., and Ui-Tei, K. (2012). Stability of miRNA 5' terminal and seed regions is correlated with experimentally observed miRNA-mediated silencing efficacy. *Scientific Reports*, 2(996).

Hohjoh, H. (2004). Enhancement of RNAi activity by improved siRNA duplexes. *FEBS letters*, 557(1-3):193–198.

Holen, T. (2006). Efficient prediction of siRNAs with siRNArules 1.0: An open-source JAVA approach to siRNA algorithms. *RNA*, 12(9):1620–1625.

Holen, T., Amarzguioui, M., Babaie, E., and Prydz, H. (2003). Similar behaviour of single-strand and double-strand siRNAs suggests they act through a common RNAi pathway. *Nucleic Acids Research*, 31(9):2401–2407.

Hong, J., Huang, Y., Li, J., Yi, F., Zheng, J., Huang, H., Wei, N., Shan, Y., An, M., Zhang, H., Ji, J., Zhang, P., Xi, Z., Du, Q., and Liang, Z. (2010). Comprehensive analysis of sequence-specific stability of siRNA. *The FASEB Journal*, 24(12):4844–4855.

Hornung, V., Guenthner-Biller, M., Bourquin, C., Ablasser, A., Schlee, M., Uematsu, S., Noronha, A., Manoharan, M., Akira, S., de Fougerolles, A., Endres, S., and Hartmann, G. (2005). Sequence-specific potent induction of IFN-α by short interfering RNA in plasmacytoid dendritic cells through TLR7. *Nature Medicine*, 11(3):263–270.

Hoshino, K., Eda, A., Kurokawa, Y., and Shimizu, N. (2002). Production of brain-derived neurotrophic factor in *Escherichia* coli by coexpression of Dsb proteins. *Bioscience, Biotechnology, and Biochemistry*, 66(2):344–350.

Hossbach, M., Gruber, J., Osborn, M., Weber, K., and Tuschl, T. (2006). Gene silencing with siRNA duplexes composed of target-mRNA-complementary and partially palindromic or partially complementary single-stranded siRNAs. *RNA Biology*, 3(2):82–89.

Huesken, D., Lange, J., Mickanin, C., Weiler, J., Asselbergs, F., Warner, J., Meloon, B., Engel, S., Rosenberg, A., Cohen, D., Labow, M., Reinhardt, M., Natt, F., and Hall, J. (2005). Design of a genome-wide siRNA library using an artificial neural network. *Nature Biotechnology*, 23(8):995–1001.

Ibáñez, C. F. and Simi, A. (2012). p75 neurotrophin receptor signaling in nervous system injury and degeneration: Paradox and opportunity. *Trends in Neurosciences*, 35(7):431–440.

Ilk, N., Schumi, C.-T., Bohle, B., Egelseer, E. M., and Sleytr, U. B. (2011). Expression of an endotoxin-free S-layer/allergen fusion protein in gram-positive *Bacillus subtilis* 1012 for the potential application as vaccines for immunotherapy of atopic allergy. *Microbial Cell Factories*, 10:6.

Iwasaki, S., Kobayashi, M., Yoda, M., Sakaguchi, Y., Katsuma, S., Suzuki, T., and Tomari, Y. (2010). Hsc70/Hsp90 chaperone machinery mediates ATP-dependent RISC loading of small RNA duplexes. *Molecular Cell*, 39(2):292–299.

Jackson, A. L., Bartz, S. R., Schelter, J., Kobayashi, S. V., Burchard, J., Mao, M., Li, B., Cavet, G., and Linsley, P. S. (2003). Expression profiling reveals off-target gene regulation by RNAi. *Nature Biotechnology*, 21(6):635–637.

Jackson, A. L., Burchard, J., Leake, D., Reynolds, A., Schelter, J., Guo, J., Johnson, J. M., Lim, L., Karpilow, J., Nichols, K., Marshall, W., Khvorova, A., and Linsley, P. S. (2006a). Position-specific chemical modification of siRNAs reduces "off-target" transcript silencing. *RNA*, 12(7):1197–1205.

Jackson, A. L., Burchard, J., Schelter, J., Chau, B. N., Cleary, M., Lim, L., and Linsley, P. S. (2006b). Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. *RNA*, 12(7):1179–1187.

Jackson, A. L. and Linsley, P. S. (2010). Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. *Nature Reviews Drug Discovery*, 9(1):57–67.

Jagla, B., Aulner, N., Kelly, P. D., Song, D., Volchuk, A., Zatorski, A., Shum, D., Mayer, T., De Angelis, D. A., Ouerfelli, O., Rutishauser, U., and Rothman, J. E. (2005). Sequence characteristics of functional siRNAs. *RNA*, 11(6):864–872.

Judge, A. D., Sood, V., Shaw, J. R., Fang, D., McClintock, K., and MacLachlan, I. (2005). Sequence-dependent stimulation of the mammalian innate immune response by synthetic siRNA. *Nature Biotechnology*, 23(4):457–462.

Jungbluth, S., Bailey, K., and Barde, Y. A. (1994). Purification and characterisation of a brain-derived neurotrophic factor/neurotrophin-3 (BDNF/NT-3) heterodimer. *European Journal of Biochemistry*, 221(2):677–685.

Kariko, K., Bhuyan, P., Capodici, J., and Weissman, D. (2004). Small interfering RNAs mediate sequence-independent gene suppression and induce immune activation by signaling through toll-like receptor 3. *Journal of Immunology*, 172(11):6545–6549.

Kato, H., Takeuchi, O., Mikamo-Satoh, E., Hirai, R., Kawai, T., Matsushita, K., Hiiragi, A., Dermody, T. S., Fujita, T., and Akira, S. (2008). Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid-inducible gene-I and melanoma differentiation-associated gene 5. *Journal of Experimental Medicine*, 205(7):1601–1610.

Katoh, T. and Suzuki, T. (2007). Specific residues at every third position of siRNA shape its efficient RNAi activity. *Nucleic Acids Research*, 35(4):e27.

Kawamata, T., Seitz, H., and Tomari, Y. (2009). Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. *Nature*, 16(9):953–960.

Kawamata, T., Yoda, M., and Tomari, Y. (2011). Multilayer checkpoints for microRNA authenticity during RISC assembly. *EMBO Reports*, 12(9):944–949.

Kenski, D. M., Butora, G., Willingham, A. T., Cooper, A. J., Fu, W., Qi, N., Soriano, F., Davies, I. W., and Flanagan, W. M. (2012). siRNA-optimized modifications for enhanced *in vivo* activity. *Molecular Therapy - Nucleic Acids*, 1:e5.

Khvorova, A., Reynolds, A., and Jayasena, S. D. (2003). Functional siRNAs and miRNAs exhibit strand bias. *Cell*, 115(2):209–216.

Kim, Y., Yeo, J., Lee, J. H., Cho, J., Seo, D., Kim, J. S., and Kim, V. N. (2014). Deletion of human tarbp2 reveals cellular MicroRNA targets and cell-cycle function of TRBP. *Cell Reports*, 9(3):1061–1074.

Kini, H. and Walton, S. P. (2007). *In vitro* binding of single-stranded RNA by human Dicer. *FEBS Letters*, 581(29):5611–6.

Kini, H. and Walton, S. P. (2009). Effect of siRNA terminal mismatches on TRBP and Dicer binding and silencing efficacy. *FEBS Journal*, 276(22):6576–85.

Kleinman, M. E., Yamada, K., Takeda, A., Chandrasekaran, V., Nozaki, M., Baffi, J. Z., Albuquerque, R. J. C., Yamasaki, S., Itaya, M., Pan, Y., Appukuttan, B., Gibbs, D., Yang, Z., Karikó, K., Ambati, B. K., Wilgus, T. A., DiPietro, L. A., Sakurai, E., Zhang, K., Smith, J. R., Taylor, E. W., and Ambati, J. (2008). Sequence- and target-independent angiogenesis suppression by siRNA via TLR3. *Nature*, 452(7187):591–597.

Knusel, B., Winslow, J. W., Rosenthal, A., Burton, L. E., Seid, D. P., Nikolics, K., and Hefti, F. (1991). Promotion of central cholinergic and dopaminergic neuron differentiation by brain-derived neurotrophic factor but not neurotrophin 3. *Proceedings of the National Academy of Sciences*, 88(3):961–5.

Köberle, C., Kaufmann, S. H. E., and Patzel, V. (2006). Selecting effective siRNAs based on guide RNA structure. *Nature Protocols*, 1(4):1832–1839.

Kodym, R., Kodym, E., and Story, M. D. (2009). 2'-5'-Oligoadenylate synthetase is activated by a specific RNA sequence motif. *Biochemical and Biophysical Research Communications*, 388(2):317–322.

Koh, H. R., Kidwell, M. A., Ragunathan, K., Doudna, J. A., and Myong, S. (2013). ATP-independent diffusion of double-stranded RNA binding proteins. *Proceedings of the National Academy of Sciences*, 110(1):151–156.

Kok, K., Ng, M., and Ching, Y. (2007). Human TRBP and PACT directly interact with each other and associate with dicer to facilitate the production of small interfering RNA. *Journal of Biological Chemistry*.

Kok, K. H., Lui, P. Y., Ng, M. H. J., Siu, K. L., Au, S. W. N., and Jin, D. Y. (2011). The double-stranded RNA-binding protein PACT functions as a cellular activator of RIG-I to facilitate innate antiviral response. *Cell Host Microbe*, 9(4):299–309.

Koller, E., Propp, S., Murray, H., Lima, W., Bhat, B., Prakash, T. P., Allerson, C. R., Swayze, E. E., Marcusson, E. G., and Dean, N. M. (2006). Competition for RISC binding predicts in vitro potency of siRNA. *Nucleic Acids Research*, 34(16):4467–4476.

Kozomara, A. and Griffiths-Jones, S. (2011). miRBase: Integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Research*, 39(Database issue):D152–7.

Kramer, M. F. (2001). Stem-loop RT-qPCR for miRNAs. In *Current Protocols in Molecular Biology*, volume 95, pages 15.10:15.10.1–15.10.15. John Wiley & Sons, Inc.

Kwak, P. B. and Tomari, Y. (2012). The N domain of Argonaute drives duplex unwinding during RISC assembly. *Nature*, 19(2):145–151.

Ladunga, I. (2006). More complete gene silencing by fewer siRNAs: Transparent optimized design and biophysical signature. *Nucleic Acids Research*, 35(2):433–440.

Lagos-Quintana, M. (2001). Identification of novel genes coding for small expressed RNAs. *Science*, 294(5543):853–858.

Lai, E. C. (2002). Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. *Nature Genetics*, 30(4):363–364.

Lambert, N. J., Gu, S. G., and Zahler, A. M. (2011). The conformation of microRNA seed regions in native microRNPs is prearranged for presentation to mRNA targets. *Nucleic Acids Research*, 39(11):4827–4835.

Laraki, G., Clerzius, G., Daher, A., and Melendez-Peña, C. (2008). Interactions between the double-stranded RNA-binding proteins TRBP and PACT define the Medipal domain that mediates protein-protein interactions. *RNA Biology*, 29(1):254–265.

Lee, H. Y. and Doudna, J. A. (2012). TRBP alters human precursor microRNA processing *in vitro*. *RNA*, 18(11):2012–9.

Lee, H. Y., Zhou, K., Smith, A. M., Noland, C. L., and Doudna, J. A. (2013). Differential roles of human Dicer-binding proteins TRBP and PACT in small RNA processing. *Nucleic Acids Research*, 41(13):6568–6576.

Lee, J. Y., Kim, H., Ryu, C. H., Kim, J. Y., Choi, B. H., Lim, Y., Huh, P. W., Kim, Y. H., Lee, K. H., Jun, T. Y., Rha, H. K., Kang, J. K., and Choi, C. R. (2004). Merlin, a tumor suppressor, interacts with transactivation-responsive RNA-binding protein and inhibits its oncogenic activity. *Journal of Biological Chemistry*, 279(29):30265–30273.

Lee, J. Y., Moon, H. J., Lee, W. K., Chun, H. J., Han, C. W., Jeon, Y. W., Lim, Y., Kim, Y. H., Yao, T. P., Lee, K. H., Jun, T. Y., Rha, H. K., and Kang, J. K. (2005). Merlin facilitates ubiquitination and degradation of transactivation-responsive RNA-binding protein. *Oncogene*, 25(8):1143–1152.

Lee, R., Kermani, P., Teng, K. K., and Hempstead, B. L. (2001). Regulation of cell survival by secreted proneurotrophins. *Science*, 294(5548):1945–8.

Leibrock, J., Lottspeich, F., Hohn, A., Hofer, M., Hengerer, B., Masiakowski, P., Thoenen, H., and Barde, Y. (1989). Molecular cloning and expression of brain-derived neurotrophic factor. *Nature*, 341(6238):149–152.

Leuschner, P. J. F., Ameres, S. L., Kueng, S., and Martinez, J. (2006). Cleavage of the siRNA passenger strand during RISC assembly in human cells. *EMBO Reports*, 7(3):314–320.

Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P., and Burge, C. B. (2003). Prediction of mammalian microRNA targets. *Cell*, 115(7):787–798.

Lima, W. F., Murray, H., Nichols, J. G., Wu, H., Sun, H., Prakash, T. P., Berdeja, A. R., Gaus, H. J., and Crooke, S. T. (2009). Human Dicer binds short single-strand and double-strand RNA with high affinity and interacts with different regions of the nucleic acids. *Journal of Biological Chemistry*, 284(4):2535–2548.

Lima, W. F., Prakash, T. P., Murray, H. M., Kinberger, G. A., Li, W., Chappell, A. E., Li, C. S., Murray, S. F., Gaus, H., Seth, P. P., Swayze, E. E., and Crooke, S. T. (2012). Single-stranded siRNAs activate RNAi in animals. *Cell*, 150(5):883–894.

Lin, X., Ruan, X., Anderson, M. G., Mcdowell, J. A., Kroeger, P. E., Fesik, S. W., and Shen, Y. (2005). siRNA-mediated off-target gene silencing triggered by a 7 nt complementation. *Nucleic Acids Research*, 33(14):4527.

Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2004). Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nature Structural & Molecular Biology*, 11(6):576–577.

Liu, J. D., Carmell, M. A., Rivas, F. V., Marsden, C. G., Thomson, J. M., Song, J. J., Hammond, S. M., Joshua-Tor, L., and Hannon, G. J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. *Science*, 305(5689):1437–1441.

Liu, Y., Ye, X., Jiang, F., Liang, C., Chen, D., Peng, J., Kinch, L. N., Grishin, N. V., and Liu, Q. (2009). C3PO, an endoribonuclease that promotes RNAi by facilitating RISC activation. *Science*, 325(5941):750–753.

Lu, B. (2003). Pro-region of neurotrophins: Role in synaptic modulation. *Neuron*, 39(5):735–8.

Lu, B., Pang, P. T., and Woo, N. H. (2005a). The yin and yang of neurotrophin action. *Nature Reviews Neuroscience*, 6(8):603–614.

Lu, P., Jones, L., and Tuszynski, M. H. (2005b). BDNF-expressing marrow stromal cells support extensive axonal growth at sites of spinal cord injury. *Experimental Neurology*, 191(2):344–360.

Lu, Z. J. and Mathews, D. H. (2008). Efficient siRNA selection using hybridization thermodynamics. *Nucleic Acids Research*, 36(2):640–647.

Lynam, D. A., Shahriari, D., Wolf, K. J., Angart, P. A., Koffler, J., Tuszynski, M. H., Chan, C., Walton, P., and Sakamoto, J. (2015). Brain derived neurotrophic factor release from layer-by-layer coated agarose nerve guidance scaffolds. *Acta Biomaterialia*, 18:128–131.

Ma, E., Macrae, I. J., Kirsch, J. F., and Doudna, J. A. (2008). Autoinhibition of human Dicer by its internal helicase domain. *Journal of Molecular Biology*, 380(1):237–243.

Ma, J. B., Ye, K., and Patel, D. J. (2004). Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature*, 429(6989):318–322.

MacRae, I. J., Ma, E., Zhou, M., Robinson, C. V., and Doudna, J. A. (2008). *In vitro* reconstitution of the human RISC-loading complex. *Proceedings of the National Academy of Sciences*, 105(2):512–517.

Macrae, I. J., Zhou, K., and Doudna, J. A. (2007). Structural determinants of rna recognition and cleavage by dicer. *Nature Structural & Molecular Biology*, 14(10):934–940.

Malefyt, A. P., Walton, S. P., and Chan, C. (2012). Endocytosis pathways for nucleic acid therapeutics. *Nano Life*, 2(3):1241005.

Malefyt, A. P., Wu, M., Vocelle, D. B., Kappes, S. J., Lindeman, S. D., Chan, C., and Walton, S. P. (2013). Improved asymmetry prediction for siRNAs. *FEBS Journal*, 281(1):320–330.

Manche, L., Green, S. R., Schmedt, C., and Mathews, M. B. (1992). Interactions between double-stranded RNA regulators and the protein kinase DAI. *Molecular and Cellular Biology*, 12(11):5238–5248.

Manoharan, M., Akinc, A., Pandey, R. K., Qin, J., Hadwiger, P., John, M., Mills, K., Charisse, K., Maier, M. A., Nechev, L., Greene, E. M., Pallan, P. S., Rozners, E., Rajeev, K. G., and Egli, M. (2011). Unique gene-silencing and structural properties of 2'-fluoro-modified siRNAs. *Angewandte Chemie*, 50(10):2284–2288.

Markham, N. R. and Zuker, M. (2005). DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Research*, 33(Web Server issue):W577–81.

Markham, N. R. and Zuker, M. (2008). UNAFold: Software for nucleic acid folding and hybridization. In Keith, J. M., editor, *Bioinformatics, Volume II: Structure, Function, and Applications*, volume 453 of *Methods in Molecular Biology*, pages 3–31. Humana Press, Totowa, NJ.

Marques, J. T., Devosse, T., Wang, D., Zamanian-Daryoush, M., Serbinowski, P., Hartmann, R., Fujita, T., Behlke, M. A., and Williams, B. R. (2006). A structural basis for discriminating between self and nonself double-stranded RNAs in mammalian cells. *Nature Biotechnology*, 24(5):559–565.

Martinez, J., Patkaniowska, A., Urlaub, H., Lührmann, R., and Tuschl, T. (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell*, 110(5):563–574.

Matranga, C., Tomari, Y., Shin, C., Bartel, D., and Zamore, P. (2005). Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell*, 123(4):607–620.

McCarty, J. H. and Feinstein, S. C. (1998). Activation loop tyrosines contribute varying roles to TrkB autophosphorylation and signal transduction. *Oncogene*, 16(13):1691–1700.

Meijering, E. H., Niessen, W. J., and Viergever, M. A. (2001). Quantitative evaluation of convolution-based methods for medical image interpolation. *Medical Image Analysis*, 5(2):111–126.

Meisenheimer, K. M. and Koch, T. H. (1997). Photocross-linking of nucleic acids to associated proteins. *Critical Reviews in Biochemistry and Molecular Biology*, 32(2):101–140.

Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. (2004). Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Molecular Cell*, 15(2):185–197.

Menei, P., Montero-Menei, C., Whittemore, S. R., Bunge, R. P., and Bunge, M. B. (1998). Schwann cells genetically modified to secrete human BDNF promote enhanced axonal regrowth across transected adult rat spinal cord. *European Journal of Neuroscience*, 10(2):607–621.

Meyer, M., Matsuoka, I., Wetmore, C., Olson, L., and Thoenen, H. (1992). Enhanced synthesis of brain-derived neurotrophic factor in the lesioned peripheral nerve: Different mechanisms are responsible for the regulation of BDNF and NGF mRNA. *Journal of Cell Biology*, 119(1):45–54.

Meyer, S. L., Lang, D. M., Forbes, M. E., Knight, Jr, E., Hirsch, J. D., Trusko, S. P., and Scott, R. W. (1994). Production and characterization of recombinant mouse brain-derived neurotrophic factor and rat neurotrophin-3 expressed in insect cells. *Journal of Neurochemistry*, 62(3):825–833.

Miyauchi, A., Ozawa, M., Mizukami, M., Yashiro, K., Ebisu, S., Tojo, T., Fujii, T., and Takagi, H. (1999). Structural conversion from non-native to native form of recombinant human epidermal growth factor by *Brevibacillus choshinensis*. *Bioscience, Biotechnology, and Biochemistry*, 63(11):1965–1969.

Mizukami, M., Hanagata, H., and Miyauchi, A. (2010). *Brevibacillus* expression system: Host-vector system for efficient production of secretory proteins. *Current Pharmaceutical Biotechnology*, 11(3):251–8.

Mowla, S. J., Farhadi, H. F., Pareek, S., Atwal, J. K., Morris, S. J., Seidah, N. G., and Murphy, R. A. (2001). Biosynthesis and post-translational processing of the precursor to brain-derived neurotrophic factor. *The Journal of Biological Chemistry*, 276(16):12660–12666.

Nagahara, A. H. and Tuszynski, M. H. (2011). Potential therapeutic uses of BDNF in neurological and psychiatric disorders. *Nature Reviews Drug Discovery*, 10(3):209–219.

Naito, Y. and Ui-Tei, K. (2012). siRNA design software for a target gene-specific RNA interference. *Frontiers in Genetics*, 3:102.

Napoli, C. and Lemieux, C. (1990). Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans. *The Plant Cell*, 2:279–289.

Navarro, F., Gutman, D., Meire, E., Caceres, M., Rigoutsos, I., Bentwich, Z., and Lieberman, J. (2009). miR-34a contributes to megakaryocytic differentiation of K562 cells independently of p53. *Blood*, 114(10):2181–2192.

Noland, C. L. and Doudna, J. A. (2013). Multiple sensors ensure guide strand selection in human RNAi pathways. *RNA*, 19(5):639–648.

Noland, C. L., Ma, E., and Doudna, J. A. (2011). siRNA repositioning for guide strand selection by human Dicer complexes. *Molecular Cell*, 43(1):110–121.

Numakawa, T., Suzuki, S., Kumamaru, E., Adachi, N., Richards, M., and Kunugi, H. (2010). BDNF function and intracellular signaling in neurons. *Histology and histopathology*, 25(2):237–258.

Nykänen, A., Haley, B., and Zamore, P. D. (2001). ATP requirements and small interfering RNA structure in the RNA interference pathway. *Cell*, 107(3):309–321.

Onishi, H., Mizukami, M., Hanagata, H., Tokunaga, M., Arakawa, T., and Miyauchi, A. (2013). Efficient production of anti-fluorescein and anti-lysozyme as single-chain antibody fragments (scfv) by *Brevibacillus* expression system. *Protein Expression and Purification*, 91(2):184–91.

Overhoff, M., Alken, M., Far, R. K.-K., Lemaitre, M., Lebleu, B., Sczakiel, G., and Robbins, I. (2005). Local RNA target structure influences sirna efficacy: A systematic global analysis. *Journal of Molecular Biology*, 348(4):871–881.

Park, K. and Shimizu, N. (1996). Expression of brain-derived neurotrophic factor by *Bacillus subtilis*. *Journal of Fermentation and Bioengineering*, 82(6):585–588.

Parker, J. S., Parizotto, E. A., Wang, M., Roe, S. M., and Barford, D. (2009). Enhancement of the seed-target recognition step in RNA silencing by a PIWI/MID domain protein. *Molecular Cell*, 33(2):204–214.

Paroo, Z., Ye, X., Chen, S., and Liu, Q. (2009). Phosphorylation of the human microRNA-generating complex mediates MAPK/Erk signaling. *Cell*, 139(1):112–122.

Patzel, V., Rutz, S., Dietrich, I., Köberle, C., Scheffold, A., and Kaufmann, S. H. E. (2005). Design of siRNAs producing unstructured guide-RNAs results in improved RNA interference efficiency. *Nature Biotechnology*, 23(11):1440–1444.

Pellino, J. L. (2007). *Dicer's role in RISC assembly during RNA silencing*. PhD thesis, Northwestern University.

Rand, T. A., Petersen, S., Du, F., and Wang, X. (2005). Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. *Cell*, 123(4):621–629.

Reynolds, A., Anderson, E. M., Vermeulen, A., Fedorov, Y., Robinson, K., Leake, D., Karpilow, J., Marshall, W. S., and Khvorova, A. (2006). Induction of the interferon response by siRNA is cell type- and duplex length-dependent. *RNA*, 12(6):988–993.

Reynolds, A., Leake, D., Boese, Q., Scaringe, S., Marshall, W. S., and Khvorova, A. (2004). Rational sirna design for rna interference. *Nature Biotechnology*, 22(3):326–330.

Rivas, F. V., Tolia, N. H., Song, J. J., Aragon, J. P., Liu, J. D., Hannon, G. J., and Joshua-Tor, L. (2005). Purified Argonaute2 and an siRNA form recombinant human RISC. *Nature Structural & Molecular Biology*, 12(4):340–349.

Robbins, M., Judge, A., and Liang, L. (2007). $2'$-O-methyl-modified RNAs act as TLR7 antagonists. *Molecular Therapy*, 15(9):1663–9.

Rosenfeld, R. and Benedek, K. (1993). Conformational changes of brain-derived neurotrophic factor during reversed-phase high-performance liquid chromatography. *Journal of Chromatography A*, 632(1-2):29–36.

Rosenthal, A., Goeddel, D. V., Nguyen, T., Martin, E., Burton, L. E., Shih, A., Laramee, G. R., Wurm, F., Mason, A., Nikolics, K., and Winslow, J. W. (1991). Primary structure and biological activity of human brain-derived neurotrophic factor. *Endocrinology*, 129(3):1289–94.

Sakurai, K., Amarzguioui, M., Kim, D., Alluin, J., Heale, B., Song, M., Gatignol, A., Behlke, M. A., and Rossi, J. J. (2011). A role for human Dicer in pre-RISC loading of siRNAs. *Nucleic Acids Research*, 39(4):1510–1525.

Salomon, W. E., Jolly, S. M., Moore, M. J., Zamore, P. D., and Serebrov, V. (2015). Single-molecule imaging reveals that Argonaute reshapes the binding properties of its nucleic acid guides. *Cell*, 162(1):84–95.

Samuel-Abraham, S. and Leonard, J. N. (2010). Staying on message: Design principles for controlling nonspecific responses to siRNA. *The FEBS Journal*, 277(23):4828–4836.

Schellman, J. A. (1997). Temperature, stability, and the hydrophobic interaction. *Biophysical Journal*, 73(6):2960–2964.

Schirle, N. T. and MacRae, I. J. (2012). The crystal structure of human Argonaute2. *Science*, 336(6084):1037–40.

Schlee, M., Hornug, V., and Hartmann, G. (2006). siRNA and isRNA: Two edges of one sword. *Molecular Therapy*, 14(4):463–470.

Schubert, S., Grünweller, A., Erdmann, V. A., and Kurreck, J. (2005). Local RNA target structure influences siRNA efficacy: Systematic analysis of intentionally designed binding regions. *Journal of Molecular Biology*, 348(4):883–893.

Schultz, N., Marenstein, D. R., De Angelis, D. A., Wang, W. Q., Nelander, S., Jacobsen, A., Marks, D. S., Massague, J., and Sander, C. (2011). Off-target effects dominate a large-scale RNAi screen for modulators of the TGF-beta pathway and reveal microRNA regulation of TGFBR2. *Silence*, 2:3.

Schwarz, D., Hutvagner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P. (2003). Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, 115(2):199–208.

Sciabola, S., Cao, Q., Orozco, M., Faustino, I., and Stanton, R. V. (2012). Improved nucleic acid descriptors for siRNA efficacy prediction. *Nucleic Acids Research*, 41(3):1383–94.

Seitz, H., Tushir, J. S., and Zamore, P. D. (2011). A 5'-uridine amplifies miRNA/miRNA* asymmetry in *Drosophila* by promoting RNA-induced silencing complex formation. *Silence*, 2:4.

Shabalina, S. A., Spiridonov, A. N., and Ogurtsov, A. Y. (2006). Computational models with thermodynamic and composition features improve siRNA design. *BMC Bioinformatics*, 7:65.

Shao, Y., Chan, C. Y., Maliyekkel, A., Lawrence, C. E., Roninson, I. B., and Ding, Y. (2007). Effect of target secondary structure on RNAi efficiency. *RNA*, 13(10):1631–1640.

Shen, M., Xu, Y., Jia, R., Zhou, X., and Ye, K. (2010). Size-independent and noncooperative recognition of dsRNA by the *Rice stripe virus* RNA silencing suppressor NS3. *Journal of Molecular Biology*, 404(4):665–679.

Sledz, C. A., Holko, M., de Veer, M. J., Silverman, R. H., and Williams, B. R. G. (2003). Activation of the interferon system by short-interfering RNAs. *Nature Cell Biology*, 5(9):834–839.

Snead, N. M. and Rossi, J. J. (2010). Biogenesis and function of endogenous and exogenous siRNAs. *Wiley Interdisciplinary Reviews: RNA*, 1(1):117–131.

Snead, N. M. and Rossi, J. J. (2012). RNA interference trigger variants: Getting the most out of RNA for RNA interference-based therapeutics. *Nucleic Acid Therapeutics*, 22(3):139–146.

Sontheimer, E. J. (1994). Site-specific RNA crosslinking with 4-thiouridine. *Molecular Biology Reports*, 20(1):35–44.

Soppet, D., Escandon, E., Maragos, J., Middlemas, D. S., Reid, S. W., Blair, J., Burton, L. E., Stanton, B. R., Kaplan, D. R., Hunter, T., Nikolics, K., and Parada, L. F. (1991). The neurotrophic factors brain-derived neurotrophic factor and neurotrophin-3 are ligands for the trkB tyrosine kinase receptor. *Cell*, 65(5):895–903.

Stalder, L., Heusermann, W., Sokol, L., Trojer, D., Wirz, J., Hean, J., Fritzsche, A., Aeschimann, F., Pfanzagl, V., Basselet, P., Weiler, J., Hintersteiner, M., Morrissey, D. V., and Meisner-Kober, N. C. (2013). The rough endoplasmatic reticulum is a central nucleation site of siRNA-mediated RNA silencing. *EMBO Journal*, 32(8):1115–1127.

Su, H., Trombly, M. I., Chen, J., and Wang, X. (2009). Essential and overlapping functions for mammalian Argonautes in microRNA silencing. *Genes & Development*, 23(3):304–317.

Sun, P. D. and Davies, D. R. (1995). The cystine-knot growth-factor superfamily. *Annual Reviews of Biophysics and Biomolecular Structure*, 24:269–291.

Sun, X., Rogoff, H. A., and Li, C. J. (2008). Asymmetric RNA duplexes mediate RNA interference in mammalian cells. *Nature Biotechnology*, 26(12):1379–1382.

Suzuki, H. I., Katsura, A., Yasuda, T., Ueno, T., Mano, H., Sugimoto, K., and Miyazono, K. (2015). Small-RNA asymmetry is directly driven by mammalian Argonautes. *Nature Structural & Molecular Biology*, 22(7):512–521.

Takagi, H., Kadowaki, K., and Udaka, S. (1989a). Screening and characterization of protein-hyperproducing bacteria without detectable exoprotease activity. *Agricultural and Biological Chemistry*, 53(3):691–699.

Takagi, H., Miyauchi, A., Kadowaki, K., and Udaka, S. (1989b). Potential use of *Bacillus brevis* HPD31 for the production of foreign proteins. *Agricultural and Biological Chemistry*, 53(8):2279–2280.

Takahashi, T., Miyakawa, T., Zenno, S., Nishi, K., Tanokura, M., and Ui-Tei, K. (2013). Distinguishable *in vitro* binding mode of monomeric TRBP and dimeric PACT with siRNA. *PLoS ONE*, 8(5):e63434.

Takahashi, T., Zenno, S., Ishibashi, O., Takizawa, T., Saigo, K., and Ui-Tei, K. (2014). Interactions between the non-seed region of siRNA and RNA-binding RLC/RISC proteins, Ago and TRBP, in mammalian cells. *Nucleic Acids Research*, 42(8):5256–5269.

Takasaki, S. (2009). Selecting effective siRNA target sequences by using Bayes' theorem. *Computational Biology and Chemistry*, 33(5):368–372.

Takasaki, S., Kotani, S., and Konagaya, A. (2004). An effective method for selecting siRNA target sequences in mammalian cells. *Cell Cycle*, 3(6):788–793.

Takeshita, T., Fujimori, K., and Shimizu, N. (1996). Synthesis of biologically active human nerve growth factor and brain-derived neurotrophic factor by a cell-free system. *Journal of Fermentation and Bioengineering*, 81(1):13–17.

Tanaka, K. and Kumano, M. (2000). Stable pharmaceutical composition of BDNF. US Patent 6,077,829.

Tanudji, M., Machalek, D., Arndt, G. M., and Rivory, L. (2010). Competition between siRNA duplexes: Impact of RNA-induced silencing complex loading efficiency and comparison between conventional-21 bp and Dicer-substrate siRNAs. *Oligonucleotides*, 20(1):27–32.

Teng, H. K., Teng, K. K., Lee, R., Wright, S., Tevar, S., Almeida, R. D., Kermani, P., Torkin, R., Chen, Z. Y., Lee, F. S., Kraemer, R. T., Nykjaer, A., and Hempstead, B. L. (2005). ProBDNF induces neuronal apoptosis via activation of a receptor complex of p75NTR and sortilin. *Journal of Neuroscience*, 25(22):5455–63.

Tluk, S., Jurk, M., Forsbach, A., Weeratna, R., Samulowitz, U., Krieg, A. M., Bauer, S., and Vollmer, J. (2009). Sequences derived from self-RNA containing certain natural modifications act as suppressors of RNA-mediated inflammatory immune responses. *International Immunology*, 21(5):607–619.

Tomari, Y., Matranga, C., Haley, B., Martinez, N., and Zamore, P. D. (2004). A protein sensor for siRNA asymmetry. *Science*, 306(5700):1377–1380.

Turner, J. J., Jones, S. W., Moschos, S. A., Lindsay, M. A., and Gait, M. J. (2006). MALDI-TOF mass spectral analysis of siRNA degradation in serum confirms an RNAse A-like activity. *Molecular BioSystems*, 3(1):43–50.

Ui-Tei, K., Naito, Y., Takahashi, F., Haraguchi, T., Ohki-Hamazaki, H., Juni, A., Ueda, R., and Saigo, K. (2004). Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference. *Nucleic Acids Research*, 32(3):936–948.

Varkonyi-Gasic, E. and Hellens, R. P. (2011). Quantitative stem-loop RT-PCR for detection of microRNAs. In Kodama, H. and Komamine, A., editors, *RNAi and Plant Gene Function Analysis: Methods and Protocols*, volume 744 of *Methods in Molecular Biology*, pages 145–157.

Vasina, J. A. and Baneyx, F. (1997). Expression of aggregation-prone recombinant proteins at low temperatures: A comparative study of the *Escherichia coli* cspA and tac promoter systems. *Protein Expression and Purification*, 9(2):211–8.

Vermeulen, A., Behlen, L., Reynolds, A., Wolfson, A., Marshall, W. S., Karpilow, J., and Khvorova, A. (2005). The contributions of dsRNA structure to Dicer specificity and efficiency. *RNA*, 11(5):674–682.

Vert, J. P., Foveau, N., Lajaunie, C., and Vandenbrouck, Y. (2006). An accurate and interpretable model for siRNA efficacy prediction. *BMC Bioinformatics*, 7(520).

Vickers, T. A., Koo, S., Bennett, C. F., Crooke, S. T., Dean, N. M., and Baker, B. F. (2003). Efficient reduction of target RNAs by small interfering RNA and RNase H-dependent antisense agents. A comparative analysis. *Journal of Biological Chemistry*, 278(9):7108–7118.

Volkov, A. A., Kruglova, N. S., Meschaninova, M. I., Venyaminova, A. G., Zenkova, M. A., Vlassov, V. V., and Chernolovskaya, E. L. (2009). Selective protection of nuclease-sensitive sites in siRNA prolongs silencing effect. *Oligonucleotides*, 19(2):191–202.

Walton, S. P., Wu, M., Gredell, J. A., and Chan, C. (2010). Designing highly active siRNAs for therapeutic applications. *The FEBS journal*, 277(23):4806–4813.

Wang, B., Li, S., Qi, H., and Chowdhury, D. (2009). Distinct passenger strand and mRNA cleavage activities of human Argonaute proteins. *Nature Structural & Molecular Biology*, 16(12):1259–66.

Wang, Y., Sheng, G., Juranek, S., Tuschl, T., and Patel, D. J. (2008). Structure of the guide-strand-containing Argonaute silencing complex. *Nature*, 456(7219):209–213.

Weber, C., Müller, C., Podszuweit, A., Montino, C., Vollmer, J., and Forsbach, A. (2012). Toll-like receptor (TLR) 3 immune modulation by unformulated small interfering RNA or DNA and the role of CD14 (in TLR-mediated effects). *Immunology*, 136(1):64–77.

Wee, L. M., Flores-Jasso, C. F., Salomon, W. E., and Zamore, P. D. (2013). Argonaute divides its RNA guide into domains with distinct functions and RNA-binding properties. *Cell*, 152(1-2):1055–1067.

Weitzer, S. and Martinez, J. (2007). The human RNA kinase hClp1 is active on 3' transfer RNA exons and short interfering RNAs. *Nature*, 447(7141):222–6.

Willkomm, S., Deerberg, A., Heidemann, J., Flugge, F., Meine, J., Hu, R., Kretschmer-Kazemi Far, R., and Restle, T. (2016). Recombinant hTRBP and hPACT modulate hAgo2-catalyzed siRNA-mediated target RNA cleavage *in vitro*. *PloS One*, 11(1):e0146814–e0146814.

Wilson, R. C., Tambe, A., Kidwell, M. A., Noland, C. L., Schneider, C. P., and Doudna, J. A. (2015). Dicer-TRBP complex formation ensures accurate mammalian microRNA biogenesis. *Molecular Cell*, 57(3):397–407.

Wittrup, A. and Lieberman, J. (2015). Knocking down disease: A progress report on siRNA therapeutics. *Nature Reviews Genetics*, 16(9):543–552.

Yamagata, H., Adachi, T., Tsuboi, A., Takao, M., Sasaki, T., Tsukagoshi, N., and Udaka, S. (1987). Cloning and characterization of the 5' region of the cell wall protein gene operon in *Bacillus brevis* 47. *Journal of Bacteriology*, 169(3):1239–1245.

Yang, J., Maurin, T., Robine, N., Rasmussen, K. D., Jeffrey, K. L., Chandwani, R., Papapetrou, E. P., Sadelain, M., O'Carroll, D., and Lai, E. C. (2010). Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis. *Proceedings of the National Academy of Sciences*, 107(34):15163–15168.

Yang, X., Sierant, M., Janicka, M., Peczek, L., Martinez, C., Hassell, T., Li, N., Li, X., Wang, T., and Nawrot, B. (2012). Gene silencing activity of siRNA molecules containing phosphorodithioate substitutions. *ACS Chemical Biology*, 7(7):1214–1220.

Ye, X., Huang, N., Liu, Y., Paroo, Z., Huerta, C., Li, P., Chen, S., Liu, Q., and Zhang, H. (2011). Structure of C3PO and mechanism of human RISC activation. *Nature Structural & Molecular Biology*, 18(6):650–657.

Yi, R., Qin, Y., Macara, I., and Cullen, B. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes & Development*, 17(24):3011–3016.

Yoda, M., Kawamata, T., Paroo, Z., Ye, X., Iwasaki, S., Liu, Q., and Tomari, Y. (2010). ATP-dependent human RISC assembly pathways. *Nature Structural & Molecular Biology*, 17(1):17–23.

Yoshinari, K., Miyagishi, M., and Taira, K. (2004). Effects on RNAi of the tight structure, sequence and position of the targeted region. *Nucleic Acids Research*, 32(2):691–699.

Zamore, P. D., Tuschl, T., Sharp, P. A., and Bartel, D. P. (2000). RNAi: Double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell*, 101(1):25–33.

Zuccato, C. and Cattaneo, E. (2009). Brain-derived neurotrophic factor in neurodegenerative diseases. *Nature Reviews Neurology*, 5(6):311–22.