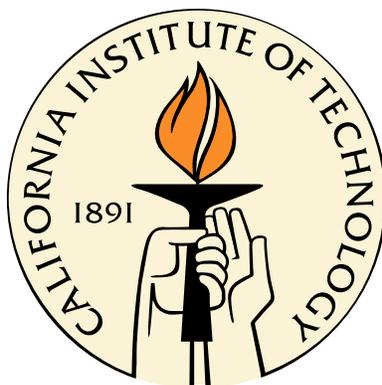


Engineering nucleic acid mechanisms for regulation and readout of gene expression: conditional Dicer substrate formation and sensitive multiplexed northern blots

Thesis by
Ma'ayan Schwarzkopf

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy



California Institute of Technology
Pasadena, California

2013
(Defended May 9, 2013)

© 2013

Ma'ayan Schwarzkopf

All Rights Reserved

To my husband Yoni, for walking this path with me.

Acknowledgments

Caltech is a special environment and the skills I learned here will no doubt help me no matter where my career leads. I would like to thank Caltech, and the Biology Department in particular, for allowing me the opportunity and freedom to conduct exciting and innovative research in the Department of Bioengineering. I thank my advisor, Dr. Niles Pierce, for introducing me to the world of nucleic acid engineering and for providing me the opportunity to work in such a unique and creative lab. I truly appreciate the support I received as well as the liberty to pursue my own path and the inspiration for excellence. To my thesis committee, Dr. David Baltimore, Dr. Raymond Deshaies and Dr. Bruce Hay, I thank you for your advice and support over the years. I would also like to thank Melinda Kirk for her administrative assistance and for keeping our lab running so well. In addition, I would like to thank the McCallum Biology Divisional fellowship, the NSF and the NIH for funding.

I am also indebted to my colleagues and friends in the Pierce lab. You gave me many great ideas and advice throughout the years as well as support when things weren't working. Most importantly, you made it fun to come into lab every day. In particular, I would like to thank Lisa Hochrein, Jonathan Sternberg, Harry Choi, Victor Beck and Jeff Vieregge. I would also like to acknowledge Jessie Ge, our very talented undergraduate Caltech SURF student. To our neighbors in the Elowitz lab, especially Fred Tan, Joe Levine and Leah Santat, thank you for your technical support and experimental advice as well as for being my enzyme supplier in times of need.

From my undergraduate studies and MSc, I thank Dr. Adi Avni, Dr. Roni Aloni and Silvia Schuster who introduced me to the fascinating world of experimental biology. I thank Silvia Schuster in particular for laying the building blocks for my experimental skills and especially time management.

A special thanks to my friends in Pasadena who made this a home away from home: Jonathan Sternberg, Kayla Birns, Or and Keren Yogev, Oren Schaedel, Shelly Tzlil, Morgan Putnam, Lee and Mati Mann, Taly Lindner and Yael and Yaron Antebi.

I would like to thank my extended family: Jimmy, Sylvia, Ran, Aude, Sharon and Idan. You welcomed me into your family many years ago and have been an important part of my life. To my brother Yoav, I love you, we both get to graduate this year, yay! To my parents, Miki and Miri Nagler, where do I begin? Thank you for making me who I am today. Thank you for your guidance, for believing in me even when I don't and for supporting the decisions I make. Thank you for giving me everything I could ever ask for. I love you and I am forever grateful for having such great parents.

Above all, I thank my husband Yoni. I couldn't have gone through this journey without you! Thank you for believing in me and for pushing me to be the best I can be. Thank you for your patience (I know it was hard at times) and for always listening to what's going on with my experiments and for making the effort to understand every single detail about my work. You truly deserve an honorary PhD in biology. And of course, thank you for being my amazing personal chef and for making sure that I eat something other than "schinzel and chips." Finally, thank you for giving me the greatest gift of all, Ori and Tom. The three of you put everything else into perspective.

Abstract

Nucleic acids are most commonly associated with the genetic code, transcription and gene expression. Recently, interest has grown in engineering nucleic acids for biological applications such as controlling or detecting gene expression. The natural presence and functionality of nucleic acids within living organisms coupled with their thermodynamic properties of base-pairing make them ideal for interfacing (and possibly altering) biological systems. We use engineered small conditional RNA or DNA (scRNA, scDNA, respectively) molecules to control and detect gene expression. Three novel systems are presented: two for conditional down-regulation of gene expression via RNA interference (RNAi) and a third system for simultaneous sensitive detection of multiple RNAs using labeled scRNAs .

RNAi is a powerful tool to study genetic circuits by knocking down a gene of interest. RNAi executes the logic: *If gene Y is detected, silence gene Y*. The fact that detection and silencing are restricted to the same gene means that RNAi is constitutively on. This poses a significant limitation when spatiotemporal control is needed. In this work, we engineered small nucleic acid molecules that execute the logic: *If mRNA X is detected, form a Dicer substrate that targets independent mRNA Y for silencing*. This is a step towards implementing the logic of conditional RNAi: *If gene X is detected, silence gene Y*. We use scRNAs and scDNAs to engineer signal transduction cascades that produce an RNAi effector molecule in response to hybridization to a nucleic acid target X. The first mechanism is solely based on hybridization cascades and uses scRNAs to produce a double-stranded RNA (dsRNA) Dicer substrate against target gene Y. The second mechanism is

based on hybridization of scDNAs to detect a nucleic acid target and produce a template for transcription of a short hairpin RNA (shRNA) Dicer substrate against target gene Y. Test-tube studies for both mechanisms demonstrate that the output Dicer substrate is produced predominantly in the presence of a correct input target and is cleaved by Dicer to produce a small interfering RNA (siRNA). Both output products can lead to gene knockdown in tissue culture. To date, signal transduction is not observed in cells; possible reasons are explored.

Signal transduction cascades are composed of multiple scRNAs (or scDNAs). The need to study multiple molecules simultaneously has motivated the development of a highly sensitive method for multiplexed northern blots. The core technology of our system is the utilization of a hybridization chain reaction (HCR) of scRNAs as the detection signal for a northern blot. To achieve multiplexing (simultaneous detection of multiple genes), we use fluorescently tagged scRNAs. Moreover, by using radioactive labeling of scRNAs, the system exhibits a five-fold increase, compared to the literature, in detection sensitivity. Sensitive multiplexed northern blot detection provides an avenue for exploring the fate of scRNAs and scDNAs in tissue culture.

3.1	Introduction	39
3.2	Results	40
3.2.1	Examination of toehold mediated branch migration in tissue culture and combatting delivery	41
3.2.2	Study of hairpin degradation in tissue culture	46
3.2.3	Silencing by mechanism hairpin C is mediated by RNAi	46
3.3	Discussion	48
3.4	Materials and methods	51
4	Engineering a conditional shRNA transcription mechanism	58
4.1	Introduction	58
4.2	Mechanism and Design	61
4.2.1	Mechanism	61
4.2.2	Design	61
4.2.3	Promoter choice	63
4.3	Results	64
4.3.1	<i>In vitro</i> studies	64
4.3.2	<i>In vivo</i> studies	64
4.4	Discussion	67
4.5	Materials and methods	71
5	Sensitive multiplexed northern blots via hybridization chain reaction (HCR)	85
5.1	Introduction	85
5.1.1	Northern blot	86
5.1.2	Northern blot probes and detection strategies	86
5.1.3	Northern blot in comparison to other methods	87
5.1.4	Hybridization chain reaction (HCR)	88

5.2	HCR as a detection method for northern blots	90
5.3	Results	91
5.3.1	HCR “dot blot”	91
5.3.2	mRNA blots	93
5.3.2.1	Specificity and blotting conditions	93
5.3.2.2	Multiplexed detection of endogenous targets	94
5.3.3	miRNA blots	99
5.3.3.1	Blotting conditions for miRNAs	99
5.3.3.2	Multiple HCR-initiator probes	102
5.3.3.3	Multiplexed miRNA detection using HCR	103
5.3.3.4	Detection limit of HCR-based miRNA northern blot detection	105
5.3.3.5	Estimation of polymer length	107
5.4	Discussion	108
5.5	Materials and methods	111
A	Supplementary information for Chapter 2	125
A.1	Stepping gel	125
A.2	Quantification of ON and OFF states	126
A.3	Quantification of catalytic property	126
A.4	<i>In vitro</i> Dicer assay	127
B	Supplementary information for Chapter 3	132
B.1	Catalytic Dicer substrate formation mechanism based on 5'-toehold hairpins	132
C	Supplementary information for Chapter 4	135
C.1	Stepping gel	135
C.1.1	Conditional shRNA transcription and Dicer gel	136
C.1.2	Quantification of ON-to-OFF ratios	136
C.2	Design of a triggered shRNA transcription mechanism using the H1 promoter	137

C.3 Materials and methods 141

List of Figures

1.1	Hairpin and toehold-mediated branch migration nomenclature and schematics	3
2.1	Conditional catalytic DsiRNA formation schematic	14
2.2	Conditional catalytic DsiRNA formation in a test tube	18
2.3	Conditional siRNA production <i>in vitro</i>	19
2.4	Transfection of conditional Dicer substrate formation into HEK293 d2EGFP DsRed cells	21
2.5	Conditional Dicer substrate formation is not functional <i>in vivo</i>	23
3.1	d2EGFP knockdown via a simplified triggered Dicer substrate formation mechanism (M1)	44
3.2	Expression of three hairpins from one plasmid, M1 mechanism	45
3.3	Hairpin C of mechanism M1 is partially degraded in tissue culture	47
3.4	Hairpin C of mechanism M2 leads to RNAi mediated knockdown	49
4.1	Conditional shRNA transcription mechanism schematic	62
4.2	Conditional shRNA transcription	65
4.3	<i>In vitro</i> transcribed shRNA down-regulates d2EGFP	67
4.4	T7 RNA expression and functionality in HEK293 d2EGFP cell line	68
5.1	HCR polymer formation schematic	90
5.2	Northern blot detection via HCR schematic	92

5.3	EGFP “dot blot” detection via HCR	93
5.4	Specific d2EGFP detection via HCR	95
5.5	Multiplexed GAPDH and ssRNA ladder detection via HCR	97
5.6	Multiplexed detection of GAPDH, 18S rRNA and ssRNA ladder via HCR	98
5.7	Multiplexed detection of synthetic miR-16a target, endogenous U6 RNA and microRNA ladder via HCR	101
5.8	Detection of endogenous miR-16a target using a FAM-labeled LNA probe	102
5.9	HCR-based detection using probes with multiple initiators schematic	103
5.10	HCR-based detection using probes with single, double and quad initiators	104
5.11	Multiplexed HCR-based detection of endogenous miR-16a	105
5.12	Multiplexed HCR-based detection of endogenous miR-16a, miR-21 and U6	106
5.13	Multiplexed HCR-based detection of miR-16a and miR-21	107
5.14	Sensitivity of HCR-based detection	107
5.15	Estimation of HCR polymer length	108
A.1	Stepping gel for the conditional catalytic DsiRNA formation mechanism	128
A.2	Quantification of B·C final product formation in ON and OFF states	129
A.3	Quantification of catalytic B·C final product formation	130
A.4	Dicer processing gel for the conditional catalytic DsiRNA formation mechanism	131
B.1	Catalytic Dicer substrate formation mechanism based on 5'-toehold hairpins (M1)	133
C.1	Conditional shRNA transcriptional template formation	145
C.2	Conditional shRNA transcription and Dicer processing	146
C.3	Quantification of transcribed shRNA product	147
C.4	Conditional shRNA transcription mechanism based on H1 promoter schematic	147
C.5	Conditional shRNA transcription template formation with H1 promoter	148

C.6 Transfection of an annealed H1-promoter-based mechanism does not lead to
d2EGFP knockdown 149

Chapter 1

Introduction

Nucleic acids, deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), are essential macromolecules for known forms of life. DNA stores the genetic information of a cell while RNA molecules participate in transmitting, expressing and controlling this information. Fortunately, one can do more with nucleic acids than carry genetic information. Using the base-pairing properties of nucleic acids, one can use nucleic acids to design molecules that self assemble into specified structures. This field is called nucleic acid nanotechnology or nucleic acid engineering.

The field emerged three decades ago with the seminal work of Nadrian C. Seeman who devised that it is possible to generate nucleic acid sequences that form 3- and 4-arm junctions rather than linear double helix duplex [1, 2]. Since then, many other complex structures have been engineered from DNA; 2-D structures such as lattices, tubes and DNA origami as well as 3-D structures [3–7]. Compared to DNA, RNA can form non-canonical base-pairs and is less hydrolytically stable. However, it has a rich conformational and catalytic space and combined with its genetic encodability make it a powerful medium for engineering nucleic acids for biological applications [8, 9]. As with DNA, a wide variety of 2-D as well as 3-D structures have been composed out of RNA [9–12].

The goal of nucleic acid engineering is to design structures for real-life applications. Structures can be made functional by using them as scaffolds for precise spatial positioning

of other molecules. For example, metals can be positioned to serve as nano-wires or proteins can potentially be arranged for X-ray crystallography or to bring together enzymatic activities [13–16]. Three-dimensional structures such as molecular cages, can potentially be used in medical applications as a drug delivery vehicle where the cargo (drug) is released in a programmable way [17–20]. Self-assembled DNA nanoparticles functionalized with a targeting ligand and fused to siRNAs can be used for targeted drug delivery into tumor cells [21]. Furthermore, the fact that nucleic acids are used as the engineering material opens up the opportunity to interface with biological samples for controlling gene expression. For example, a biomolecular computer can be designed to analyze the levels of messenger RNA species and in response produces a molecule that affects the level of gene expression [22].

A majority of nucleic acid structures described in the literature are static structures formed by thermal annealing (a process in which the reactants are mixed together, heated and then gradually cooled until they reach the desired structure). The target structure is designed to have the lowest free energy compared to all other possible structures and is therefore thermodynamically favorable to form. Dynamically formed nucleic acids structures can also be designed utilizing an isothermal process of strand displacement [23]. In this process, two nucleic acid strands hybridize to each other while displacing a strand that was previously hybridized. The newly revealed sequence can also initiate the reveal of another sequence, resulting in a cascade. Strand displacement reactions are facilitated by single-stranded regions called ‘toeholds’ that mediate a branch migration. Reactions are driven forward by the gain of base-pairs in each step of the cascade and entropy gain from disassembly reactions. Strand displacement cascades have been used to create logic circuits [24, 25], walkers [26–29] and catalytic self-assembly [28].

In this field, our lab has introduced the concept of kinetically controlling strand displacement cascades based on hairpin motifs (a type of small conditional RNA or DNA (scRNA and scDNA, respectively)). The hairpin motif is composed of three features: a toehold, a double-stranded stem and a loop (Figure 1.1(a)). Strand displacement reac-

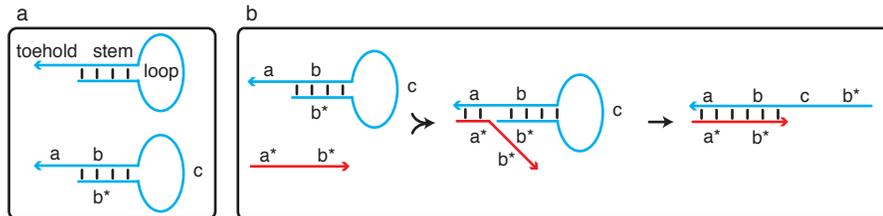


Figure 1.1: **Hairpin and toehold-mediated branch migration nomenclature and schematics.** 5' and 3' polarity are differentiated by an arrow representation of the 3' end. Letters marked with * are complementary to the corresponding unmarked letter (e.g., 'a' is complementary to 'a*'). Base-pairs are represented by a black line. (a) Hairpin schematic. A hairpin consists of three domains: a single-stranded toehold, a double-stranded stem and an unpaired loop. Letters represent domains in the structure ('a' is a toehold, 'b' and 'b*' compose the stem and 'c' is the loop). (b) Toehold-mediated strand displacement schematic. The red strand (termed initiator) displaces the stem of the hairpin. Domain 'a*' of the initiator binds to toehold 'a' of the hairpin; next, branch migration occurs and domain 'b*' of the initiator displaces domain 'b' of the hairpin. The binding of the initiator to the hairpin results in a new structure, initiator-hairpin, with an exposed new toehold 'c'. Either 'c' or both 'c -b*' can serve to initiate a downstream reaction mediated by another component (not shown).

tions cause the hairpin to change conformation and open up, resulting in a new initiator that instead of being released is still bound to the initial initiator (Figure 1.1(b)). Using hairpin-based cascades we have demonstrated autonomous locomotion [30], self-assembly of branched dendrimers and catalytic self-assembly [28] and formation of a long dsDNA or dsRNA polymer (hybridization chain reaction, HCR) [31, 32]. By functionalizing scRNAs with fluorophores, a fluorescent HCR polymer can be grown. This strategy was used for multiplexed detection of fluorescent in situ hybridization (FISH) with improved signal-to-background compared to traditional FISH [32].

The focus of the work presented here is to expand the repertoire of applications available from engineered nucleic acids using toehold-mediated strand displacement. Engineered systems have the capacity to sense molecular cues (e.g., mRNA) and to produce an output that leads to a change in the cell. Chapters 2, 3 and 4 present mechanisms with the aim

of conditionally down-regulating gene expression via RNA interference (RNAi) in response to detection of an input signal of an unrelated mRNA molecule. Chapter 2 presents a catalytic system based on scRNAs that generates an RNAi effector molecule in response to an mRNA target. In Chapter 3 we examine some possible explanations as to why conditional RNAi is not obtained in tissue culture. Chapter 4 presents a system based on scDNAs that conditionally transcribes an RNAi effector molecule. Finally, Chapter 5 expands upon the work presented by Choi et al. [32] to use fluorescent scRNAs as an output signal for a sensitive, multiplexed northern blot detection of mRNAs and miRNAs. The presented method has the potential to improve the study of transfected or expressed scRNAs described in Chapter 3.

Bibliography

- [1] Seeman, N. C. Nucleic acid junctions and lattices. *Journal of theoretical biology* **99**(2), 237–47 (1982).
- [2] Seeman, N. C., Maestre, M. F., Ma, R. I., and Kallenbach, N. R. Physical characterization of a nucleic acid junction. *Progress in clinical and biological research* **172A**, 99–108 (1985).
- [3] Chen, J. H. and Seeman, N. C. Synthesis from DNA of a molecule with the connectivity of a cube. *Nature* **350**(6319), 631–3 (1991).
- [4] Lin, C., Liu, Y., Rinker, S., and Yan, H. DNA tile based self-assembly: building complex nanoarchitectures. *Chemphyschem : a European journal of chemical physics and physical chemistry* **7**(8), 1641–7 (2006).
- [5] Feldkamp, U. and Niemeyer, C. M. Rational design of DNA nanoarchitectures. *Angewandte Chemie* **45**(12), 1856–76 (2006).
- [6] Rothemund, P. W. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**(7082), 297–302 (2006).
- [7] Ke, Y., Ong, L. L., Shih, W. M., and Yin, P. Three-dimensional structures self-assembled from DNA bricks. *Science* **338**(6111), 1177–83 (2012).
- [8] Krishnan, Y. and Bathe, M. Designer nucleic acids to probe and program the cell. *Trends in cell biology* **22**(12), 624–33 (2012).
- [9] Guo, P., Haque, F., Hallahan, B., Reif, R., and Li, H. Uniqueness, advantages, challenges, solutions, and perspectives in therapeutics applying rna nanotechnology. *Nucleic acid therapeutics* **22**(4), 226–45 (2012).
- [10] Jaeger, L. and Leontis, N. B. Tecto-RNA: One-Dimensional Self-Assembly through Tertiary Interactions. *Angewandte Chemie* **39**(14), 2521–2524 (2000).

- [11] Afonin, K. A., Bindewald, E., Yaghoubian, A. J., Voss, N., Jacovetty, E., Shapiro, B. A., and Jaeger, L. In vitro assembly of cubic RNA-based scaffolds designed in silico. *Nature nanotechnology* **5**(9), 676–82 (2010).
- [12] Grabow, W. W., Zakrevsky, P., Afonin, K. A., Chworos, A., Shapiro, B. A., and Jaeger, L. Self-assembling RNA nanorings based on RNAI/II inverse kissing complexes. *Nano letters* **11**(2), 878–87 (2011).
- [13] Seeman, N. C. DNA in a material world. *Nature* **421**(6921), 427–31 (2003).
- [14] Yan, H., Park, S. H., Finkelstein, G., Reif, J. H., and LaBean, T. H. DNA-templated self-assembly of protein arrays and highly conductive nanowires. *Science* **301**(5641), 1882–4 (2003).
- [15] Liu, Y., Lin, C., Li, H., and Yan, H. Aptamer-directed self-assembly of protein arrays on a DNA nanostructure. *Angewandte Chemie* **44**(28), 4333–8 (2005).
- [16] Jiang, Q., Song, C., Nangreave, J., Liu, X., Lin, L., Qiu, D., Wang, Z. G., Zou, G., Liang, X., Yan, H., and Ding, B. DNA origami as a carrier for circumvention of drug resistance. *Journal of the American Chemical Society* **134**(32), 13396–403 (2012).
- [17] Goodman, R. P., Heilemann, M., Doose, S., Erben, C. M., Kapanidis, A. N., and Turberfield, A. J. Reconfigurable, braced, three-dimensional DNA nanostructures. *Nature nanotechnology* **3**(2), 93–6 (2008).
- [18] Andersen, E. S., Dong, M., Nielsen, M. M., Jahn, K., Subramani, R., Mamdouh, W., Golas, M. M., Sander, B., Stark, H., Oliveira, C. L., Pedersen, J. S., Birkedal, V., Besenbacher, F., Gothelf, K. V., and Kjems, J. Self-assembly of a nanoscale DNA box with a controllable lid. *Nature* **459**(7243), 73–6 (2009).
- [19] Lo, P. K., Karam, P., Aldaye, F. A., McLaughlin, C. K., Hamblin, G. D., Cosa, G.,

- and Sleiman, H. F. Loading and selective release of cargo in DNA nanotubes with longitudinal variation. *Nature chemistry* **2**(4), 319–28 (2010).
- [20] Douglas, S. M., Bachelet, I., and Church, G. M. A logic-gated nanorobot for targeted transport of molecular payloads. *Science* **335**(6070), 831–4 (2012).
- [21] Lee, H., Lytton-Jean, A. K., Chen, Y., Love, K. T., Park, A. I., Karagiannis, E. D., Sehgal, A., Querbes, W., Zurenko, C. S., Jayaraman, M., Peng, C. G., Charisse, K., Borodovsky, A., Manoharan, M., Donahoe, J. S., Truelove, J., Nahrendorf, M., Langer, R., and Anderson, D. G. Molecularly self-assembled nucleic acid nanoparticles for targeted in vivo siRNA delivery. *Nature nanotechnology* **7**(6), 389–93 (2012).
- [22] Benenson, Y., Gil, B., Ben-Dor, U., Adar, R., and Shapiro, E. An autonomous molecular computer for logical control of gene expression. *Nature* **429**(6990), 423–9 (2004).
- [23] Zhang, D. Y. and Seelig, G. Dynamic DNA nanotechnology using strand-displacement reactions. *Nature chemistry* **3**(2), 103–13 (2011).
- [24] Seelig, G., Soloveichik, D., Zhang, D. Y., and Winfree, E. Enzyme-free nucleic acid logic circuits. *Science* **314**(5805), 1585–8 (2006).
- [25] Qian, L. and Winfree, E. Scaling up digital circuit computation with DNA strand displacement cascades. *Science* **332**(6034), 1196–201 (2011).
- [26] Sherman, W. B. and Seeman, N. C. A precisely controlled DNA biped walking device. *Nano letters* **4**(7), 1203–1207 (2004).
- [27] Shin, J. S. and Pierce, N. A. A synthetic DNA walker for molecular transport. *Journal of the American Chemical Society* **126**(35), 10834–5 (2004).
- [28] Yin, P., Choi, H. M., Calvert, C. R., and Pierce, N. A. Programming biomolecular self-assembly pathways. *Nature* **451**(7176), 318–22 (2008).

- [29] Omabegho, T., Sha, R., and Seeman, N. C. A bipedal DNA Brownian motor with coordinated legs. *Science* **324**(5923), 67–71 (2009).
- [30] Venkataraman, S., Dirks, R. M., Rothmund, P. W., Winfree, E., and Pierce, N. A. An autonomous polymerization motor powered by DNA hybridization. *Nature nanotechnology* **2**(8), 490–4 (2007).
- [31] Dirks, R. M. and Pierce, N. A. Triggered amplification by hybridization chain reaction. *Proceedings of the National Academy of Sciences of the United States of America* **101**(43), 15275–8 (2004).
- [32] Choi, H. M., Chang, J. Y., Trinh le, A., Padilla, J. E., Fraser, S. E., and Pierce, N. A. Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nature biotechnology* **28**(11), 1208–12 (2010).

Chapter 2

Engineering a conditional catalytic DsiRNA formation mechanism

2.1 Introduction

RNA is a versatile molecule responsible for many processes within a cell. It serves both as a template and as a component responsible for protein translation, RNA processing and as a regulatory element. Regulatory RNAs are non-coding RNAs (ncRNAs) that control gene expression; control of gene expression may be initiated by ribozymes, riboswitches, riboregulators, antisense and RNA interference (RNAi) [1]. In this work, RNA's ability to down-regulate gene expression through RNA interference (RNAi) [2, 3] is utilized. RNAi has the potential to silence any gene, which has made it an attractive tool to probe gene function and serve as a potential therapeutic [4–6].

RNAi is a mechanism of post transcriptional gene silencing, induced by small interfering RNAs (siRNAs) in a sequence specific manner. siRNAs are short dsRNAs 21–25 nucleotides in length, with a phosphate at the 5' end and a two-base overhang at the 3' end [7–10]. siRNAs can be introduced directly into the cell or they may be processed in the cytoplasm from long double-stranded RNAs (dsRNAs) or short-hairpin RNAs (shRNAs) by an RNaseIII endonuclease called Dicer [11–13]. Following siRNA formation, a single strand of the siRNA, the guide strand, is incorporated into a complex of proteins known

as the RNA-induced silencing complex (RISC) [14], while the second strand, the passenger strand, gets degraded [15–18]. Next, RISC uses the guide strand to find the complementary mRNA sequence via Watson-Crick base-pairing and endonucleolytically cleaves the target mRNA [10, 19]. Once activated, RISC can undergo multiple rounds of mRNA cleavage, mediating a robust response against the target gene [20].

The logic operation RNAi implements using an siRNA for gene Y is: *silence gene Y*. As a result, RNAi is constitutively on. This may pose a limitation on the study of essential genes as well as therapeutics. There are several approaches for spatio-temporal control of RNAi, divided into “traditional” and “engineered” approaches. Traditional approaches include targeted delivery [21] and controlled shRNA expression either by tissue-specific promoters and/or activation/inactivation of promoters by small molecules or enzymatic means [22–26] (discussed in Chapter 4). Engineered approaches rely on the use of non-coding RNAs to control RNAi activity following detection of an input signal. These non-coding RNAs harbor an RNAi effector molecule which, in its initial state, may or may not be functional. The presence of an input signal results in a conformational change of the molecule allowing RNAi to be turned on or off.

Small molecules have been used as the input signal to obtain conditional RNAi in cells [27–31]. These mechanisms use an aptamer as the input signal that controls RNAi. In addition to conditionality, small molecule activators also allow the output signal to be tuned (e.g., more ligand, stronger signal). Current small-ligand based mechanisms rely on cellular expression of the non-coding RNA and exogenous addition of the input signal.

As an alternative to small molecules, nucleic acids can be used to control the conformation of non-coding RNAs; this opens the possibility of controlling RNAi via endogenous nucleic acids such as mRNA. Several groups have attempted to achieve this goal. Xie et al. [32] have engineered a sensor that generates an siRNA in response to a 140-nt RNA in *Drosophila* embryo cell-free extract. While the system succeeds in detecting a long RNA molecule, the detection sequence differs by two nucleotides from the sense strand of the

output siRNA resulting in the logic: If gene Y is detected, silence gene Y'.

With careful design, a conditional RNAi system can be engineered to implement the logic operation: *If gene X is detected, silence independent gene Y.*

So far, two systems [33, 34] have been designed toward this goal. While both designs detect and silence two independent sequences, they rely on the use of a short synthetic target and not on a full-length mRNA. In the first system, Kumar et al. [34] based their design on the miRNA pathway. In this pathway, a long stem-loop structure called primary-miRNA is being processed by Drosha to produce a precursor-miRNA (pre-miRNA) which can then be processed by Dicer [35, 36]. In their system, a non-coding RNA was expressed in mammalian cells; upon the transfection of a short target X, the non-coding RNA underwent a conformational change which resulted in the formation of a pri-miRNA. In the second system, Masu et al. [33] engineered a system that, when annealed in a test tube with a short synthetic RNA target X, generates a Dicer substrate. Neither system worked by detecting an endogenous target. For RNAi activation in cells, the systems were either reacted outside the cells or expressed inside the cells and triggered by addition of a chemically modified synthetic target.

In this chapter, we present a conditional RNAi mechanism, that is intended to implement the logic: *If gene X is detected, silence independent gene Y.* Upon the detection of mRNA X, the mechanism produces a Dicer substrate targeting gene Y. Conditional RNAi activation is mediated by small conditional RNAs (scRNAs) through toehold-mediated strand displacement. Activation of the mechanism is demonstrated in the presence of a short synthetic nucleic acid target and full-length mRNA in a test tube. The mechanism exhibits good ON-to-OFF ratio; in the absence of a detection target, minimal Dicer substrate is formed. The Dicer substrate can be processed by Dicer in vitro while initial and intermediate components remain intact. When generated in a test tube, the Dicer substrate leads to down-regulation of gene expression in tissue culture.

2.2 Mechanism

We envision a conditional RNAi mechanism based on Dicer processing. As such, the output of the mechanism, the final product, must be an RNAi effector molecule: an siRNA, shRNA or longer dsRNA. A long dsRNA was chosen since it allows more flexibility in the design. Moreover, Kim et al. [37] have demonstrated that synthetic dsRNA duplexes 25–30 nucleotides long are more potent RNAi activators than siRNAs. This enhanced potency is attributed to the fact that longer dsRNAs are Dicer substrates, directly linking siRNA production and incorporation into RISC.

Our mechanism is based on metastable hairpins. By metastable we mean that the hairpin conformation is not the global minima and it is kinetically trapped in a hairpin state. A good mechanism should have good ON and OFF states. In the OFF state the hairpins are kinetically trapped in the monomer state. In the presence of a detection target (X) the mechanism is turned ON and the hairpins are “released” from their trap and interact to form a Dicer substrate. The system components are hairpins with 3′ toeholds. This way, as opposed to the mechanism presented in Appendix B, the siRNA antisense sequence is never exposed as a single-stranded region during mechanism transduction and therefore cannot interact with the silencing target prematurely.

The mechanism depicted in Figure 2.1, is an extension of the catalytic hairpin cascades pioneered by Yin et al. [38]. The mechanism reactants consist of three hairpins: the first hairpin (A) detects the detection target X, while the other two hairpins (B and C) serve to produce a Dicer substrate against the silencing target Y. In the presence of detection target X, toehold ‘a*’ of hairpin A binds to ‘a’ of the detection target, initiating a branch migration of hairpin A, ending in the complex X·A and exposing ‘c’ on hairpin A (step 1, Figure 2.1). Next, toehold ‘c*’ of hairpin B binds to ‘c’ in the single-stranded region of complex X·A, followed by a branch migration leading to the opening of hairpin B and to the formation of complex X·A·B, exposing ‘x’ on B (step 2, Figure 2.1). In the third step, toehold ‘x*’ of hairpin C binds to ‘x’ in the single-stranded region of complex X·A·B,

leading to a branch migration and formation of complex X·A·B·C (step 3a, Figure 2.1). Finally, entropy drives the release of complex B·C from X·A (step 3b, Figure 2.1). The formed B·C complex harbors the output signal, domains ‘v-w-x-y-z’, which serve as a substrate for Dicer, triggering the RNAi pathway.

The output of the mechanism (B·C in Figure 2.1) is inspired by Dicer substrate interfering RNA (DsiRNA). Typical DsiRNAs contain a 25bp stem and a two nucleotide 3’ overhang in the antisense strand [37, 39, 40]. The incorporation of a 3’ overhang on only one end introduces a preference for Dicer processing to start from that end since Dicer acts as a molecular ruler, measuring its cleavage site from the 3’ end overhang [39, 41]. Due to design constraints, the B·C final product contains a longer stem and a 5’ overhang instead of a blunt end.

A notable feature of this mechanism is catalytic production of the final Dicer substrate B·C. When complex B·C dissociates from X·A·B·C then X·A is released (step 3b, Figure 2.1). The re-emergence of complex X·A allows it to interact with a new B hairpin without the need to detect a new target molecule. This way, the detection of one target molecule (by one A hairpin) can lead to the formation of multiple Dicer substrates (B·C duplexes), limited by the amount of B and C hairpins present.

2.3 Design

The design follows the logic operation: *If gene X is detected, silence independent gene Y*. As can be seen in Figure 2.1, complete sequence independence is observed between the detection target X (sequence ‘a-b-c-d’) and the silencing target Y (sequence ‘v-w-x-y-z’). This makes it possible to re-program the mechanism to detect and silence different genes. The sequence design space is constrained by the detection and silencing targets. The sequences are constrained by either the detection target (DsRed mRNA) or silencing target (destabilized eGFP mRNA, d2EGFP) with the exception of domain ‘e’. The purpose of this domain is to allow some structural flexibility to the mRNA·A·B complex. Domain ‘e’ was

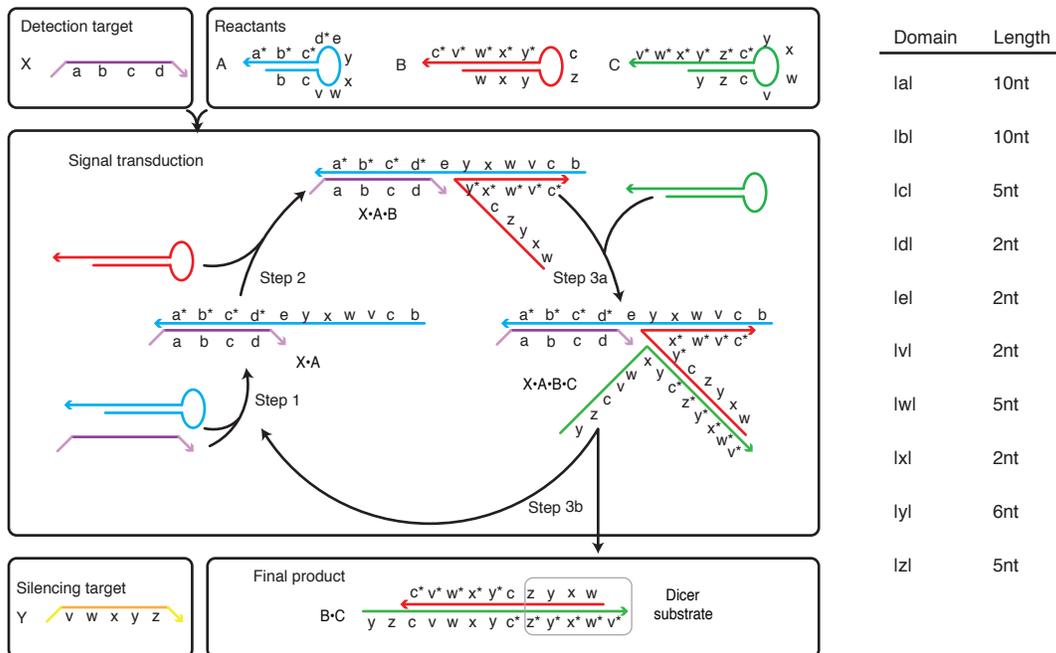


Figure 2.1: **Conditional catalytic DsiRNA formation schematic.** The detection target X is detected by hairpin A (step 1) initiating a cascade of reactions resulting in the opening of hairpins B (step 2) and C (step 3a) leading to catalytic formation of Dicer substrate B·C (step 3b). Domain dimensions are listed.

set initially as two nucleotides but it can be made longer, shorter or removed altogether. Domain ‘d’ in the loop of hairpin A is part of the detection feature of the hairpin. The purpose of this domain is to help hairpin A stay bound to the target by the addition of extra binding through the loop. This segment can also be removed if deemed unnecessary.

Sequences were assigned to our structures using the concentration-based multi-state design feature on NUPACK [42, 43](Wolfe, unpublished data). It was previously demonstrated in our lab that region 591–623 of DsRed is a good detection region. We therefore specified the design code to chose the detection sequence from this region. The siRNA output was selected from the full-length eGFP coding region. To assign sequences to structures, the design code locates regions in the mRNA that minimize the sum of the ensemble defect (a measurement of how far the system is from an ideal design) [43, 44]. Designed structures include the initial hairpin components (A, B, C), the final structure B·C as well as the structures produced in the first and second steps (X·A and X·A·B, respectively). Concentration based design was used to ensure that structure X·A·B dominates over dimer X·A and monomer B in a dilute solution. The sequence segment ‘e’ that is not constrained by the detection or silencing sequence was assigned by NUPACK to minimize the ensemble defect.

Following the initial design of hairpins on NUPACK we used NUPACK’s thermodynamic analysis tools [42, 45–47] to pick the set of designed sequences that performed best according to the design specifications (starting materials form the correct structure and that only desired complexes are formed). Based on thermodynamic analysis we then modified the dimensions of hairpins B and C to improve the design. Two nucleotides were removed from the 5’ end of hairpin B and the corresponding two nucleotides were removed from the toehold of hairpin C to ensure complementarity. This change made the stem of hairpin B shorter while making the toehold longer and was predicted to improve the binding of hairpin B to complex $X_{short} \cdot A$ (X_{short} corresponds to the sequence of the detected region on DsRed mRNA). To prevent Dicer from cleaving the starting and intermediate

structures, we made the molecules shorter than a standard Dicer substrate and/or used 2'-OMe modifications.

Following initial *in vitro* studies (data not shown) the stem of hairpin C was shortened by two nucleotides in order to obtain a longer toehold with the aim of improving conversion. We also changed the 2'-OMe modification pattern of hairpin B. In the original design the toehold of hairpin B was 2'-OMe while the rest of the hairpin was made of RNA. In the second iteration, the toehold of hairpin B was changed to RNA and five bases at the 3' end of the stem of hairpin B were changed to 2'-OMe. With this change we hoped to achieve two goals: improve conversion and reduce leakage. When the toehold of hairpin B binds to hairpin A we make 2'-OMe:RNA base-pairs rather than 2'-OMe:2'-OMe base-pairs, potentially making a stronger bond which should improve our conversion. By placing 2'-OMe bases at the stem we change one end of the stem from RNA:RNA to RNA:2'-OMe, again making a tighter bond [48, 49], which should reduce leakage. Alternative modification patterns for all three hairpins were also explored (data not shown); we chose to proceed with modifications which resulted in the best ON:OFF ratios and did not lead to unwanted Dicer cleavage of reactants.

Hairpin A is made entirely of 2'-OMe, hairpin B has a stem which is partially modified with 2'-OMe and the stem of hairpin C is half RNA and half 2'-OMe. The sequences listed in Table 2.1 are the final versions used to generate the data presented in this chapter.

2.4 Results

2.4.1 *In vitro* studies

2.4.1.1 Study of ON:OFF properties of triggered Dicer substrate formation

A good mechanism is turned OFF in the absence of a detection target (hairpins do not interact), while forming the final product (Dicer substrate) in its ON state upon the presence of a detection target. We demonstrate the different steps of our mechanism using

native polyacrylamide gel electrophoresis (Figure 2.2(a) and Figure A.1 in Appendix A). In the OFF state, a minimal amount of Dicer substrate B·C is produced. The OFF state is represented either using no target (lane 4, Figure 2.2(a)), an off-target mRNA Z (GAPDH mRNA, lane 9, Figure 2.2(a)) or the output silencing target Y (d2EGFP mRNA, lane 10, Figure 2.2(a)). The silencing target was used as an OFF state measurement since both B and C hairpins contain sequences that are complementary to d2EGFP and we wanted to verify that the output target cannot turn the mechanism ON by reversing the steps of the mechanism. In the ON state, the Dicer substrate B·C is produced. The ON state is demonstrated using both a short synthetic target (X_{short}) and full-length DsRed mRNA (X) (lanes 7,8 of Figure 2.2(a), respectively). The amount of B·C formed in the ON and OFF states is quantified and normalized relative to production using X_{short} (Figure 2.2(a) and Figure A.2 in Appendix A).

2.4.1.2 Study of the catalytic properties of the mechanism

To demonstrate the catalytic property of the mechanism we used sub-stoichiometric amounts of X_{short} to trigger the formation of complex B·C. While the consumption of hairpin A is limited by the amount of X_{short} , if the mechanism is indeed catalytic then B and C will still be consumed. As expected, the amount of B·C formed is greater than the amount of X_{short} present in the reaction; B and C are nearly consumed with as little as $0.3 \times X_{short}$ and roughly 50% are consumed with $0.1 \times X_{short}$ within two hours (Figure 2.2(b), left to right and Figure A.3 in Appendix A). The amount of B·C formed is quantified relative to B·C production using $1 \times X_{short}$.

2.4.1.3 Dicer cleavage assays

The final product of our mechanism, complex B·C, was designed to be processed by Dicer to produce an siRNA. We show here that indeed the Dicer substrate can be cut by Dicer while the initial and intermediate components remain uncut. Components in the absence

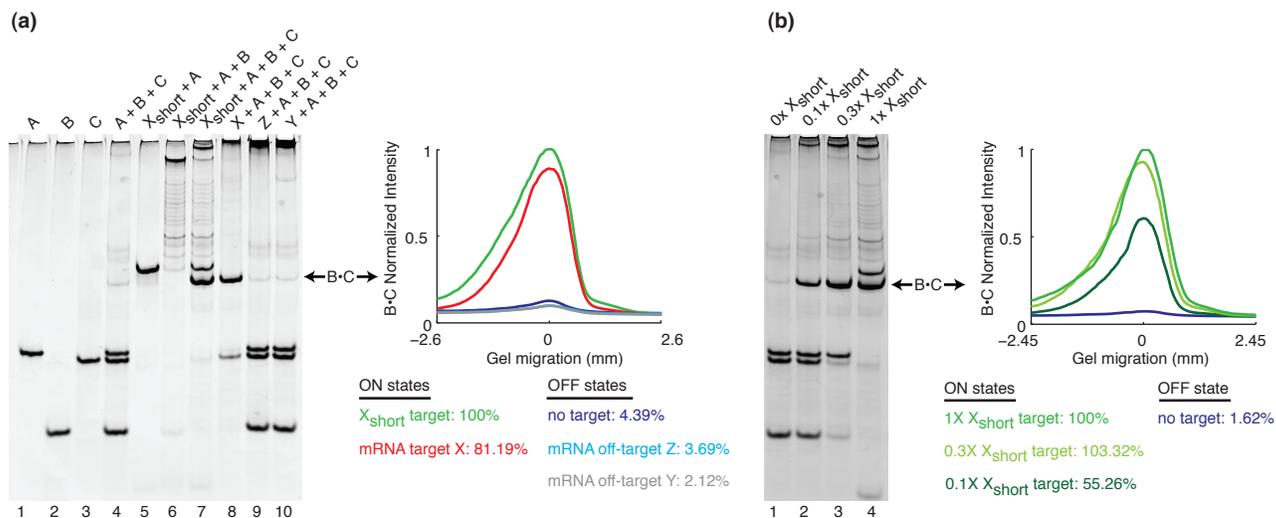


Figure 2.2: **Conditional catalytic DsiRNA formation in a test tube.** (a) Conditional Dicer substrate formation *in vitro*. ON and OFF states of the mechanism in the presence of a short synthetic target X_{short} (lane 7), DsRed mRNA target X (lane 8), no target (lane 4), GAPDH off-target mRNA Z (lane 9) and silencing off-target mRNA d2EGFP Y (lane 10). (b) Catalytic Dicer substrate formation. Sub-stoichiometric amounts of target X_{short} generate above stoichiometric amounts of B·C (target levels increase between lanes 1 through 4). All the lanes have a $0.5\mu\text{M}$ mixture of hairpins A, B and C. Plots represent the normalized intensity of BC in each lane. Quantification was done as described in methods.

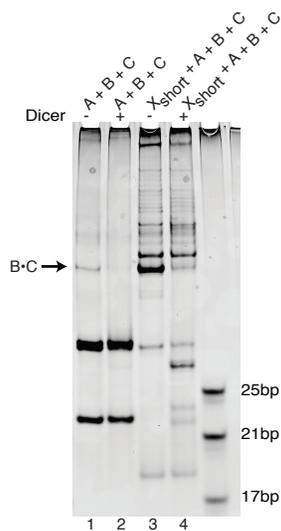


Figure 2.3: **Conditional siRNA production *in vitro*.** In the OFF state (lane 1), hairpins A, B and C are not cleaved by Dicer (lane 2). In the ON state, the presence of X_{short} leads to formation of the Dicer substrate B·C (lane 3). B·C is cleaved by Dicer to produce siRNAs (lane 4). (-/+) indicate absence/presence of Dicer.

or presence of X_{short} were subjected to an *in vitro* Dicer cleavage assay. In the OFF state, none of the hairpins are processed by Dicer (Lane 2 in Figure 2.3), whereas in the ON state, bands corresponding to an siRNA and higher molecular weight leftover products are formed while the B·C band disappears (lane 4 in Figure 2.3). None of the other intermediate complexes are processed by Dicer *in vitro* (see Figure A.4 in Appendix A).

2.4.1.4 Cell studies

We examine here the functionality of our mechanism in tissue culture. The sequences of our mechanism were designed to detect DsRed and silence d2EGFP. The ON state of the mechanism can be examined in cells expressing both DsRed and d2EGFP while the OFF state can be examined in cells expressing only d2EGFP. Ideally, d2EGFP levels should be knocked down in DsRed/d2EGFP expressing cells and remain unchanged in d2EGFP cells.

We analyze our mechanism by examining the effect of each component on its own, the full mechanism, a full mechanism with a short synthetic target as well as a DsiRNA control against the same d2EGFP region. As can be seen in Figure 2.4(a), transfection of hairpin A on its own results in DsRed down-regulation and d2EGFP up-regulation. Hairpin A is designed to detect the target DsRed, the binding of hairpin A to DsRed may result in DsRed knockdown by means of antisense or another pathway. RNAi is less likely since hairpin A is entirely modified with 2'-OMe. The up-regulation of d2EGFP in response to DsRed down-regulation might be due to more translation of d2EGFP now that DsRed is not being generated. Hairpin B on its own seems to lead to some knockdown of d2EGFP. *In vitro* Dicer cleavage assays show that hairpin B is not cut by Dicer (Figure 2.3 and Figure A.4 in Appendix A). It is possible that the cellular environment allows hairpin B to be cut by Dicer, or silencing may be mediated without Dicer processing (see discussion). Hairpin C on its own does not significantly down-regulate DsRed or d2EGFP, as expected. When the full mechanism (A+B+C) is introduced into cells DsRed is down-regulated and surprisingly d2EGFP is highly up-regulated. Transfection of a full mechanism with a short synthetic target ($X_{short}+A+B+C$), of an annealed final product B·C or of a DsiRNA did not result in d2EGFP knockdown.

When designing the mechanism, the d2EGFP silencing region was selected by the NUPACK design code to optimize the ensemble defect of the mechanism and without using prior knowledge regarding the effectiveness of the silenced region. Unfortunately, as is evident from Figure 2.4(a), this region is not amenable to RNAi. Lack of silencing is due to sequence and not transfection conditions as is exemplified by comparing DsiRNA to DsiRNA2 in Figure Figure 2.4(b).

Due to the choice of d2EGFP silencing region, the current design was not capable of leading to d2EGFP knockdown even if a Dicer substrate is formed once transfected. We therefore modified the sequence of the design to detect the same DsRed region but silence a d2EGFP region which works well according to data in the lab (for sequences see

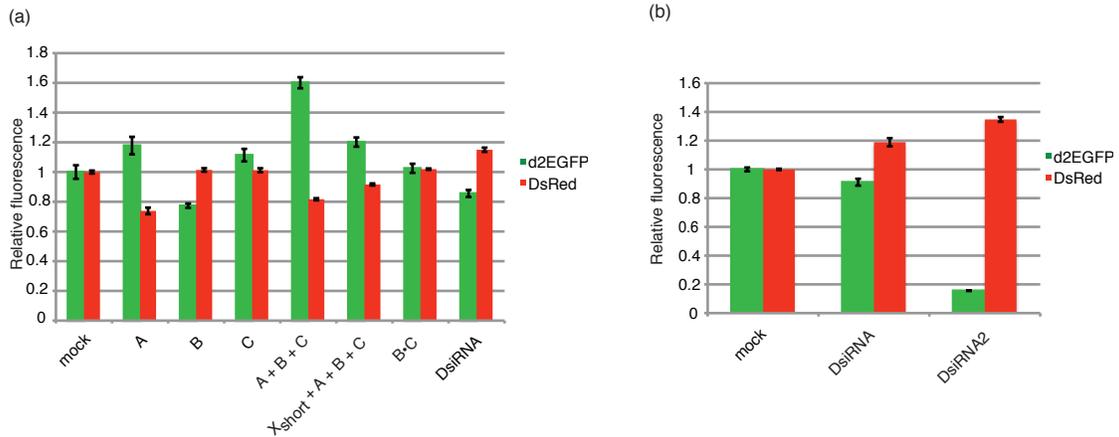


Figure 2.4: **Transfection of conditional Dicer substrate formation into HEK293 d2EGFP DsRed cells.** Reverse transfection of 20pmol of each oligonucleotide or annealed B·C. Flow cytometry was used to determine fluorescence. Mean fluorescence was normalized relative to mock treated cells. Green bars represent relative d2EGFP fluorescence, red bars represent relative DsRed fluorescence. Error bars represent the standard error of the mean of three samples. **(a)** Transfection of mechanism components. B·C was annealed prior to transfection. **(b)** d2EGFP silencing via DsiRNA transfection. DsiRNA targets the same d2EGFP region as the mechanism, DsiRNA2 was used as a control. See table 2.2 for regions and sequences.

Table 2.3). The ON and OFF states of the mechanism were compared by transfecting the same mixture into cells lacking or expressing the DsRed detection target (Figure 2.5(a) and (b), respectively). In both cell lines, transfection of a pre-annealed Dicer substrate B2·C2 resulted in $\sim 80\text{-}90\%$ d2EGFP knockdown, demonstrating that the final product is functional and that the chosen d2EGFP silencing region works well. Transfection of the three hairpins with a short synthetic target ($X_{short}+A2+B2+C2$) resulted in $\sim 40\%$ d2EGFP knockdown in d2EGFP cells and $\sim 60\%$ d2EGFP knockdown in d2EGFP DsRed cells. X_{short} was added to the hairpins immediately before the transfection reagent was added to the mix so the reaction components were pre-incubated for a minimal time prior to transfection. Comparing the silencing efficiency of B2·C2 relative to $X_{short}+A2+B2+C2$ suggests that the reaction with X_{short} does not go to completion and less B2·C2 is formed. When the hairpin components of the mechanism are transfected ($A2+B2+C2$), $\sim 15\%$ d2EGFP knockdown is observed in d2EGFP cells (similar to B2 transfection on its own) and $\sim 40\%$ d2EGFP knockdown in d2EGFP DsRed cells (Figure 2.5(a) and (b), respectively). In cells expressing both d2EGFP and DsRed both B2 and C2 hairpins knockdown d2EGFP by $\sim 20\%$ (Figure 2.5(b)).

Next, we examine whether the $\sim 40\%$ d2EGFP knockdown observed by $A2+B2+C2$ transfection into cells expressing for d2EGFP and DsRed is due to the mechanism being turned ON. Due to the variability of the d2EGFP silencing efficiency between the two cell lines we chose to use the DsRed expressing cell line for this study. Unfortunately, a $\sim 20\%$ variability in d2EGFP knockdown exists between the two transfections of the ON state ($A2+B2+C2$ samples in Figures 2.5(b) and (c)) making it difficult to compare the separate experiments. Still, d2EGFP appears to be down-regulated by transfection of the un-annealed B2 and C2 hairpins (sample B2 + C2 in Figure 2.5(c)). It is not clear if this is due to leakage or a cumulative effect of each hairpin silencing d2EGFP on its own.

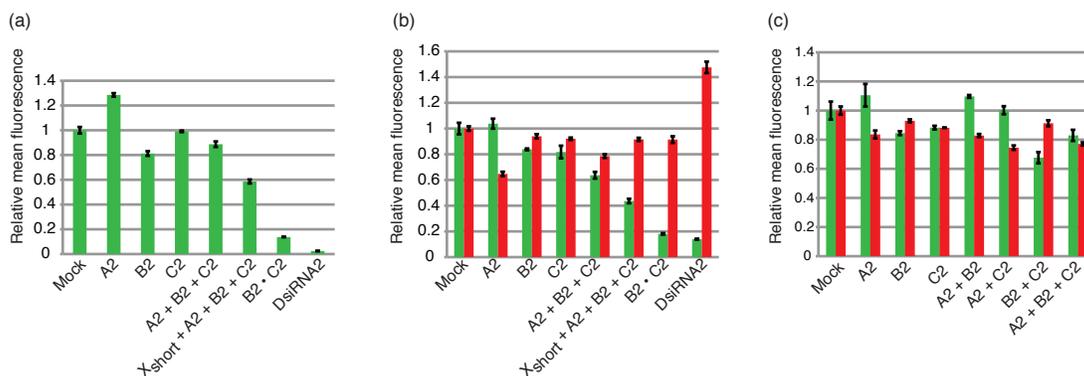


Figure 2.5: **Conditional Dicer substrate formation is not functional *in vivo*.** Reverse transfection of 20pmol of each oligonucleotide or annealed B2·C2. Flow cytometry was used to determine fluorescence. Mean fluorescence was normalized relative to mock treated cells. Green bars represent relative d2EGFP fluorescence, red bars represent relative DsRed fluorescence. Error bars represent the standard error of the mean of three samples. (a) Validation of OFF state. Transfection of mechanism components into cells lacking the DsRed detection target. B2·C2 was annealed prior to transfection. (b) Validation of ON state. Transfection of mechanism components into cells expressing the DsRed detection target. B2·C2 was annealed prior to transfection. (c) Transfection of leakage controls into cells expressing DsRed and d2EGFP.

2.5 Discussion

In this chapter we presented a catalytic mechanism that conditionally produces a Dicer substrate upon the detection of an mRNA target. While traditional RNAi implements the logic *silence gene Y*, the mechanism described in this chapter implements the logic *if gene X is detected then produce a Dicer substrate targeting independent gene Y*. Our mechanism is comprised of three scRNA hairpins which form a signal transduction cascade in which the detection of an mRNA target X results in conformational change of the hairpins leading to formation of a Dicer substrate targeting independent gene Y.

We demonstrated conditional Dicer substrate formation in a test tube. In the absence of a detection target the hairpins do not interact and minimal Dicer substrate is formed. Upon the presence of a short synthetic target or a full-length mRNA detection target the output of the mechanism results in a Dicer substrate. The mechanism is catalytic, detection of sub-stoichiometric amounts of target result in production of Dicer substrate with observed turnover of approximately 100% with as little as $0.3\times$ target and roughly 50% turnover is observed with as little as $0.1\times$ target.

The produced Dicer substrate was inspired by DsiRNAs yet it is a non-canonical Dicer substrate. Like a DsiRNA, it has a two nucleotide 3' overhang on one end. The purpose of this overhang is to introduce a preference for Dicer processing to start from that end. Unlike a DsiRNA, its opposite end has a 5' overhang instead of a blunt end. Despite being a non-canonical substrate, it is cleaved by recombinant Dicer *in vitro*. Furthermore, when transfected into cells it results in efficient gene knockdown, comparable to that of a DsiRNA.

The silencing observed by the generated non-canonical Dicer substrate suggests that the output of a functioning mechanism should result in gene knockdown. To date, we have not been successful in achieving conditional RNAi in cells. Several factors can contribute to our mechanism being non-functional *in vivo*. For the mechanism to work, three hairpins need to be co-delivered into cells. Our current delivery strategy might be insufficient;

alternative delivery methods need to be explored. One such possibility is expressing the scRNAs off a plasmid, this is further discussed in Chapter 3.

Our hairpins were designed to be small as well as chemically modified so that they are not cleaved by Dicer. *In vitro* studies show that we were successful in this goal. Yet, some silencing was observed by single hairpins in cells. It is possible that silencing is mediated by non-Dicer pathways. Other than achieving gene-knockdown, it could also be that the hairpins are bound or degraded by cellular proteins and are therefore not available to the mechanism. This is discussed in further detail in Chapter 3.

For signal transduction to occur, the mechanism relies on toehold-mediated branch migration. The toehold dimensions and/or effective concentration may not be sufficiently high. Some DsRed knockdown is observed by hairpins A and A2 (Figure 2.4 and 2.5, respectively). Both hairpins are fully modified with 2'-OMe and therefore are not expected to be enzymatically processed. Therefore, we postulate that DsRed silencing is observed via an antisense mechanism, suggesting that the hairpin is bound to the DsRed mRNA and that toehold-mediated branch migration is observed. Further studies are needed to validate this hypothesis.

The selected detection and silencing targets used are fluorescent proteins. Interestingly, in some cases when d2EGFP expression is down-regulated, DsRed expression is up-regulated and vice versa (see Figures 2.4 and 2.5). One explanation would be that the mRNA of one fluorescent protein is degraded allowing for more translation of the other fluorescent protein. However, this phenomenon is not consistent across different transfected samples. For example, in Figure 2.5(b) transfection of DsiRNA2 results in d2EGFP knockdown and DsRed up-regulation while transfection of annealed B2·C2 down-regulates d2EGFP to the same extent of sample DsiRNA2 but DsRed expression does not appear to change. The cause for this discrepancy could be activation of an immune response such as protein kinase R (PKR) by B2·C2 but not by DsiRNA2. Activation of PKR would inhibit translation [50, 51], potentially causing DsRed expression to remain unchanged. Further

studies are needed to validate this hypothesis.

While still facing challenges *in vivo*, implementing conditional RNAi has profound applications as both a research tool and as a therapeutic agent. As a research tool, triggered RNAi mechanisms will allow gene Y to be silenced in a specific tissue or at a specific developmental stage by appropriately selecting gene X. Alternatively, the spatio-temporal expression of any gene can be reported visually by specifying gene Y as a fluorescent protein or a regulator of a fluorescent protein. As a therapeutic, triggered RNAi could potentially treat any disease that is encoded genetically and could benefit from down-regulation of gene expression or specific cell death such as cancer or autoimmune diseases. For example, to treat cancers, it could be possible to detect an mRNA cancer marker and silence a housekeeping gene to kill the diseased cell.

The mechanism proposed in this work produces a dsRNA substrate designed to interact with the RNAi pathway. However, it is not limited to RNAi. By altering the final product this mechanism has the potential to (conditionally) interface with biology through other means. The final product can have immunostimulatory effects through induction of proinflammatory cytokines and type I interferon via interaction with receptors such as RIG-I and TLRs etc [52–54]. The final product can also be designed to serve a double function as both immunostimulatory and gene downregulator [55]. Furthermore, gene knockdown is not limited to eukaryotes encoding the RNAi pathway. It has recently been discovered that bacteria and archaea have nucleic acid based adaptive immune systems termed CRISPR. This system relies on small RNAs for sequence specific silencing of foreign nucleic acids [56, 57]. As understanding of this system grows it is becoming evident that CRISPR, in a similar fashion to RNAi, can be programmed as well [58, 59]. The above described mechanisms may offer the potential to conditionally knock down genes not only in eukaryotes but in bacteria and archaea as well.

2.6 Materials and methods

Oligonucleotides. Oligonucleotides were synthesized and HPLC purified by Integrated DNA Technologies (IDT). Strands were diluted to the desired concentration in $1\times$ duplex buffer (20 mM Hepes, pH 7.5, 100 mM Potassium Acetate). Oligonucleotide concentrations were determined and adjusted using A_{260} absorbance on a NanoDrop8000 (Thermo Scientific). Further adjustments were performed by incubating different ratios of individual strands for 2 hours at 37°C followed by gel electrophoresis until correct stoichiometry was obtained.

Hairpins were snap cooled by heating them to 95°C for 90 seconds followed by a 30 second incubation on ice and room temperature incubation of at least 30 minutes. Complexes were annealed by heating to 90°C for 3 minutes followed by a controlled gradual cooling at -1°C per minute to 23°C in a PCR block.

Oligonucleotide sequences. For a list of sequences see Tables 2.1, 2.2, 2.3.

To separate $X_{short}\cdot A$ from $B\cdot C$ on a 20% native polyacrylamide gel, a target longer than 27 nucleotides is needed. Therefore, a 33-nucleotides-long target was used. Three bases were added to the 5' end of the target and three bases were added to the 3' end of the target (see Table 2.1). These extra bases do not affect the properties of our mechanism (data not shown).

Polyacrylamide gel electrophoresis. Hairpins and X_{short} were used at $0.5\mu\text{M}$ each, mRNA targets were used at $1\mu\text{M}$. Reactions were carried out for two hours at 37°C in $1\times$ duplex buffer. 20% native polyacrylamide gels were cast and run in $1\times$ TBE (Tris-Borate-EDTA) at 200V. Denaturing polyacrylamide gels were cast and run in $1\times$ TBE at 500V unless otherwise specified. Denaturing gels were pre-run at 500V for 1–2hr (unless otherwise specified). Gels were stained in $1\times$ SYBR Gold (Life Technologies) for 10 minutes at room temperature and imaged using an FLA-5100 imaging system (Fuji Photo Film).

Quantification and band intensity plots. Multi Gauge ver3.0 (Fujifilm) software was used for quantification and intensity plot data. Bands were quantified using the “Quant

Strand	Sequence
X_{short}	GGCAAGCUGGACAUCACCUCCCACAACGAGGAC
A	UCACCUCCCACAACGCUUCAAGUCCGCCAUCUCUCGUUGUGGGAGGU GAUGUCCAGCUU
B	UCAAGUCCGCCAUGCCCG CAACGAUGGCGGACUUGAAGCGUUG
C	CGCCAUGCCCGCAACGCUUCAAGUCCGCCAUCGUUG*CGGGCAUGGCG GACUUGAAG

Table 2.1: List of strands for triggered Dicer substrate formation mechanism. DsRed region: 592–618, d2EGFP region: 252–271. In red are sequences corresponding to DsRed, in green are sequences corresponding to EGFP, in black are random bases. 2'-OMe modifications are underlined. * The nucleotide at the 5' end of the guide strand is part of the DsRed coding sequence. This is not part of the overall design but rather a result of trimming of hairpin C (see Section 2.3). Since silencing potency was not affected by this, we left it as is.

Strand	Sequence
DsiRNA sense	UCAAGUCCGCCAUGCCCGCAACGAU
DsiRNA antisense	AUCGUUGCGGGCAUGGCGGACUUGAAG
DsiRNA2 sense	UACGGCAAGCUGACCCUGAAGUCUC
DsiRNA2 antisense	GAGACUUCAGGGUCAGCUUGCCGUACA

Table 2.2: List of DsiRNA sequences. DsiRNA targets region 252–271 of d2EGFP. DsiRNA2 targets region 118–140 of d2EGFP.

Strand	Sequence
X_{short}	GGCAAGCUGGACAUCACCUCCCACAACGAGGAC
A2	UCACCUCCCACAACGUGACCCUGAAGUCCACUCUCGUUGUGGGAGGU GAUGUCCAGCUU
B2	GACCCUGAAGUUCAUCUGCAACGGAACUUCAGGGUCAGCGUUG
C2	AAGUUCAUCUGCAACGUGACCCUGAAGUCCGUUG*CAGAUGAACUU CAGGGUCAG

Table 2.3: List of strands for triggered Dicer substrate formation mechanism 2. DsRed region: 592–618, d2EGFP region: 127–148. In red are sequences corresponding to DsRed, in green are sequences corresponding to EGFP, in black are random bases. 2'-OMe modifications are underlined. * The nucleotide at the 5' end of the guide strand is part of the DsRed coding sequence. This is not part of the overall design but rather a result of trimming of hairpin C (see Section 2.3). Since silencing potency was not affected by this, we left it as is.

Measure mode.” Data points for band intensity plots were gathered using the profile feature. ON-to-OFF ratio was determined by setting the ON ratio with a short target to 100%.

***In vitro* Dicer assay.** Dicer reactions were performed using the Recombinant Human Turbo Dicer Enzyme kit (Genlantis) according to the manufacturer with some modifications. Reactions were performed at $0.5\mu\text{M}$ in $10\mu\text{L}$ using 0.5 unit of turbo Dicer. Hairpins were snap cooled prior to Dicer reaction. Dicer, target and reactants were all mixed at the same time (i.e., the reactants were not pre-incubated with their target prior to addition of Dicer). Dicer reactions were carried out for 2 hours at 37°C , reactions were stopped by the addition of the appropriate loading dye. siRNA formation was determined by polyacrylamide gel electrophoresis.

Cell lines and transfections. HEK293 d2EGFP cells were a generous gift from Dr. Chase Beisel. The destabilized EGFP sequence comes from pd2EGFP-1 plasmid (Clontech, PT3205-5 catalog #6008-1). HEK293 d2EGFP DsRed cells were generated by a stable transfection of pDsRed2-1-C1 (Clontech, PT3603-5 catalog #632407) into HEK293 d2EGFP cells. Cells were maintained at 37°C 5% CO_2 in DMEM (Invitrogen) supplemented with 10% fetal bovine serum (Invitrogen). Transfections were carried out at the specified oligonucleotide concentrations using Lipofectamine RNAiMAX transfection reagent (Invitrogen) according to the reverse transfection protocol. $1-2 \times 10^5$ cells were plated per well of a 24-well plate. Cell counts and viability were determined using the Countess automated cell counter according to the manufacturer (Invitrogen).

Flow cytometry. 24–48 hours post transfection samples were trypsinized for 5 minutes at 37°C , the trypsin was then quenched with DMEM growth media supplemented with 10% FBS. Samples were run on a BD Accuri C6 flow cytometer (BD Biosciences). Healthy cells were gated according to their scatter using untreated cells. Results represent the normalized mean and error bars represent the standard deviation of the mean of three transfected samples.

pTnT-DsRed construction. The DsRed mRNA coding sequence was amplified off

pDsRed2-C1 (Clontech, catalog #632407) using Taq DNA polymerase. The forward and reverse primers included the MluI and NotI restriction sites (respectively) for directional cloning. The DsRed mRNA coding sequence was cloned into pTnT vector (Promega, catalog #L5610) between the MluI and NotI restriction sites. The construct was verified by sequencing.

DsRed mRNA sequence

```

1 ATGGCCTCCT CCGAGAACGT CATCACCGAG TTCATGCGCT TCAAGGTGCG CATGGAGGGC
61 ACCGTGAACG GCCACGAGTT CGAGATCGAG GCGGAGGGCG AGGGCCGCC CTACGAGGGC
121 CACAACACCG TGAAGCTGAA GGTGACCAAG GCGGGCCCC TGCCCTTCGC CTGGGACATC
181 CTGTCCCCC AGTTCAGTA CGGCTCCAAG GTGTACGTGA AGCACCCCGC CGACATCCCC
241 GACTACAAGA AGCTGTCCTT CCCCAGGGC TTCAAGTGGG AGCGCGTGAT GAACTTCGAG
301 GACGGCGGCG TGGCGACCGT GACCCAGGAC TCCTCCCTGC AGGACGGCTG CTCATCTAC
361 AAGGTGAAGT TCATCGGCGT GAACTTCCCC TCCGACGGCC CCGTGATGCA GAAGAAGACC
421 ATGGGCTGGG AGGCCTCCAC CGAGCGCCTG TACCCCGCG ACGGCGTGCT GAAGGGCGAG
481 ACCCACAAGG CCCTGAAGCT GAAGGACGGC GGCCACTACC TGGTGGAGTT CAAGTCCATC
541 TACATGGCCA AGAAGCCCGT GCAGCTGCCC GGCTACTACT ACGTGGACGC CAAGCTGGAC
601 ATCACCTCCC ACAACGAGGA CTACACCATC GTGGAGCAGT ACGAGCGCAC CGAGGGCCGC
661 CACCACCTGT TCCTGAGATC TCGAGCTCAA GCTTCGAATT CTGCAGTCGA CGGTACCGCG
721 GGCCCGGGAT CCACCGGATC TAGATAA

```

pGEM-T easy-d2EGFP construction. The d2EGFP mRNA coding sequence was cloned from cells expressing d2EGFP (generous gift from Dr. Beisel) based on the pd2EGFP-1 (Clontech, catalog #6008-1) sequence and cloned into pGEM-T easy vector (Promega, catalog #A1360)

d2EGFP mRNA sequence

```

1 ATGGTGAGCA AGGGCGAGGA GCTGTTCCACC GGGGTGGTGC CCATCCTGGT CGAGCTGGAC

```

61 GCGGACGTAA ACGGCCACAA GTTCAGCGTG TCCGGCGAGG GCGAGGGCGA TGCCACCTAC
 121 GGCAAGCTGA CCCTGAAGTT CATCTGCACC ACCGGCAAGC TGCCCGTGCC CTGGCCCACC
 181 CTCGTGACCA CCCTGACCTA CGGCGTGCAG TGCTTCAGCC GCTACCCCGA CCACATGAAG
 241 CAGCACGACT TCTTCAAGTC CGCCATGCCC GAAGGCTACG TCCAGGAGCG CACCATCTTC
 301 TTCAAGGACG ACGGCAACTA CAAGACCCGC GCCGAGGTGA AGTTCGAGGG CGACACCCTG
 361 GTGAACCGCA TCGAGCTGAA GGGCATCGAC TTCAAGGAGG ACGGCAACAT CCTGGGGCAC
 421 AAGCTGGAGT ACAACTACAA CAGCCACAAC GTCTATATCA TGGCCGACAA GCAGAAGAAT
 481 GGCATCAAGG TGAACTTCAA GATCCGCCAC AACATCGAGG ACGGCAGCGT GCAGCTCGCC
 541 GACCACTACC AGCAGAACAC CCCCATCGGC GACGGCCCCG TGCTGCTGCC CGACAACCAC
 601 TACCTGAGCA CCCAGTCCGC CCTGAGCAAA GACCCCAACG AGAAGCGCGA TCACATGGTC
 661 CTGCTGGAGT TCGTGACCGC CGCCGGGATC ACTCTCGGCA TGGACGAGCT GTACAAGAAG
 721 CTTAGCCATG GCTTCCCGCC GGAGGTGGAG GAGCAGGATG ATGGCACGCT GCCCATGTCT
 781 TGTGCCCAGG AGAGCGGGAT GGACCGTCAC CCTGCAGCCT GTGCTTCTGC TAGGATCAAT
 841 GTGTAG

mRNA *in vitro* transcription. pGEMTeasy-GAPDH was a gift from Lisa Hochrein. Plasmids were linearized by restriction digestion. DsRed was *in vitro* transcribed using T7-Scribe Standard RNA IVT Kit (CELLSCRIPT) according to the manufacturer. d2EGFP and GAPDH were *in vitro* transcribed using SP6-Scribe Standard RNA IVT Kit (CELLSCRIPT) according to the manufacturer. Transcribed mRNA was purified using RNeasy Protect Mini Kit (Qiagen) according to the manufacturer. mRNA concentration was determined using A_{260} absorbance on a NanoDrop8000 (Thermo Scientific).

Bibliography

- [1] Isaacs, F. J., Dwyer, D. J., and Collins, J. J. RNA synthetic biology. *Nature Biotechnology* **24**(5), 545–54 (2006).
- [2] Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**(6669), 806–11 (1998).
- [3] Hannon, G. J. RNA interference. *Nature* **418**(6894), 244–51 (2002).
- [4] Hannon, G. J. and Rossi, J. J. Unlocking the potential of the human genome with RNA interference. *Nature* **431**(7006), 371–8 (2004).
- [5] Bumcrot, D., Manoharan, M., Koteliansky, V., and Sah, D. W. RNAi therapeutics: a potential new class of pharmaceutical drugs. *Nature chemical biology* **2**(12), 711–9 (2006).
- [6] Kim, D. H. and Rossi, J. J. Strategies for silencing human disease using RNA interference. *Nature reviews. Genetics* **8**(3), 173–84 (2007).
- [7] Hamilton, A. J. and Baulcombe, D. C. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* **286**(5441), 950–2 (1999).
- [8] Hammond, S. M., Bernstein, E., Beach, D., and Hannon, G. J. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* **404**(6775), 293–6 (2000).
- [9] Zamore, P. D., Tuschl, T., Sharp, P. A., and Bartel, D. P. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**(1), 25–33 (2000).
- [10] Elbashir, S. M., Lendeckel, W., and Tuschl, T. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes & development* **15**(2), 188–200 (2001).

- [11] Bernstein, E., Caudy, A. A., Hammond, S. M., and Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**(6818), 363–6 (2001).
- [12] Ketting, R. F., Fischer, S. E., Bernstein, E., Sijen, T., Hannon, G. J., and Plasterk, R. H. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes & development* **15**(20), 2654–9 (2001).
- [13] Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**(6836), 494–8 (2001).
- [14] Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R., and Tuschl, T. Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* **110**(5), 563–74 (2002).
- [15] Liu, J., 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. Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**(5689), 1437–41 (2004).
- [16] Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Molecular Cell* **15**(2), 185–97 (2004).
- [17] Rand, T. A., Petersen, S., Du, F., and Wang, X. Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. *Cell* **123**(4), 621–9 (2005).
- [18] Rivas, F. V., Tolia, N. H., Song, J. J., Aragon, J. P., Liu, J., Hannon, G. J., and Joshua-Tor, L. Purified Argonaute2 and an siRNA form recombinant human RISC. *Nature structural & molecular biology* **12**(4), 340–9 (2005).

- [19] Pratt, A. J. and MacRae, I. J. The RNA-induced silencing complex: a versatile gene-silencing machine. *The Journal of biological chemistry* **284**(27), 17897–901 (2009).
- [20] Hutvagner, G. and Zamore, P. D. A microRNA in a multiple-turnover RNAi enzyme complex. *Science* **297**(5589), 2056–60 (2002).
- [21] Oliveira, S., Storm, G., and Schifflers, R. M. Targeted delivery of siRNA. *Journal of biomedicine & biotechnology* **2006**(4), 63675 (2006).
- [22] Tiscornia, G., Tergaonkar, V., Galimi, F., and Verma, I. M. CRE recombinase-inducible RNA interference mediated by lentiviral vectors. *Proceedings of the National Academy of Sciences of the United States of America* **101**(19), 7347–51 (2004).
- [23] Gupta, S., Schoer, R. A., Egan, J. E., Hannon, G. J., and Mittal, V. Inducible, reversible, and stable RNA interference in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* **101**(7), 1927–32 (2004).
- [24] Wiznerowicz, M., Szulc, J., and Trono, D. Tuning silence: conditional systems for RNA interference. *Nature methods* **3**(9), 682–8 (2006).
- [25] Cullere, X., Lauterbach, M., Tsuboi, N., and Mayadas, T. N. Neutrophil-selective CD18 silencing using RNA interference in vivo. *Blood* **111**(7), 3591–8 (2008).
- [26] Lee, S. K. and Kumar, P. Conditional RNAi: towards a silent gene therapy. *Advanced drug delivery reviews* **61**(7-8), 650–64 (2009).
- [27] An, C. I., Trinh, V. B., and Yokobayashi, Y. Artificial control of gene expression in mammalian cells by modulating RNA interference through aptamer-small molecule interaction. *RNA* **12**(5), 710–6 (2006).
- [28] Beisel, C. L., Bayer, T. S., Hoff, K. G., and Smolke, C. D. Model-guided design of ligand-regulated RNAi for programmable control of gene expression. *Molecular systems biology* **4**, 224 (2008).

- [29] Tuleuova, N., An, C. I., Ramanculov, E., Revzin, A., and Yokobayashi, Y. Modulating endogenous gene expression of mammalian cells via RNA-small molecule interaction. *Biochemical and biophysical research communications* **376**(1), 169–73 (2008).
- [30] Kumar, D., An, C. I., and Yokobayashi, Y. Conditional RNA interference mediated by allosteric ribozyme. *Journal of the American Chemical Society* **131**(39), 13906–7 (2009).
- [31] Beisel, C. L., Chen, Y. Y., Culler, S. J., Hoff, K. G., and Smolke, C. D. Design of small molecule-responsive microRNAs based on structural requirements for Droscha processing. *Nucleic acids research* **39**(7), 2981–94 (2011).
- [32] Xie, Z., Liu, S. J., Bleris, L., and Benenson, Y. Logic integration of mRNA signals by an RNAi-based molecular computer. *Nucleic acids research* **38**(8), 2692–701 (2010).
- [33] Masu, H., Narita, A., Tokunaga, T., Ohashi, M., Aoyama, Y., and Sando, S. An activatable siRNA probe: trigger-RNA-dependent activation of RNAi function. *Angewandte Chemie* **48**(50), 9481–3 (2009).
- [34] Kumar, D., Kim, S. H., and Yokobayashi, Y. Combinatorially inducible RNA interference triggered by chemically modified oligonucleotides. *Journal of the American Chemical Society* **133**(8), 2783–8 (2011).
- [35] Carthew, R. W. and Sontheimer, E. J. Origins and Mechanisms of miRNAs and siRNAs. *Cell* **136**(4), 642–55 (2009).
- [36] Zeng, Y. and Cullen, B. R. Efficient processing of primary microRNA hairpins by Droscha requires flanking nonstructured RNA sequences. *The Journal of biological chemistry* **280**(30), 27595–603 (2005).
- [37] Kim, D. H., Behlke, M. A., Rose, S. D., Chang, M. S., Choi, S., and Rossi, J. J. Syn-

- thetic dsRNA Dicer substrates enhance RNAi potency and efficacy. *Nature Biotechnology* **23**(2), 222–6 (2005).
- [38] Yin, P., Choi, H. M., Calvert, C. R., and Pierce, N. A. Programming biomolecular self-assembly pathways. *Nature* **451**(7176), 318–22 (2008).
- [39] Rose, S. D., Kim, D. H., Amarzguioui, M., Heidel, J. D., Collingwood, M. A., Davis, M. E., Rossi, J. J., and Behlke, M. A. Functional polarity is introduced by Dicer processing of short substrate RNAs. *Nucleic acids research* **33**(13), 4140–56 (2005).
- [40] Collingwood, M. A., Rose, S. D., Huang, L., Hillier, C., Amarzguioui, M., Wiüger, M. T., Soifer, H. S., Rossi, J. J., and Behlke, M. A. Chemical modification patterns compatible with high potency dicer-substrate small interfering RNAs. *Oligonucleotides* **18**(2), 187–200 (2008).
- [41] Macrae, I. J., Zhou, K., Li, F., Repic, A., Brooks, A. N., Cande, W. Z., Adams, P. D., and Doudna, J. A. Structural basis for double-stranded RNA processing by Dicer. *Science* **311**(5758), 195–8 (2006).
- [42] Zadeh, J. N., Steenberg, C. D., Bois, J. S., Wolfe, B. R., Pierce, M. B., Khan, A. R., Dirks, R. M., and Pierce, N. A. NUPACK: Analysis and design of nucleic acid systems. *Journal of computational chemistry* **32**(1), 170–3 (2011).
- [43] Zadeh, J. N., Wolfe, B. R., and Pierce, N. A. Nucleic acid sequence design via efficient ensemble defect optimization. *Journal of computational chemistry* **32**(3), 439–52 (2011).
- [44] Dirks, R. M., Lin, M., Winfree, E., and Pierce, N. A. Paradigms for computational nucleic acid design. *Nucleic acids research* **32**(4), 1392–403 (2004).
- [45] Dirks, R. M. and Pierce, N. A. A partition function algorithm for nucleic acid sec-

- ondary structure including pseudoknots. *Journal of computational chemistry* **24**(13), 1664–77 (2003).
- [46] Dirks, R. M. and Pierce, N. A. An algorithm for computing nucleic acid base-pairing probabilities including pseudoknots. *Journal of computational chemistry* **25**(10), 1295–304 (2004).
- [47] Dirks, R. M., Bois, J. S., M., S. J., Winfree, E., and Pierce, N. A. Thermodynamic analysis of interacting nucleic acid strands. *SIAM Review* **49**, 65–88 (2007).
- [48] Sugimoto, N., Nakano, S., Katoh, M., Matsumura, A., Nakamuta, H., Ohmichi, T., Yoneyama, M., and Sasaki, M. Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. *Biochemistry* **34**(35), 11211–6 (1995).
- [49] Majlessi, M., Nelson, N. C., and Becker, M. M. Advantages of 2'-O-methyl oligoribonucleotide probes for detecting RNA targets. *Nucleic acids research* **26**(9), 2224–9 (1998).
- [50] Manche, L., Green, S. R., Schmedt, C., and Mathews, M. B. Interactions between double-stranded RNA regulators and the protein kinase DAI. *Molecular and cellular biology* **12**(11), 5238–48 (1992).
- [51] Williams, B. R. PKR; a sentinel kinase for cellular stress. *Oncogene* **18**(45), 6112–20 (1999).
- [52] Janeway, C. A., J. and Medzhitov, R. Innate immune recognition. *Annual review of immunology* **20**, 197–216 (2002).
- [53] Yoneyama, M., Kikuchi, M., Natsukawa, T., Shinobu, N., Imaizumi, T., Miyagishi, M., Taira, K., Akira, S., and Fujita, T. The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nature immunology* **5**(7), 730–7 (2004).

- [54] Takeda, K. and Akira, S. Toll-like receptors in innate immunity. *International immunology* **17**(1), 1–14 (2005).
- [55] Schlee, M., Hornung, V., and Hartmann, G. siRNA and isRNA: two edges of one sword. *Molecular therapy: the journal of the American Society of Gene Therapy* **14**(4), 463–70 (2006).
- [56] Sorek, R., Kunin, V., and Hugenholtz, P. CRISPR—a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nature reviews. Microbiology* **6**(3), 181–6 (2008).
- [57] Horvath, P. and Barrangou, R. CRISPR/Cas, the immune system of bacteria and archaea. *Science* **327**(5962), 167–70 (2010).
- [58] Hale, C. R., Majumdar, S., Elmore, J., Pfister, N., Compton, M., Olson, S., Resch, A. M., Glover, C. V., r., Graveley, B. R., Terns, R. M., and Terns, M. P. Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs. *Molecular cell* **45**(3), 292–302 (2012).
- [59] Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**(6096), 816–21 (2012).

Chapter 3

Characterization of cellular modes of failure for conditional Dicer substrate formation in tissue culture

3.1 Introduction

In Chapter 2 we presented the design of a conditional catalytic RNAi mechanism that intended to implement the logic operation: *If gene X is detected, then silence independent gene Y*. We demonstrated that minimal Dicer substrate was formed in the absence of X while in the presence of either a short synthetic or a full-length mRNA target X the Dicer substrate for Y was generated. While the generated Dicer substrate is non-canonical, it is functional. Dicer can process the substrate *in vitro* and when transfected into cells as a pre-made complex, knockdown of Y was observed. Despite being able to design a system that is functional *in vitro*, the mechanism does not lead to gene silencing in tissue culture. The mode of failure may be due to a poor choice of sequences, a mechanistic flaw, branch migration not occurring in cells, a problem in delivery, protein sequestration and/or degradation of components or a combination thereof.

In our design process, both the hairpins as well as intermediate components were de-

signed to not be processed by Dicer. This was done by keeping hairpin stem dimensions smaller than standard shRNAs [1] and/or by using 2'-OMe chemical modifications so that the duplexes are not cleaved by Dicer. However, recent findings suggest that molecules shorter than standard siRNAs and shRNAs can also lead to gene silencing. Asymmetric interfering RNA (aiRNA) duplexes with a full antisense strand and a 15bp sense strand can lead to gene knockdown via RNAi [2–4] as can short siRNA with a 16bp duplex and two base overhangs [3]. RNAi can also be achieved by segmenting the sense strand of siRNAs into two short sequences in order to generate a small internally segmented interfering RNA (sisiRNA) [5]. shRNAs with shorter stems (16–19bp, termed sshRNA) are not processed by Dicer *in vitro*, however they do maintain their gene silencing properties [6]. It is speculated that a nuclease other than Dicer is responsible for cleavage of these sshRNAs. Supporting this hypothesis, Ago2 has been recently found to catalyze the maturation of pre-miRNA-451 [7–9]. The secondary structure of pre-miRNA-451 is similar to that of an sshRNA; it contains a 17bp stem and is not processed by Dicer.

This chapter aims to characterize the reason why triggered RNAi is not observed in tissue culture. To overcome delivery issues we constructed plasmids to express hairpins or a subset of our mechanism. The fact that many non-canonical substrates can lead to efficient RNAi raises the question whether our molecules also interact with the RNAi pathway in a non-desired manner. We examine the fate of some of our hairpins inside the cells. We address the question of whether the hairpins are being processed using northern blots and of whether they can lead to Ago2-mediated silencing using 5' rapid amplification of complementary DNA (cDNA) ends (5' RACE).

3.2 Results

We tested multiple designs for a conditional Dicer substrate formation mechanism in test tube as well as in tissue culture. While the work in Chapter 2 was done using a 3'-toehold-based mechanism, the work presented in this chapter uses hairpins both from a

5'-toehold-based mechanism (see Appendix B) as well as a 3'-toehold-based mechanism. To address the issues of poor choice of sequences and/or a mechanistic flaw we tried different designs using different subsequences of DsRed and d2EGFP, as well as testing different dimensions of toeholds, stems and loops. None of these variations resulted in a functional mechanism in tissue culture (data not shown).

3.2.1 Examination of toehold mediated branch migration in tissue culture and combatting delivery

The conditional Dicer substrate formation mechanism is complex. Three hairpins must enter the same cell, detect an endogenous mRNA target as well as interact with one another. We next tried to reduce the complexity of the mechanism by using the minimal components necessary to induce RNAi. The minimal requirements to form a Dicer substrate are the single stranded region of hairpin M1.B (termed M1.B_{short}) bound to hairpin M1.C (Figure 3.1(a), highlighted parts). If transfection of M1.B_{short} together with hairpin M1.C will lead to down-regulation of d2EGFP in cells, this is an indication that branch migration may be occurring within cells. Figure 3.1(b) demonstrates that transfection of either component on its own results in minimal down-regulation of d2EGFP; this is despite the fact that M1.B_{short} is the antisense sequence to d2EGFP. When transfected together (in the same vesicle), M1.B_{short} and M1.C can interact during the transfection process but may not necessarily interact inside the cells. To verify, the transfection was split into two: one complex contained only M1.B_{short} while the second complex contained only M1.C, both complexes were added to the cells simultaneously. Split transfection of M1.B_{short} and M1.C leads to approximately 50% knockdown of d2EGFP while transfection of the annealed duplex or a control DsiRNA results in approximately 90% knockdown. No increase in down-regulation was observed when the transfection concentration is increased ~4-fold. It is possible that, despite forming a separate transfection complex for each strand, the M1.B_{short} vesicles and the M1.C vesicles fuse during the transfection processes resulting in

binding of M1.B_{short} and M1.C outside of the cells and not in the cells. To rule this out, hairpin M1.C was transfected on its own and three hours later an additional transfection of M1.B_{short} was done (Figure 3.1(b), 83nM C, then B_{short}). Similar levels of d2EGFP knockdown were observed compared to a split transfection. This data is suggestive of the opening of hairpin M1.C by M1.B_{short} inside cells. The cells were not washed prior to the second transfection and so it is still possible that some vesicles containing M1.C remained outside the cells and that those vesicles fused with the B_{short} vesicles prior to entering the cells. We therefore sought to express both strands in cells; expression of the components rules out their interaction during transfection.

Strands were cloned into pSilencer plasmid (Ambion) under the control of an H1 promoter between the BamHI and HindIII sites. Co-delivery of pSilencer-M1.B_{short} with hairpin M1.C or co-delivery of pSilencer-M1.B_{short} and pSilencer-M1.C did not result in d2EGFP knockdown (data not shown). To overcome the need to deliver two plasmids into the same cell we next sought to express both strands off the same plasmid. M1.B_{short} was cloned into pSilencer under the control of a U6 promoter between the BamHI and HindIII sites. A shorter version of M1.B_{short} without domain 'b' (see Figure 3.1(a)) was used. Domain 'b' is the two nucleotide overhang of the Dicer substrate, this overhang will be introduced from the termination sequence of Pol III polymerase and is therefore unnecessary. The U6 promoter: M1.B_{short} cassette was amplified off the plasmid and cloned between the NarI sites in pSilencer H1 promoter: M1.C plasmid. Expression of the simplified system did not lead to d2EGFP knockdown. The observed reduction in d2EGFP levels appears to be due to expression of hairpin M1.C, this is contradictory to data obtained from transfecting hairpin M1.C (Figure 3.1(b) and (c)). Northern blot analysis was done to confirm the expression of M1.C and M1.B_{short}, however no signal was obtained (data not shown). Future studies are needed to determine whether this is due to a low expression level or an unoptimized blotting protocol. Expression of a GFP shRNA in combination with a DsRed shRNA into d2EGFP DsRed expressing cells resulted in knockdown of both proteins as was

determined by flow cytometry, suggesting that a different plasmid expressing two hairpins is functional (data not shown).

An additional attempt to combat insufficient delivery of hairpins and/or degradation was made by generating an expression plasmid that can transcribe three different inserts under separate promoters. Such a plasmid will enable a single cell to transcribe the three hairpins of the mechanism while a continuous production of hairpins would compensate for hairpin degradation. Each insert was cloned using a different set of restriction sites allowing for convenient exchange of inserts (Figure 3.2(a), for a list of constructed plasmids, refer to the Materials and methods section). Transfection of a plasmid expressing a full conditional Dicer substrate formation mechanism (sample A B C in Figure 3.2(b)) did not lead to down-regulation of d2EGFP. The observed reduced expression of d2EGFP appears to be due to expression of hairpin M1.C as can be seen by sample hairpin C and two negative control shRNAs in Figure 3.2(b). The negative control shRNA does not lead to d2EGFP down-regulation as can be seen by transfection of a plasmid expressing three negative control shRNAs (Figure 3.2(b)). To examine whether the plasmid expresses all components, a northern blot was performed. The data suggests that only hairpin M1.C is expressed in significant levels in cells. An extremely faint band was detected for hairpin M1.B and no signal was detected for hairpin M1.A.

Both U6 and H1 promoters are transcribed by RNA polymerase III and are therefore competing for resources. Work comparing the efficiency of U6 versus the H1 promoters for lentiviral delivery of shRNAs shows that the U6 promoter is superior to H1 in both tissue culture and *in vivo* [10]. In the expression plasmid created, hairpin M1.C is under the control of a U6 promoter whereas hairpins M1.A and M1.B are under the control of an H1 promoter. This could result in the observed variation of hairpin expression. To further pursue the expression of a full mechanism, the plasmid should be changed so that all hairpins are under the control of the same promoter with the goal of achieving a similar expression level for all hairpins.

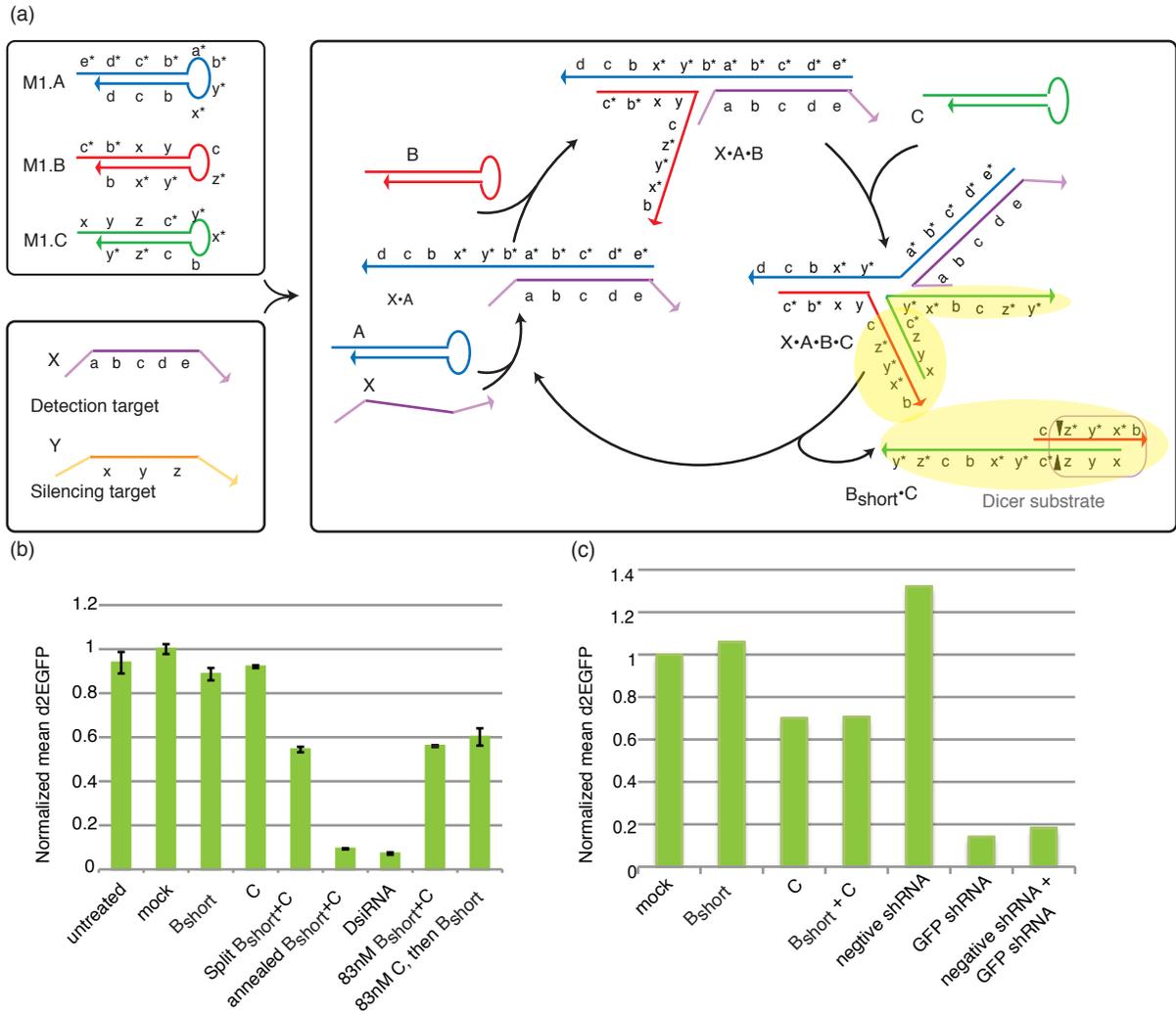


Figure 3.1: **d2EGFP knockdown via a simplified triggered Dicer substrate formation mechanism (M1).** (a) Schematic representation of the simplified mechanism. Highlighted regions represent the transfected strands. (b) Relative d2EGFP fluorescence 26 hours post transfection. Transfections were done in triplicate using 20nM final concentration of each strand unless otherwise specified. (c) Plasmid expression of a simplified mechanism. Relative d2EGFP fluorescence 48 hours post transfection. M1.B_{short}+M1.C and negative shRNA+GFP shRNA are plasmids expressing two different strands from the same plasmid. One sample per transfection was analyzed.

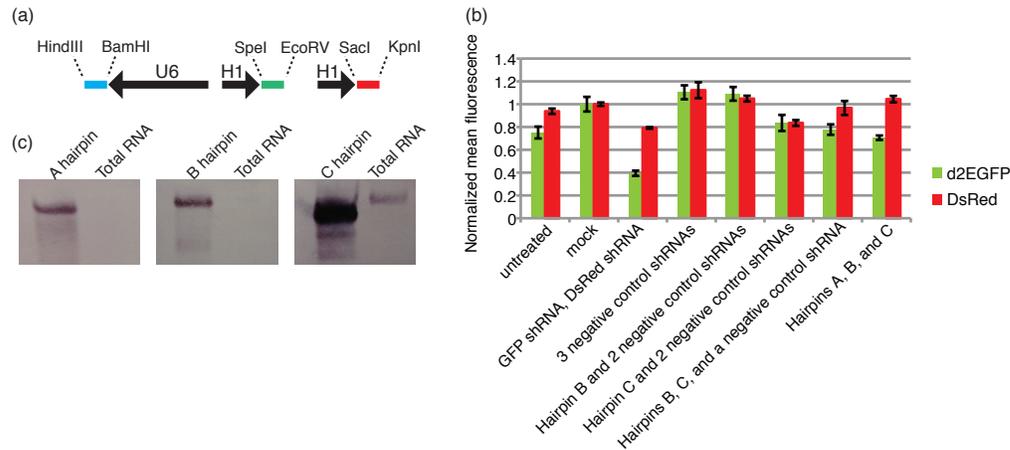


Figure 3.2: **Expression of three hairpins from one plasmid, M1 mechanism.** (a) Schematic representation of the expression construct. Black arrows represent promoters and their direction. Colored lines represent three different inserts. Restriction enzymes used for the cloning of each hairpin are listed. Drawing is not to scale. (b) Relative d2EGFP and DsRed fluorescence 48 hours post triple-expression plasmid transfection. Transfections were done in triplicate, error bars represent the standard deviation of the mean. Negative shRNA sequence was adapted from the pSilencer negative control (Ambion). (c) Northern blot analysis for the expression of hairpins from a plasmid expressing M1.A, M1.B and M1.C hairpins. A triple expression plasmid was transfected into HEK293 d2EGFP cells. Total RNA was extracted using Trizol 24 hours post transfection, $30\mu\text{g}$ total RNA were run per blot. Two pmol of synthetic M1.A and M1.B hairpins and five pmol of M1.C hairpin were blotted as controls. Blotting conditions were according to the *mirVana* miRNA blotting procedure (see Materials and methods in Chapter 5) using 40nM of biotin labeled DNA probes. Probes are the reverse complement of each hairpin. Detection was carried out using a biotin chromogenic detection kit (Thermo Scientific) according to the manufacturer.

3.2.2 Study of hairpin degradation in tissue culture

Expression of hairpin M1.C in cells leads to d2EGFP knockdown which indicates that the hairpin may get cleaved inside the cells. Hairpin M1.C has an 18bp stem, a 5' eight-base toehold and a 16-nucleotide loop with some secondary structure (Figure 3.3(a)). While it has a shorter stem than a canonical Dicer substrate (as well as a non-canonical 5' toehold) the secondary structure of the loop may be considered as part of the stem. Indeed, this hairpin can be processed by Dicer *in vitro* (see Appendix B Figure B.1(b)). We used 2'-OMe blot analysis to determine whether hairpin M1.C is being cut inside the cells. Northern blot confirms that a fraction of the hairpin is being cut 10 hours post transfection (Figure 3.3, C hairpin). Chemical modification of hairpin M1.C with 2'-OMe reduced cleavage as expected (Figure 3.3, C2 hairpin and Appendix B Figure B.1(c)). It is yet to be determined whether the observed lower band in Figure 3.3(c) (C2 hairpin) is due to some cleavage or impurity in the IDT synthesis. Cleavage products were observed both via SYBRGold staining pre-transfer and using a 2'-O-Me blot (Figure 3.3(b) and (c), respectively). Not all of the cleavage product bands are detected by blotting, this is probably due to the choice of a full complement probe which has a hairpin secondary structure. To see if loop size and secondary structure matter for cleavage we changed the loop of hairpin C to four nucleotides (Figure 3.3, C3); the shorter loop did not affect cleavage.

3.2.3 Silencing by mechanism hairpin C is mediated by RNAi

In Chapter 2 we presented a system based on hairpins with 3' toeholds. This section focuses on studies done with hairpin M2.C from a 3' toehold system. This hairpin is shorter than the hairpin in Chapter 2. Hairpin M2.C is predicated to have a 15bp stem and a six-nucleotide 3' toehold (Figure 3.4(a), M2.C) and is therefore not a conventional Dicer substrate. However, *in vitro* Dicer assays show that this hairpin can be somewhat cleaved. To reduce Dicer processing hairpin M2.C was modified with 2'-O-Me across the 5' end of the stem and halfway into the loop (Figure 3.4(a), M2.C2). This modification pattern

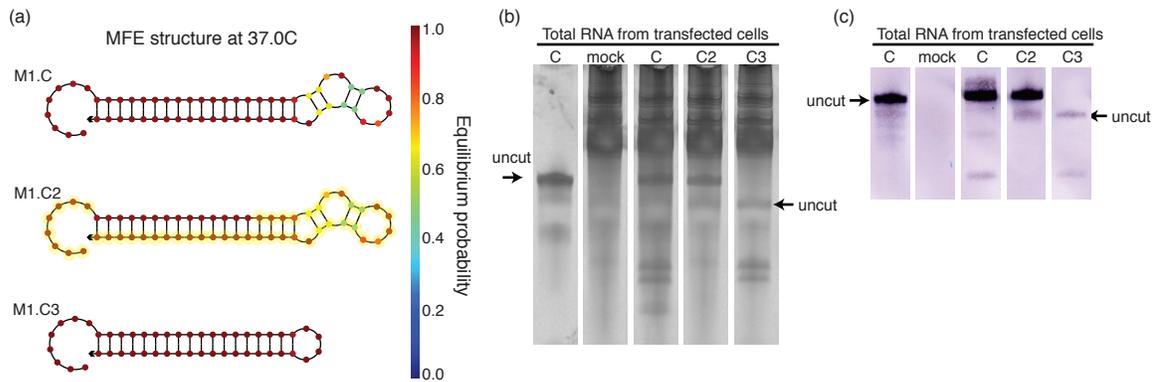


Figure 3.3: Hairpin C of mechanism M1 is partially degraded in tissue culture. (a) Predicted minimum free energy secondary structure of hairpins analyzed by NUPACK. Highlighted bases represent 2'-O-Me modifications. (b) Pre-transfer SYBRGold staining of 2'-O-Me blot gels. 50nM of each hairpin was transfected into HEK293A cells, transfections were done in duplicate. Total RNA was extracted using Trizol 10 hours post transfection, all of the total RNA collected from two transfections was run per blot. One pmol of synthetic M1.C hairpin was blotted as a control. (c) Northern blot analysis for the degradation of hairpin C in cells. Blotting conditions were according to the *mirVana* miRNA blotting procedure (see Materials and methods in chapter 5) using 20nM of biotin labeled DNA probe (hairpin C reverse complement). Detection was carried out using a biotin chromogenic detection kit (Thermo Scientific) according to the manufacturer.

abrogated Dicer cleavage *in vitro* (Figure 3.4(b)). Mechanism M2 targets d2EGFP region 597–615. This region is not a good silencing target, as can be seen by the siRNA transfection in Figure 3.4(c). Nevertheless, hairpin M2.C can lead to d2EGFP knockdown in similar levels to an siRNA. As expected, M2.C2, which is chemically modified, exhibits reduced d2EGFP knockdown. Supporting evidence comes from 2'O-Me blot analysis demonstrating that hairpin C was digested to an siRNA-like size in tissue culture. While the chosen probe and/or blotting conditions are not sufficient for detection of full-length M2.C and M2.C2, a band corresponding an siRNA in size is observed in cells transfected with M2.C. Again, hairpin M2.C2 does not appear to be cleaved (Figure 3.4(d)).

Finally, we use 5' rapid amplification of cDNA ends (5' RACE) to examine whether knockdown is mediated by the RNAi pathway. Ago2 has a defined cleavage site and is expected cut the target mRNA between bases nine and ten relative to the 5' end of the mRNA (10–11 nt downstream from the 5' end of the guide strand) [11, 12]. If M2.C silence via Ago2, the d2EGFP mRNA is expected to be cleaved between nucleotides 605 and 606. Figure 3.4(e) demonstrates the mRNA cut sites obtained by 5' RACE from cells transfected with M2.C or M2.C2. Indeed, cells transfected with M2.C mostly display a cleavage pattern around the predicted site (Figure 3.4(e)). It is possible that the reason why a uniform cleavage point is not observed is due to the non-canonical structure of M2.C. It was difficult to obtain 5' RACE data for M2.C2 cleavage. This is most likely due to the fact that it results in low knockdown levels. Still, some data mapped to the siRNA region, however not to the expected site (Figure 3.4(e)).

3.3 Discussion

We have discussed in this chapter several reasons why conditional Dicer substrate formation is not functional in a cell. Issues such as delivery, the ability to carry out toehold-mediated branch migration, hairpin processing in cells and mode of d2EGFP down-regulation were explored. Initial studies of toehold-mediated branch migration using a minimal system

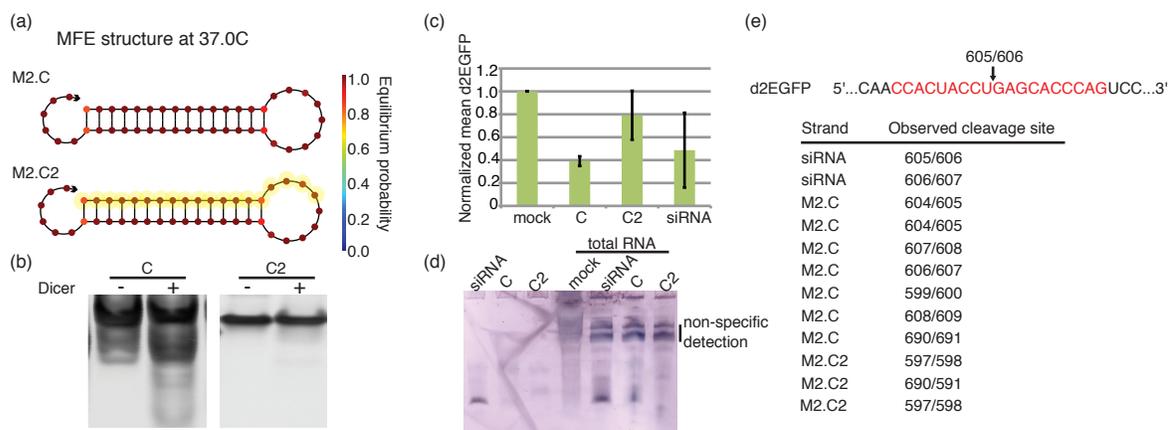


Figure 3.4: **Hairpin C of mechanism M2 leads to RNAi mediated knockdown.**

(a) Predicted minimum free energy secondary structure of hairpins analyzed by NUPACK. Highlighted bases represent 2'-O-Me modifications. (b) Hairpin M2.C but not 2'-O-Me chemically modified M2.C2 is cut by Dicer *in vitro* (for methods see Materials and methods section in Chapter 2). (c) Relative d2EGFP knockdown following hairpin or siRNA transfection. Mean fluorescence represents data from 3–4 independent experiments, error bars represent the standard deviation of the mean. (d) Northern blot analysis for the degradation of hairpin C in cells. Blotting conditions were according to the *mirVana* miRNA blotting procedure (see Materials and methods in Chapter 5) using 20nM of biotin labeled 2'-O-Me probe (5'-AmGmCmCmAmCmUmAmCmCmUmGmAmGmCmAmCmCmAmG-3'). Detection was carried out using a biotin chromogenic detection kit (Thermo Scientific) according to the manufacturer. (e) Observed d2EGFP knockdown is mediated by the RNAi pathway. 5' RACE data obtained from transfection of a M2.C, M2.C2 and siRNA demonstrating the cleavage point on d2EGFP.

were promising, suggesting that toehold-mediated branch migration is likely occurring. Attempts to express such a system resulted in silencing by hairpin M1.C, contradictory to transfection of a synthetic hairpin (Figure 3.1(b) and (c)). Similar results were obtained when trying to express a full mechanism (Figure 3.2(b)). In addition, expression of a full system resulted in the sole expression of hairpin C. This is most likely due to choice of promoters. U6 promoter, which controls the expression of hairpin C, has a stronger expression than the H1 promoter [10], which controls the expression of hairpins A and B. For future expression of a full mechanism, the same promoter must be used to express all three hairpins.

Due to the observed silencing by hairpin C, we next focused on examining the fate of hairpin C once transfected into cells. Hairpin C was studied for both a 5'- and a 3'-toehold mechanism (Figures 3.3 and 3.4, respectively). Despite being an sshRNA, hairpin C can be processed by Dicer *in vitro*, and is cleaved *in vivo*; when chemically modified with 2'O-Me, processing is negated. Hairpin C appears to down-regulate d2EGFP by an RNAi-dependent mechanism of action as based on 5' RACE data.

The short stem of hairpin C is shorter than that of a conventional shRNA, as for an sshRNA. These hairpins are thought to be processed by an enzyme other than Dicer, possibly Ago2, in cells [6–9]. The activity of sshRNAs appears dependent on their designation. Hairpins with the guide strand upstream (5') to the loop region are designated as Left-hand shRNAs (L shRNA) and those with the guide strand downstream to the loop (3') are designated as Right-hand (R shRNA) [6]. For R shRNAs it appears that endonucleolytic cleavage of the loop region is required for functional RNAi [13]. When the loop region was changed to 2'O-Me, the potency of the same sshRNA has decreased. However, when 2'O-Me and phosphorothioate (PS) modifications were placed on the passenger arm across from Ago2's slicer activity, the silencing ability of R sshRNAs was not significantly affected. Conversely, L shRNAs do not depend on loop cleavage, but rather depend on slicing activity (stem cleavage) [13]. Both types of hairpin C (5'- and 3'-toehold) are con-

sidered to be R type hairpins. Processing of the loop may explain the observed silencing of M2.C2, whereas M1.C2, whose entire loop is 2'O-Me, does not down-regulate d2EGFP (Figure 3.4, data not shown for M1.C2). It would be interesting to see whether changing hairpin M2.C2 to an RNA loop would improve its silencing capabilities.

To conclude, the work presented in this chapter suggests that at least part of the transfected hairpins undergo unwanted processing. Similar analysis must be done for hairpins A and B as well. It may be beneficial to alter their sequence in order to observe d2EGFP knockdown and mRNA cleavage. Further studies are needed to determine whether hairpin processing in cells is mediated by Dicer, Ago2 or another protein.

3.4 Materials and methods

Cell lines. HEK293 d2EGFP cells were a generous gift from Dr. Chase Beisel. The destabilized EGFP sequence comes from pd2EGFP-1 plasmid (Clontech, PT3205-5 catalog #6008-1). HEK293 d2EGFP DsRed cells were generated by a stable transfection of pDsRed2-1-C1 (Clontech, PT3603-5 catalog #632407) into HEK293 d2EGFP cells. Cells were maintained at 37°C 5% CO₂ in DMEM (Invitrogen) supplemented with 10% fetal bovine serum (Invitrogen).

Oligo transfections. Transfections were carried out at the specified oligonucleotide concentrations using HiPerFect transfection reagent (Qiagen) according to the Fast-Forward transfection protocol. $1-2 \times 10^5$ cells were plated per well of a 24-well plate. Cell counts and viability were determined using the Countess automated cell counter according to the manufacturer (Invitrogen).

Plasmid transfections. Transfections were done using Lipofectamine2000 (Invitrogen) according to the manufacturer. Cell counts were determined as described above.

Flow cytometry. 24–48 hours post transfection samples were trypsinized for 5 minutes at 37°C, the trypsin was then quenched with DMEM growth media supplemented with 10% FBS. Samples were run on a BD Accuri C6 flow cytometer (BD Biosciences). Healthy

cells were gated according to their scatter using untreated cells. Results represent the normalized mean and error bars represent the standard deviation of the mean of three transfected samples.

Construction of a three hairpin expression plasmid. A custom plasmid (pIDTSMART backbone) was synthesized by IDT containing the following insert: NotI restriction site - H1 promoter - SacI restriction site - GFP shRNA - polIII termination sequence - KpnI restriction site. The expression cassette was amplified using the following primers and cloned into a pSilencer 2.1 U6 backbone between the EcoRI site:

pIDTSMART GFP shRNA for: 5'-ACGTAGGAATTCAGATCTGCGGCCGCAATTCATATTTGC-3' this primer adds the EcoRI and BglII restriction sites to the amplicon

pIDTSMART GFP shRNA rev: 5'-AGCTAGGAATTCATCGATGGTACCTTCCAAAAAAGACCCTG-3' this primer adds the EcoRI and ClaI restriction sites to the amplicon

A second custom plasmid (pIDTSMART backbone) was synthesized by IDT containing the following insert: BglII restriction site - H1 promoter - SpeI restriction site - GFP shRNA - polIII termination sequence - EcoRV restriction site - NotI restriction site. The expression cassette was amplified using the following primers and cloned into a pSilencer 2.1 U6 H1 promoter backbone (see above) between the BglII and NotI sites:

pIDTSMARTv2 for for: 5'-GATTCTGAATTCAGATCTGTCAGGCTATGGCGCG-3' this primer adds the EcoRI restriction site to the amplicon

pIDTSMARTv2 rev: 5'-ATGACAGAATTCGATATCTTCCAAAAAAGACCC-3' this primer adds the EcoRI restriction site to the amplicon

All plasmids were verified to contain the correct sequence via sequencing (Laragen).

Plasmid	U6: BamHI HindIII	H1: SacI KpnI	H1: SpeI EcoRV
3×negative shRNA	negative shRNA	negative shRNA	negative shRNA
M1.A 2×negative shRNA	negative shRNA	negative shRNA	M1.A
M1.B 2×negative shRNA	negative shRNA	M1.B	negative shRNA
M1.C 2×negative shRNA	M1.C	negative shRNA	negative shRNA
M1.A M1.B negative shRNA	negative shRNA	M1.B	M1.A
M1.B M1.C negative shRNA	M1.C	M1.B	negative shRNA
M1.A M1.C negative shRNA	M1.C	negative shRNA	M1.A
M1.A M1.B M1.C	M1.C	M1.B	M1.A

Expression plasmid hairpin sequences

M1.A: 5'-CUCGAUCUCGAACUCGUGGCUGGUCAGCUUGCCGUACACGAGUUCG-3'

M1.B: 5'-CGAACUCGUGUACGGCAAGCUGACCGAGACUUCAGGGUCAGCUUGC
CGUACA-3'

M1.C: 5'-UACGGCAAGCUGACCCUGAAGUCUCGGUCAGCUUGCCGUACACGAG
ACUUCAGGGUCAGC-3'

Negative shRNA: 5'-GUCAGGCUAUCGCGUAUCGUUCAAGAGACGAUACGCGAUA
GCCUGAC-3'

5' RACE. 30 pmol C hairpin were snap cooled and transfected into HEK293 d2EGFP cells using RNAiMAX reverse transfection protocol (Invitrogen). 48 hours post transfection the were samples were trypsinized for 5 minutes at 37°C, the trypsin was then quenched with DMEM growth media supplemented with 10% FBS. Triplicates were combined and total RNA was extracted using ZR RNA MiniPrep kit (Zymo research) with an In-column DNaseI digestion according to the manufacturer or RNA spin mini (GE healthcare) according to the manufacturer. 1.5µg total RNA were ligated to 20pmol of GeneRacer RNA oligo (5'-CGACUGGAGCACGAGGACACUGACAUGGACUGAAGGAGUAGAAA-3'): total RNA and oligo were combined in a total volume of 10µl, heated to 65°C for 5 minutes followed by a 2 minute incubation on ice. Ligation was carried out for 1 hour at 37°C

using 20 units of RNA ligase 1 (NEB), 40 units RiboGuard (Epicenter), 1mM ATP 10% PEG-8000 in a total volume of 20 μ l. Ligation products were then purified using RNeasy mini kit (Qiagen) according to the manufacturer. cDNA was transcribed using SuperScriptIII (Invitrogen) according to the manufacturer using d2EGFP 839 rev primer (5'-TTGATCCTAGCAGAAGCACAGGCT-3'). The cDNA was amplified using OneTaq hot start 2 \times master mix with standard buffer (NEB) using touch-down PCR and the following primers: GeneRacer 5' primer (5'-CGACTGGAGCACGAGGACACTGA-3') and d2EGFP 822 reverse primer (5'-ACAGGCTGCAGGGTGACGGTCCAT-3'). PCR program: 94°C for 2 minutes, 5 cycles of 94°C for 30 seconds and 1 minute at 72°C, 5 cycles of 94°C for 30 seconds and 1 minute at 70°C, 20 cycles of 94°C for 30 seconds, 66°C for 30 seconds, 68°C for 1 minute, followed by a 5 minute incubation at 68°C. A second nested PCR reaction was carried out using the following primers: GeneRacer 5' nested primer (5'-GGACACTGACATGGACTGAAGGAGTA-3') and d2EGFP 728 reverse primer (5'-TGGCTAAGCTTCTTGACAGCTCG-3'). PCR conditions are as listed above with the following modification: for the 20 cycles annealing was done at 68°C. PCR products were sequenced by Laragen.

Sequences used:

M1.B_{short}: 5'-ACCGAGACUUCAGGGUCAGCUUGCCGUACA-3'

M1.C: 5'-UACGGCAAGCUGACCCUGAAGUCUCGGUCAGCUUGCCGUACACGAGACUUCAGGGUCAGC-3'

M1.C2: 5'-mUmAmCmGmGmCmAmAGCUGACCCUGAAGUmCmUmCmGmGmUmCmAmGmCmUmUmGmCmCmGmUmAmCmAmCmGmAmGmAmCmUmUmCmAmGmGmGmUmCmAmGmC-3'

M1.C3: 5'-UACGGCAAGCUGACCCUGAAGUCUCGGUCACGAGACUUCAGGGUCAGC-3'

M2.C: 5'-UACCUGAGCACCCAGCCACUACCUCUGGGUGCUCAGGUAGUGGCU-3'

M2.c2: 5'-mUmAmCmCmUmGmAmGmCmAmCmCmAmGmCmCmAmCUACCUC

UGGGUGCUCAGGUAGUGGCU-3'

Bibliography

- [1] McManus, M. T., Petersen, C. P., Haines, B. B., Chen, J., and Sharp, P. A. Gene silencing using micro-RNA designed hairpins. *RNA* **8**(6), 842–50 (2002).
- [2] Sun, X., Rogoff, H. A., and Li, C. J. Asymmetric RNA duplexes mediate RNA interference in mammalian cells. *Nature biotechnology* **26**(12), 1379–82 (2008).
- [3] Chu, C. Y. and Rana, T. M. Potent RNAi by short RNA triggers. *RNA* **14**(9), 1714–9 (2008).
- [4] 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. K. Asymmetric shorter-duplex siRNA structures trigger efficient gene silencing with reduced nonspecific effects. *Molecular therapy : the journal of the American Society of Gene Therapy* **17**(4), 725–32 (2009).
- [5] Bramsen, J. B., Laursen, M. B., Damgaard, C. K., Lena, S. W., Babu, B. R., Wengel, J., and Kjems, J. Improved silencing properties using small internally segmented interfering RNAs. *Nucleic acids research* **35**(17), 5886–97 (2007).
- [6] Ge, Q., Ilves, H., Dallas, A., Kumar, P., Shorestein, J., Kazakov, S. A., and Johnston, B. H. Minimal-length short hairpin RNAs: the relationship of structure and RNAi activity. *RNA* **16**(1), 106–17 (2010).
- [7] Cheloufi, S., Dos Santos, C. O., Chong, M. M., and Hannon, G. J. A Dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* **465**(7298), 584–9 (2010).
- [8] Cifuentes, D., Xue, H., Taylor, D. W., Patnode, H., Mishima, Y., Cheloufi, S., Ma, E., Mane, S., Hannon, G. J., Lawson, N. D., Wolfe, S. A., and Giraldez, A. J. A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* **328**(5986), 1694–8 (2010).

- [9] Yang, J. S., Maurin, T., Robine, N., Rasmussen, K. D., Jeffrey, K. L., Chandwani, R., Papapetrou, E. P., Sadelain, M., O'Carroll, D., and Lai, E. C. Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **107**(34), 15163–8 (2010).
- [10] Makinen, P. I., Koponen, J. K., Karkkainen, A. M., Malm, T. M., Pulkkinen, K. H., Koistinaho, J., Turunen, M. P., and Yla-Herttuala, S. Stable RNA interference: comparison of U6 and H1 promoters in endothelial cells and in mouse brain. *The journal of gene medicine* **8**(4), 433–41 (2006).
- [11] Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Molecular cell* **15**(2), 185–97 (2004).
- [12] Rana, T. M. Illuminating the silence: understanding the structure and function of small RNAs. *Nature reviews. Molecular cell biology* **8**(1), 23–36 (2007).
- [13] Dallas, A., Ilves, H., Ge, Q., Kumar, P., Shorestein, J., Kazakov, S. A., Cuellar, T. L., McManus, M. T., Behlke, M. A., and Johnston, B. H. Right- and left-loop short shRNAs have distinct and unusual mechanisms of gene silencing. *Nucleic acids research* **40**(18), 9255–71 (2012).

Chapter 4

Engineering a conditional shRNA transcription mechanism

4.1 Introduction

As introduced in Chapter 2, RNAi can be induced by siRNAs, dsRNAs, and shRNAs. Once in cells, these molecules turn RNAi on. The need for spatio-temporal induction of RNAi has led to the development of methods to conditionally activate RNAi. In this chapter, we present a conditional shRNA transcription mechanism. We engineered an inactive promoter upstream to an shRNA coding sequence. Upon the presence of an input nucleic acid detection molecule, the promoter changes conformation into an active state, thus driving the expression of an shRNA.

shRNAs can be transfected directly into cells, however, it is common practice to incorporate them into vectors and have them transcribed in cells under the control of polymerase II (pol II) or polymerase III (pol III) promoters [1–4]. Once transcribed, the shRNA is exported from the nucleus to the cytoplasm by Exportin-5 and processed by Dicer into an siRNA [2, 5, 6]. Pol III promoters such as the U6 and H1 promoters are ideal for the production of shRNAs as they naturally transcribe small RNA transcripts lacking a 5' cap and a polyadenylation tail [1]. In addition, their transcription initiation point and termination signal (4–6 thymidines) are well defined, ensuring expression of an shRNA with an exact

sequence and without added bases, a property critical for siRNA generation [7, 8]. Pol II promoter-driven shRNAs can be expressed in a tissue specific matter and are transcribed as micro-RNA precursors that have a 5' cap and polyadenylation tail that must be processed [9, 10].

Constitutively expressed shRNAs have the advantage of mediating prolonged silencing phenotypes compared to chemically synthesized siRNAs, yet they limit the analysis of genes essential for cell survival, cell cycle regulation and cell development. Another disadvantage of constitutive shRNA expression is the risk of off-target effects due to saturation of endogenous silencing pathways [11, 12]. Therefore, a conditional shRNA expression system may be favorable. To this end, tissue specific promoters [10, 13, 14] or inducible promoters [8, 10] can be used. Inducible promoters can be either reversible or irreversible. Reversible induction of shRNA expression systems are most commonly based on tetracycline or ecdysone [10, 15–19]. Irreversible systems are based on the removal of a genetic element that inhibits or promotes transcription and are mostly based on Cre-loxP or Flp recombination [10, 20–22].

Recently, the use of non-coding RNAs to control shRNAs has been developed [23–27]. In this approach, non-coding RNAs containing an shRNA sequence and an aptamer are expressed either in an on (shRNA accessible to RNAi machinery) or in an off (shRNA inaccessible to RNAi machinery) conformation. Addition of a small molecule induces a conformational change in the non-coding RNA which results in turning RNAi from an on state to an off state (or vice versa).

With the exception of tissue specific promoters, a limitation of the approaches described so far is that they rely on the addition of an external molecule (small ligand or an enzyme). An advantage for enzymes is that they can be co-expressed with the shRNA or genetically encoded into the cells. An advantage for the use of a ligand is that ligand concentration can tune the level of shRNA expression. A disadvantage is that the need to add an inducer in addition to an expression construct complicates the system.

An alternative approach to achieve conditional shRNA transcription is to rely on nucleic acids as the inducer molecule. Similar to aptamer-based approaches, nucleic acids can lead to structural changes in conformation of a nucleic acid molecule thus making it active or inactive. Nucleic acid triggers offer a further advantage over aptamers in that they can be programmable, whereas aptamers cannot. An important aspect of a nucleic acid-based approach is that the inducer can be an endogenous nucleic acid target. This overcomes the need to deliver or express an inducer. By choosing the inducer appropriately (e.g., mRNA target expressed only in a subset of cells), spatio-temporal control of expression can be achieved.

We present a method to control shRNA transcription based on nucleic acids. Similar to the mechanism in Chapter 2, this mechanism follows the logic operation: *If gene X is detected, silence independent gene Y*. This mechanism is fundamentally different from the mechanism described in Chapter 2 because in this design, the Dicer substrate (shRNA) is made by enzymatic means and not solely based on strand displacement cascades. To achieve such a design, the transduction molecules must be made of DNA, whereas the design described in Chapter 2 is made of RNA and 2'-OMe. The concept behind this mechanism is that a promoter is active only when a sufficient number of base pairs are formed. Kim et al. [28] have previously demonstrated the use of conditional hybridization of a double-stranded DNA promoter in order to obtain conditional *in vitro* transcription. Here, conditional shRNA transcription is mediated by small conditional DNAs (scDNAs) through toehold-mediated strand displacement. Activation of the mechanism is demonstrated in the presence of a short synthetic nucleic acid target in a test tube. The mechanism exhibits good ON-to-OFF ratio; in the absence of a detection target, minimal shRNA is transcribed. The transcribed shRNA can be processed by Dicer *in vitro* as well as lead to d2EGFP knockdown in tissue culture.

4.2 Mechanism and Design

4.2.1 Mechanism

We have designed a conditional RNAi mechanism based on conditional shRNA transcription. The concept behind this mechanism is that a promoter is active only when a sufficient number of base pairs are present. In the absence of a detection target X (OFF state), the hairpins maintain their secondary structure and should not interact with one another. In the presence of a detection target X (ON state), interaction between the hairpins occurs, a functional promoter is formed and an shRNA targeting gene Y is transcribed.

Our mechanism is comprised of two metastable DNA hairpins kinetically trapped in the hairpin state. The presence of a detection target X starts a cascade of events that allows the hairpins to reach a thermodynamic equilibrium that otherwise cannot be reached in short time scales. The system (Figure 4.1) has the following domains: a target recognition sequence ('b*-a*'), a disrupted promoter sequence ('p1-p2'), an shRNA coding sequence ('z-y*-z*') and a transcription termination sequence ('t'). In the presence of a detection target, toehold 'b*' of hairpin A can bind to the target leading to a branch migration in which complex X·A is formed (Step 1). Next, toehold 'p1*' of hairpin B nucleates with segment 'p1' of hairpin A resulting in a branch migration in which complex X·A·B is formed (Step 2). Complex X·A·B now has a full dsDNA promoter followed by an shRNA coding sequence which can be transcribed (Step 3). This mechanism is catalytic in the sense that one template can be transcribed into multiple shRNAs.

4.2.2 Design

The design follows the logic operation: *If gene X is detected, silence independent gene Y.* As can be seen in Figure 4.1, complete sequence independence is observed between the detection target X (sequence 'a-b') and the silencing target Y (sequence 'z'). This allows one to re-program the mechanism to detect and silence different genes. The design space

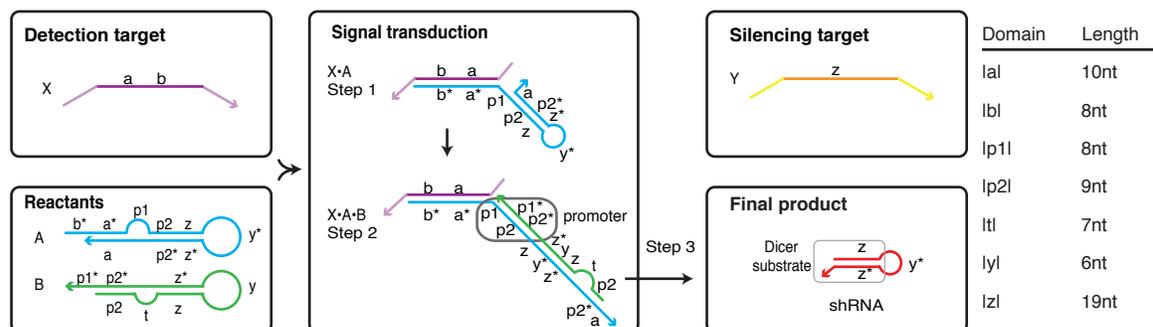


Figure 4.1: **Conditional shRNA transcription mechanism schematic.** The detection target X is recognized by A (Step 1), leading to the binding and opening of B (Step 2), which in turn forms an intact promoter and drives the transcription of a shRNA against the silencing target Y (Step 3).

is constrained for this mechanism. All sequences but the loop of the shRNA (domain ‘y’) are constrained by the detection target, silencing target, termination signal and promoter sequence. Some flexibility exists in the choice of dimensions for the shRNA (stem and loop size); typically, shRNA with 19–29bp stems separated by a 4–9 nucleotide loop are used [1, 29, 30]. We have chosen to use a 19bp stem with a six nucleotide loop. Careful consideration had to be put into the length specification of ‘p1’. It needs to be long enough to make ‘p2-p2*’ a non-functional promoter, yet short enough in order to minimize the interaction between hairpins A in B in the absence of a detection target (OFF state, also termed ‘leakage’). Leakage is also controlled by the length choice of domain ‘a’: the longer ‘a’, the less favorable it is for hairpin A to bind to hairpin B in the absence of a detection target. Sequences were assigned to our structures using the design feature on NUPACK [31, 32]. Sequences were specified for the promoter choice (discussed below) and for the shRNA. The shRNA sequence targets a destabilized version of enhanced green fluorescent protein (d2EGFP) and was adapted from an siRNA used by Ohert et al. [33]. To allow more flexibility in the design, for the work presented here, the target sequence was unspecified and the sequence was assigned by NUPACK to minimize the ensemble defect (a measurement of how far the system is from an ideal system)[32, 34]. We then

used the thermodynamic analysis tool on NUPACK [31, 35–37] to handpick the set of designed sequences that performed best based on further analysis (starting materials form the correct structure and that only desired complexes are formed).

4.2.3 Promoter choice

A suitable promoter for use in triggered shRNA transcription should have a defined transcription start site and preferably a transcription stop site. The promoter should be minimally active when not in full duplex form. Ideally, it should also be short for design and synthesis purposes. Natural candidates were the short and well-characterized bacteriophage T3 [38], SP6 [39] and T7 [40] promoters as well as the mammalian H1 promoter [41]. In this work we use the T7 and H1 promoters (for work pertaining to the H1 promoter see appendix C).

T7 RNA polymerase is not present in mammalian cells and therefore a T7-promoter based mechanism is less likely to interfere with endogenous pathways. However, this also means that T7 RNA polymerase needs to be introduced into mammalian cells, which is a disadvantage. The T7 promoter is extremely short, requiring only 17 nucleotides and has a well-defined transcription start site, both of which are a great advantage. Considering its short length, we postulated that by making half of the promoter single stranded, the promoters' function will be disrupted. Some termination of transcription is achieved by the synthesis of an RNA hairpin followed by a stretch of uridine (U) residues [42–44]. We have therefore added a poly-U termination sequence after the shRNA (domain 't', Figure 4.1). Another major advantage is the availability of enzyme for *in vitro* transcription. The use of T7 RNA polymerase *in vitro* to transcribe the strands of an siRNA as well as shRNAs to mediate RNAi has been well documented [2, 8, 45]. Conditional transcription has also been demonstrated *in vitro* by controlling the activity of the promoter using nucleic acid hybridization [28, 46]. Several groups have demonstrated that mammalian cells expressing the T7 RNA polymerase enzyme can trigger RNAi when presented with an

siRNA or shRNA under the control of a T7 promoter [47–49]. This strategy could be made conditional by expressing T7 RNA polymerase under the control of an inducible promoter.

4.3 Results

4.3.1 *In vitro* studies

The mechanism has two parts: conditional transcription template formation and shRNA transcription. For both parts, our aim was to verify that in the absence of a detection target, the OFF state is preserved, and that the presence of a detection target switches the mechanism to its ON state (complex $X \cdot A \cdot B$ is formed and an shRNA is transcribed). Native polyacrylamide gel electrophoresis demonstrates that minimal $A \cdot B$ duplex formation is observed in the OFF state, while a higher order complex $X_{short} \cdot A \cdot B$ is formed in the presence of a short synthetic DNA target X_{short} (lanes 3 and 5 (respectively) in Figure 4.2(a) and Figure C.1(a) in Appendix C). The system also converts with X_{short} that is made of RNA as is demonstrated in Figure C.1(b) in Appendix C. Radioactive *in vitro* transcription was used to verify that the T7 promoter is off when the target is absent and that it can transcribe an shRNA in the presence of a detection target. In the OFF state, minimal shRNA is produced (3.2%) while in the ON state an shRNA is formed (lanes 1 and 3 in Figure 4.2(b) and (c)); the shRNA transcribed can be cut into an siRNA by Dicer as demonstrated in Figure 4.2(b) lanes 2 and 4. Minimal transcription is observed from the reactants or intermediates in the absence of a target (Figure C.2 in Appendix C). The transcribed shRNA contains the expected shRNA sequence as can be seen by northern blot using a probe against the expected shRNA sequence (Figure 4.2(d)).

4.3.2 *In vivo* studies

So far we have demonstrated the conditional transcription of an shRNA that harbors a sequence targeting d2EGFP. The termination signal for T7 RNA polymerase is not very

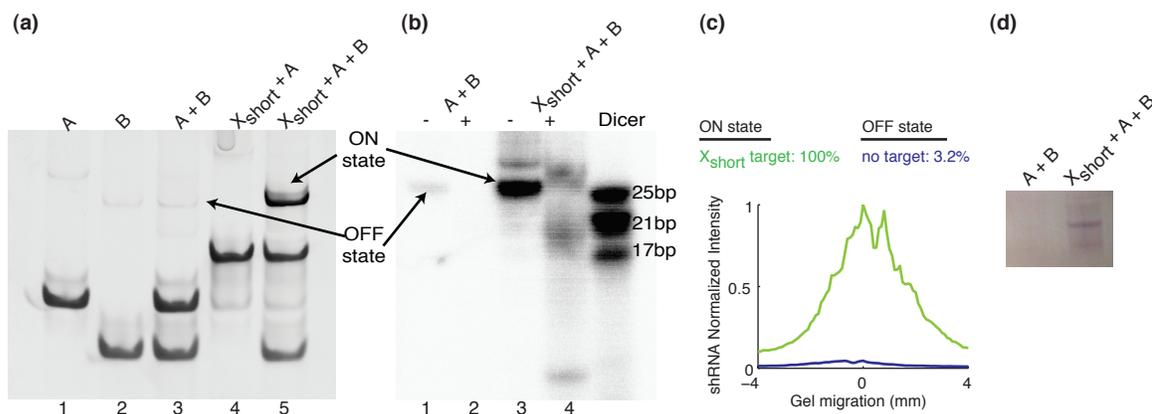


Figure 4.2: Conditional shRNA transcription. (a) Transcription template formation *in vitro*. In the OFF state, a minimal amount of A·B is produced (lane 3). In the ON state, the transcription template $X_{short} \cdot A \cdot B$ is produced (lane 5). (b) Conditional shRNA transcription and siRNA production *in vitro*. Radioactive shRNA transcription followed by Dicer processing. In the OFF state, a minimal amount of shRNA is transcribed (lane 1). In the ON state, a significant amount of shRNA is transcribed (lane 3). The shRNA is cleaved by Dicer to produce an siRNA (lanes 2,4). (-/+ indicates the absence/presence of Dicer). (c) ON:OFF shRNA transcription ratio. The shRNA band in panel (b) (- dicer lanes) was quantified and band intensity plotted. (d) Northern blot analysis demonstrating that the transcribed shRNA binds to the expected shRNA probe sequence.

strong and therefore the transcribed shRNA can have varying lengths of a 3' overhang (up to 16 nucleotides for a run-off transcript). Studies suggest that for shRNAs with 3' overhangs longer than three nucleotides, Dicer no longer uses the "3' end rule" to determine the cleavage site but rather cleaves the shRNA in a similar pattern as if it contained a three-nucleotide overhang [50]. Since the transcribed shRNA might not have a canonical shRNA overhang, we next tested whether it can down-regulate d2EGFP. Figure 4.3 demonstrates that a 20nM transfection of *in vitro* transcribed shRNA into HEK293 d2EGFP cells can down-regulate d2EGFP. This suggests that if the mechanism functions in cells, then a functional shRNA will be produced.

To test the functionality of our system in tissue culture, we needed cells that have the RNAi machinery and that express T7 RNA polymerase. We generated a stable HEK293 d2EGFP cell line expressing a cytoplasmic version of T7 RNA polymerase (see Materials and methods). The expression of T7 RNA polymerase was confirmed by western blot and compared to a T7 RNA polymerase enzyme solution (Figure 4.4(a)). The *in vivo* activity of the T7 RNA polymerase was demonstrated by reverse transcription PCR (RT-PCR) following the transfection of a plasmid (pT7CAT [51]) containing the chloramphenicol acetyltransferase (CAT) gene flanked by T7 promoter and termination signals. Figure 4.4(b) demonstrates that the CAT gene is being transcribed in cells transfected with the pT7CAT plasmid that express the T7 RNA polymerase mRNA (Figure 4.4(c)).

Next, we examined if, under the control of a T7 promoter, a linear DNA template containing an shRNA against d2EGFP can lead to d2EGFP down-regulation. Multiple attempts using various transfection reagents (Lipofectamine2000 (Invitrogen), Trans-IT-Oligo (Mirus)) as well as different amounts of DNA, the use of a carrier plasmid, and scrambled shRNA and promoter sequences were used. No significant down regulation of d2EGFP that can be attributed to RNAi was observed (data not shown). It appeared that when mild d2EGFP knockdown was observed it was sequence independent and/or due to toxicity during the transfection. Due to the lack of a proper positive control we did

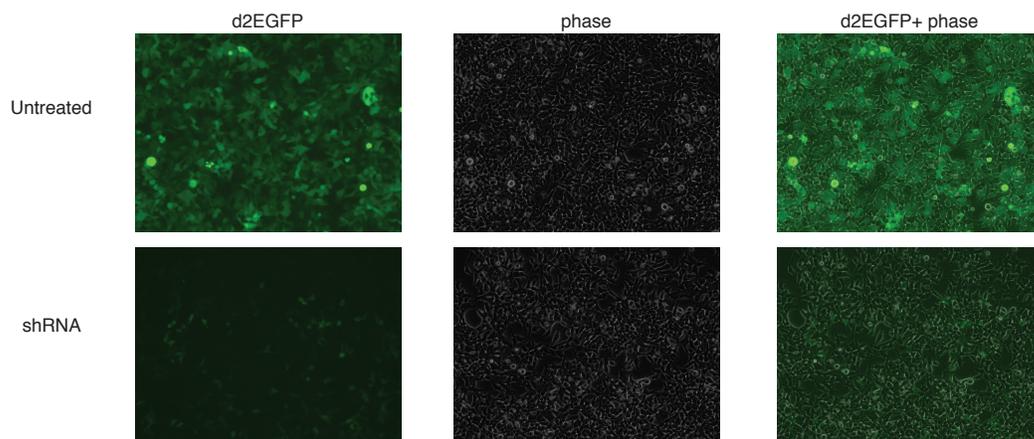


Figure 4.3: ***In vitro* transcribed shRNA down-regulates d2EGFP.** 20nM of *in vitro* transcribed shRNA were reverse-transfected (HiPerFect, Qiagen) into HEK293 d2EGFP cells. Transcribed shRNA was purified using a MasterPure RNA purification kit (Epicentre) according to the manufacturer. Images were taken with an inverted fluorescent microscope 48 hours post transfection.

not pursue the examination of the conditional shRNA transcription mechanism in tissue culture.

4.4 Discussion

In this chapter we introduced a conditional shRNA transcription mechanism based on T7 RNA polymerase. The mechanism implements the logic: *If gene X mRNA is detected, then produce an shRNA specific to independent gene Y.* While following the same logic operation as the mechanism presented in Chapter 2, conditional shRNA transcription is fundamentally different. This mechanism is based on DNA and relies on cellular machinery for shRNA transcription. This mechanism is potentially cheaper, more nuclease resistant, and is programmable using different promoters. As such, the mechanism was re-programmed to use an H1 pol III promoter as demonstrated in Appendix C.

For the T7-promoter-based system, we showed that in a test tube, the mechanism

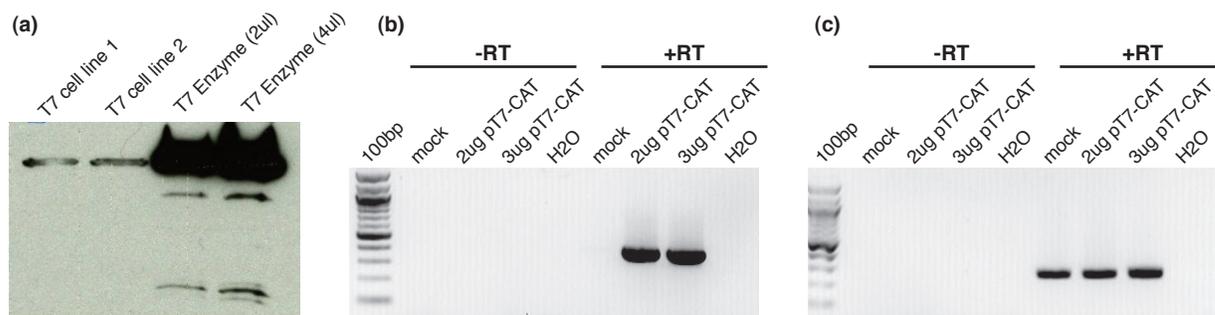


Figure 4.4: **T7 RNA expression and functionality in HEK293 d2EGFP cell line.** (a) Western blot of T7 RNA polymerase from protein extracted from HEK293 d2EGFP T7 cell line. Lanes 1 and 2: protein extracted from two cell line clones. Lanes 3 and 4: T7 RNA polymerase enzyme solution (Epicentere). (b) RT-PCR using CAT primers following 2µg or 3µg pT7CAT plasmid transfection into HEK293d2EGFP T7 cell line. (c) RT-PCR using T7 RNA polymerase primers following 2µg or 3µg pT7CAT plasmid transfection into HEK293d2EGFP T7 cell line.

conditionally transcribes an shRNA and that the shRNA is cleaved by Dicer to produce an siRNA. The mechanism exhibits a good ON-to-OFF ratio; minimal shRNA is transcribed in the absence of a detection target, while the presence of a detection target leads to significant shRNA transcription. We also demonstrated that the transcribed shRNA is functional, when introduced into cells expressing d2EGFP, the fluorescence intensity is knocked down.

T7 RNA polymerase is not endogenous to mammalian cells. We have therefore generated a stable cell line expressing d2EGFP and T7 RNA polymerase to test conditional shRNA transcription *in vivo*. However, we were not able to down-regulate d2EGFP via T7 transcription of shRNA from a linear DNA template and therefore conditional shRNA transcription has not been attempted. RNAi knockdown using linear DNA templates containing either a T7, H1 or U6 promoter has been previously demonstrated [49, 52, 53] suggesting this approach should work with further optimization. It is possible that the expression of T7 RNA polymerase is not high enough. T7 RNA polymerase is expressed bicistronically with an antibiotic selection marker, increasing the selection pressure will

select for cells expressing more enzyme. Further optimization of the transfection of short dsDNA oligonucleotides may also be needed.

Some modifications to the design may be needed to obtain efficient conditional shRNA transcription *in vivo*. In the current design, the transcription template is not released from the bound target. To achieve transcription while bound to a target mRNA, it may be necessary to add a spacer sequence between the promoter region and the mRNA binding site to accommodate the RNA polymerase; we have yet to explore this case. The immunostimulatory effects of a T7-RNA-polymerase-transcribed shRNA also need to be addressed. The transcription product of T7 RNA polymerase contains a 5' triphosphate. It was recently discovered that triphosphates induce type I interferon through activation of retinoic acid-inducible protein 1 (RIG-1) [54, 55]. Consequently, siRNAs and shRNAs transcribed by T7 RNA polymerase may induce an interferon response [56]. Some work suggests that the presence of 3' overhangs [57, 58] or that additional guanine residues at the 5' end or a 5' overhang abrogate interferon induction [58, 59]. For the transcribed shRNA, it should be verified that knockdown is due to RNAi and not interferon. While our design already contains a 3' overhang, modifications may be needed at the 5' end sequence if interferon induction is observed.

The proposed mechanism can be extended from RNAi to other cellular pathways that interact with nucleic acids. Many bacteria and archaea also use RNA to guide the destruction of foreign genetic elements [60, 61]. In response to viral or plasmid challenge, short nucleic acid elements from the invader are integrated into the host genome into a clustered regularly interspaced short palindromic repeat (CRISPR) [62–64]. The CRISPR loci is transcribed and the transcript is processed into short CRISPR-derived RNAs (crRNAs) [65] which guide the destruction of complementary genetic elements [64, 65]. Three types of CRISPR systems exist, which differ in the crRNA precursor (single-stranded vs. double-stranded RNA) [66, 67] and in whether they target DNA or RNA for destruction [68]. It has been demonstrated *in vitro* that engineered crRNAs can direct the cleavage of

a selected target [69, 70]. While the CRISPR mechanism is quite distinct from RNAi, the use of complementary nucleic acids for gene silencing suggests that the described mechanism can be adjusted to trigger gene silencing in bacteria as well by transcribing a crRNA precursor.

4.5 Materials and methods

Strand sequences. The hairpins and target for this mechanism are DNA. The T7 promoter sequence used is TAATACGACTCACTATAG, where **G** is the first base incorporated into the transcript. The shRNA sequence is GCAGCACGACUUCUUCAAGAGCUGACUUGAAGAAGUCGUGCUGCUU where the last two U's are incorporated from the transcription termination signal. The detection target used for this design is a random sequence, the transcribed shRNA targets nucleotides 240–258 on eGFP.

Strand	Sequence
X	ATAAGCCCTCATCCAAC
A	AGTTGGATGAGGGCTTATTAATACGACTCACTATAGCAGCACGACTT CTTCAAGAGCTGACTTGAAGAAGTCGTGCTGCTATAGTGAGATAAG CCCTC
B	CTCACTATAAAAAAAGCAGCACGACTTCTTCAAGTCAGCTCTTGAA GAAGTCGTGCTGCTATAGTGAGTCGTATTA

Table 4.1: List of strands for T7-promoter-based shRNA transcripton.

Oligonucleotides. Oligonucleotides were synthesized by Integrated DNA Technologies (IDT). DNA hairpins were synthesized as two pieces (each PAGE purified by IDT) which were then ligated to produce the full hairpin. Ligation was performed using T4 DNA ligase (New England Biolabs) and were purified using denaturing polyacrylamide gel electrophoresis followed by an ethanol precipitation.

Oligonucleotide concentrations were determined and adjusted using A_{260} absorbance on a NanoDrop8000 (Thermo Scientific). Further adjustments were performed by incubating different ratios of individual strands for 2 hours at 37°C followed by gel electrophoresis until correct stoichiometry was obtained.

Hairpins were snap-cooled by heating them to 95°C for 90 seconds followed by a 30 second incubation on ice and room temperature incubation of at least 30 minutes. Complexes were annealed by heating to 90°C for 3 minutes followed by a controlled gradual cooling at -1°C per minute to 23°C in a PCR block.

Polyacrylamide gel electrophoresis. Reactants were incubated at $0.5\mu\text{M}$ each for two hours at 37°C in $1\times$ SPSC buffer ($50\text{mM Na}_2\text{HPO}_4$, 0.5M NaCl , $\text{pH}7.5$). Native polyacrylamide gels were cast and run in $1\times$ TBE (Tris-Borate-EDTA) at 200V . Denaturing polyacrylamide gels were cast and run in $1\times$ TBE at 500V unless otherwise specified. Denaturing gels were pre-run at 500V for 1–2hr (unless otherwise specified). Gels were stained in $1\times$ SYBR Gold (Life Technologies) for 10 minutes at room temperature and imaged using an FLA-5100 imaging system (Fuji Photo Film).

***In vitro* transcription.** Oligonucleotides were snap cooled prior to transcription. Transcription was carried out using the T7-Scribe standard RNA IVT kit (CELLSCRIPT) according to the manufacturer, 2pmol of each DNA strand were included per $20\mu\text{l}$ reaction. For radioactive transcription only 50nmole of UTP was used in combination with $3\text{--}4\mu\text{l}$ of $[\alpha\text{--}^{32}\text{P}]$ UTP (10mCi/ml , MP Biomedicals). Transcription reactions were incubated for 3 hours at 37°C followed by 20 minutes of DnaseI treatment at 37°C . Transcription products were recovered by Organic extraction / Chromatography / Ethanol precipitation. The reaction volume was adjusted to $200\mu\text{l}$ using RNase-free water and extracted using 1:1 (v/v) TE-saturated phenol/chloroform. Unincorporated NTP's were removed from the aqueous phase by spin column chromatography using NucAway spin columns (Life Technologies) according to the manufacturer. Ethanol precipitation was done by incubation on ice for 15 minutes in 1:10 (v/v) of 3M sodium acetate and $2.5\times$ (v/v) 95% EtOH. The RNA was pelleted followed a 70% EtOH wash. The pellet was dried and resuspended in $1\times$ duplex buffer. Counts were measured on a Beckman LS-5000TD Liquid Scintillation Counter.

5' end labeling. microRNA and siRNA markers (New England Biolabs) were 5' end labeled with $[\gamma\text{--}^{32}\text{P}]$ ATP (10mCi/ml , MP Biomedicals) using T4 polynucleotide kinase (New England Biolabs). Unincorporated $[\gamma\text{--}^{32}\text{P}]$ ATP were removed by spin column chromatography using Illustra MicroSpin G-25 columns (GE Healthcare) according to the manufacturer.

***In vitro* Dicer assay.** Dicer reactions were performed using the Recombinant Human

Turbo Dicer Enzyme kit (Genlantis) according to the manufacturer with some modifications. The volume corresponding to 10,000–20,000 cpm of labeled shRNA from “ON” reaction X·A·B was used for the Dicer reactions (i.e., if 2 μ l of shRNA from X·A·B transcription reaction correspond to 20,000 cpm then 2 μ l were used from the other reactions as well). One Dicer unit per 20,000 cpm was used. Hairpins were snap cooled prior to Dicer reaction. Reactions were carried out for 2 hours at 37°C, reactions were stopped by the addition of the appropriate loading dye. siRNA formation was determined by EMSA. Radioactive gels were exposed overnight onto an image plate (Fujifilm type BAS-MS) and scanned using an FLA-5100 imaging system (Fuji Photo Film).

Quantification and band intensity plots. Multi Gauge ver2.0 (Fujifilm) software was used for quantification and intensity plot data. Bands were quantified using the “Quant Measure mode.” Data points for band intensity plots were gathered using the profile feature. ON-to-OFF ratio was determined by setting the ON ratio with a short target to 100%.

Northern blot. See miRNA blotting procedure with Ultrahyb-Oligo in materials and methods of Chapter 5. Biotin chromogenic detection (Thermo Scientific) according to the manufacturer. Probe: 5'-biotin-AAG CAG CAC GAC TTC TTC AAG TCA GCT CTT GAA GAA GTC GTG CTG C -3'

HEK293 d2EGFP-T7 RNA polymerase cell line generation. HEK293 d2EGFP cells were a generous gift from Dr. Chase Beisel. T7 RNA polymerase was cloned from pT7POL26 (Gentaur) into pIRESHyg3 vector (clontech) by Keyclone Technologies to generate pIRESHyg-T7 pol. pIRESHyg-T7 pol plasmid was linearized with Bstz17I (NEB) followed by nucleotide removal using QIAquick nucleotide removal kit (Qiagen). 400ng of linearized plasmid were transfected into 293 d2EGFP cells using Attractene transfection reagent (Qiagen) according to the manufacturer. Transfected cells were selected in DMEM supplemented with 10% FBS, 1% Penicillin-Streptomycin 125 μ g/ml Hygromycin B. Cells were grown at 37°C 5% CO₂.

Western blot. Protein was extracted from cells using CellLytic M (Sigma-Aldrich) ac-

ording to the manufacturer. Bradford reagent (Sigma Aldrich) was used to determine the amount of protein according to the manufacturer. 36 μ g of protein were denatured by heating to 95°C for five minutes and separated on a 10% TRIS-HCL SDS polyacrylamide gel (Bio-Rad). Protein was transferred to a nitrocellulose membrane using semi dry transfer (Owl separation systems). The membrane was blocked overnight at 4°C in 1 \times TBST (10mM Tris, 150mM NaCl, 0.05%(v/v) Tween20, pH 7.5) 5% non-fat milk and 2% BSA. Following blocking the membrane was washed 3 times in 1 \times TBST and then blotted in 1 \times TBST with 1:1,000 mouse anti T7 RNA polymerase monoclonal antibody (Novagen catalog #70566-3). The membrane was then washed 3 times in 1 \times TBST and incubated in 1 \times TBST containing 1:2000 Anti-mouse IgG HRP-linked antibody (Cell Signaling catalog #7076). The membrane was washed in 1 \times TBST and signal was developed using SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific) according to the manufacturer.

Plasmid transfection for reverse transcription PCR. The pT7CAT plasmid was a generous gift from Dr. Bernard Moss (NIH). Transfections were carried out using Lipofectamine2000 (Invitrogen) according to the manufacturer. RNA was extracted 24 hours post transfection using Trizol (Invitrogen) followed by a DnaseI treatment (NEB) according to the manufacturer. EDTA was added to the sample to a 0.5M final concentration and the Dnase was heat inactivated by a 10 minute incubation at 4°C. cDNA was generated using the High Capacity cDNA Reverse Transcription kit (Applied Biosystems catalog #4368814) according to the manufacturer. The cDNA was then PCR amplified using the Taq PCR Core kit (Qiagen catalog #201223) with the following program: 3 minute incubation at 94°C followed by 30 cycles of 0.5 minute at 94°C, 0.5 minute at 60°C, 1 minute at 72°C, an additional 10 minute incubation at 72°C was added at the end. CAT primers: Forward 5'-ATTCACATTCTTGCCCGCCTGATG-3' Reverse 5'-GGAAGCCATCACAAACGGCATGAT-3'

T7 polymerase primers: Forward 5'-AACTCCCGATGAAACCGGAAGACA-3' Reverse

5'-ACCTTGCGGGTTGAACATTGACAC-3'

Bibliography

- [1] Brummelkamp, T. R., Bernards, R., and Agami, R. A system for stable expression of short interfering RNAs in mammalian cells. *Science* **296**(5567), 550–3 (2002).
- [2] Paddison, P. J., Caudy, A. A., Bernstein, E., Hannon, G. J., and Conklin, D. S. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes & development* **16**(8), 948–58 (2002).
- [3] Paul, C. P., Good, P. D., Winer, I., and Engelke, D. R. Effective expression of small interfering RNA in human cells. *Nature biotechnology* **20**(5), 505–8 (2002).
- [4] Sui, G., Soohoo, C., Affar el, B., Gay, F., Shi, Y., and Forrester, W. C. A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* **99**(8), 5515–20 (2002).
- [5] Yi, R., Qin, Y., Macara, I. G., and Cullen, B. R. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes & development* **17**(24), 3011–6 (2003).
- [6] Lund, E., Guttinger, S., Calado, A., Dahlberg, J. E., and Kutay, U. Nuclear export of microRNA precursors. *Science* **303**(5654), 95–8 (2004).
- [7] Baer, M., Nilsen, T. W., Costigan, C., and Altman, S. Structure and transcription of a human gene for H1 RNA, the RNA component of human RNase P. *Nucleic acids research* **18**(1), 97–103 (1990).
- [8] Amarzguioui, M., Rossi, J. J., and Kim, D. Approaches for chemically synthesized siRNA and vector-mediated RNAi. *FEBS letters* **579**(26), 5974–81 (2005).
- [9] Zeng, Y., Wagner, E. J., and Cullen, B. R. Both natural and designed micro RNAs can

- inhibit the expression of cognate mRNAs when expressed in human cells. *Molecular cell* **9**(6), 1327–33 (2002).
- [10] Lee, S. K. and Kumar, P. Conditional RNAi: towards a silent gene therapy. *Advanced drug delivery reviews* **61**(7-8), 650–64 (2009).
- [11] Yi, R., Doehle, B. P., Qin, Y., Macara, I. G., and Cullen, B. R. Overexpression of exportin 5 enhances RNA interference mediated by short hairpin RNAs and microRNAs. *RNA* **11**(2), 220–6 (2005).
- [12] Grimm, D., Streetz, K. L., Jopling, C. L., Storm, T. A., Pandey, K., Davis, C. R., Marion, P., Salazar, F., and Kay, M. A. Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature* **441**(7092), 537–41 (2006).
- [13] Cullere, X., Lauterbach, M., Tsuboi, N., and Mayadas, T. N. Neutrophil-selective CD18 silencing using RNA interference in vivo. *Blood* **111**(7), 3591–8 (2008).
- [14] Giering, J. C., Grimm, D., Storm, T. A., and Kay, M. A. Expression of shRNA from a tissue-specific pol II promoter is an effective and safe RNAi therapeutic. *Molecular therapy : the journal of the American Society of Gene Therapy* **16**(9), 1630–6 (2008).
- [15] van de Wetering, M., Oving, I., Muncan, V., Pon Fong, M. T., Brantjes, H., van Leenen, D., Holstege, F. C., Brummelkamp, T. R., Agami, R., and Clevers, H. Specific inhibition of gene expression using a stably integrated, inducible small-interfering-RNA vector. *EMBO reports* **4**(6), 609–15 (2003).
- [16] Czauderna, F., Santel, A., Hinz, M., Fechtner, M., Durieux, B., Fisch, G., Leenders, F., Arnold, W., Giese, K., Klippel, A., and Kaufmann, J. Inducible shRNA expression for application in a prostate cancer mouse model. *Nucleic acids research* **31**(21), e127 (2003).

- [17] Lin, X., Yang, J., Chen, J., Gunasekera, A., Fesik, S. W., and Shen, Y. Development of a tightly regulated U6 promoter for shRNA expression. *FEBS letters* **577**(3), 376–80 (2004).
- [18] Higuchi, M., Tsutsumi, R., Higashi, H., and Hatakeyama, M. Conditional gene silencing utilizing the lac repressor reveals a role of SHP-2 in cagA-positive *Helicobacter pylori* pathogenicity. *Cancer science* **95**(5), 442–7 (2004).
- [19] Gupta, S., Schoer, R. A., Egan, J. E., Hannon, G. J., and Mittal, V. Inducible, reversible, and stable RNA interference in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* **101**(7), 1927–32 (2004).
- [20] Kasim, V., Miyagishi, M., and Taira, K. Control of siRNA expression using the Cre-loxP recombination system. *Nucleic acids research* **32**(7), e66 (2004).
- [21] Tiscornia, G., Tergaonkar, V., Galimi, F., and Verma, I. M. CRE recombinase-inducible RNA interference mediated by lentiviral vectors. *Proceedings of the National Academy of Sciences of the United States of America* **101**(19), 7347–51 (2004).
- [22] Ventura, A., Meissner, A., Dillon, C. P., McManus, M., Sharp, P. A., Van Parijs, L., Jaenisch, R., and Jacks, T. Cre-lox-regulated conditional RNA interference from transgenes. *Proceedings of the National Academy of Sciences of the United States of America* **101**(28), 10380–5 (2004).
- [23] An, C. I., Trinh, V. B., and Yokobayashi, Y. Artificial control of gene expression in mammalian cells by modulating RNA interference through aptamer-small molecule interaction. *RNA* **12**(5), 710–6 (2006).
- [24] Beisel, C. L., Bayer, T. S., Hoff, K. G., and Smolke, C. D. Model-guided design of ligand-regulated RNAi for programmable control of gene expression. *Molecular systems biology* **4**, 224 (2008).

- [25] Tuleuova, N., An, C. I., Ramanculov, E., Revzin, A., and Yokobayashi, Y. Modulating endogenous gene expression of mammalian cells via RNA-small molecule interaction. *Biochemical and biophysical research communications* **376**(1), 169–73 (2008).
- [26] Kumar, D., An, C. I., and Yokobayashi, Y. Conditional RNA interference mediated by allosteric ribozyme. *Journal of the American Chemical Society* **131**(39), 13906–7 (2009).
- [27] Beisel, C. L., Chen, Y. Y., Culler, S. J., Hoff, K. G., and Smolke, C. D. Design of small molecule-responsive microRNAs based on structural requirements for Droscha processing. *Nucleic acids research* **39**(7), 2981–94 (2011).
- [28] Kim, J., White, K. S., and Winfree, E. Construction of an in vitro bistable circuit from synthetic transcriptional switches. *Molecular systems biology* **2**, 68 (2006).
- [29] Collingwood, M. A., Rose, S. D., Huang, L., Hillier, C., Amarzguioui, M., Wiüger, M. T., Soifer, H. S., Rossi, J. J., and Behlke, M. A. Chemical modification patterns compatible with high potency Dicer-substrate small interfering RNAs. *Oligonucleotides* **18**(2), 187–200 (2008).
- [30] Ge, Q., Ilves, H., Dallas, A., Kumar, P., Shorestein, J., Kazakov, S. A., and Johnston, B. H. Minimal-length short hairpin RNAs: the relationship of structure and RNAi activity. *RNA* **16**(1), 106–17 (2010).
- [31] Zadeh, J. N., Steenberg, C. D., Bois, J. S., Wolfe, B. R., Pierce, M. B., Khan, A. R., Dirks, R. M., and Pierce, N. A. NUPACK: Analysis and design of nucleic acid systems. *Journal of computational chemistry* **32**(1), 170–3 (2011).
- [32] Zadeh, J. N., Wolfe, B. R., and Pierce, N. A. Nucleic acid sequence design via efficient ensemble defect optimization. *Journal of computational chemistry* **32**(3), 439–52 (2011).

- [33] Ohrt, T., Merkle, D., Birkenfeld, K., Echeverri, C. J., and Schwille, P. In situ fluorescence analysis demonstrates active siRNA exclusion from the nucleus by exportin 5. *Nucleic acids research* **34**(5), 1369–80 (2006).
- [34] Dirks, R. M., Lin, M., Winfree, E., and Pierce, N. A. Paradigms for computational nucleic acid design. *Nucleic acids research* **32**(4), 1392–403 (2004).
- [35] Dirks, R. M. and Pierce, N. A. A partition function algorithm for nucleic acid secondary structure including pseudoknots. *Journal of computational chemistry* **24**(13), 1664–77 (2003).
- [36] Dirks, R. M. and Pierce, N. A. An algorithm for computing nucleic acid base-pairing probabilities including pseudoknots. *Journal of computational chemistry* **25**(10), 1295–304 (2004).
- [37] Dirks, R. M., Bois, J. S., M., S. J., Winfree, E., and Pierce, N. A. Thermodynamic analysis of interacting nucleic acid strands. *SIAM Review* **49**, 65–88 (2007).
- [38] Adhya, S., Basu, S., Sarkar, P., and Maitra, U. Location, function, and nucleotide sequence of a promoter for bacteriophage T3 RNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America* **78**(1), 147–51 (1981).
- [39] Stump, W. T. and Hall, K. B. SP6 RNA polymerase efficiently synthesizes RNA from short double-stranded DNA templates. *Nucleic acids research* **21**(23), 5480–4 (1993).
- [40] Davanloo, P., Rosenberg, A. H., Dunn, J. J., and Studier, F. W. Cloning and expression of the gene for bacteriophage T7 RNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America* **81**(7), 2035–9 (1984).
- [41] Myslinski, E., Ame, J. C., Krol, A., and Carbon, P. An unusually compact external promoter for RNA polymerase III transcription of the human H1 RNA gene. *Nucleic acids research* **29**(12), 2502–9 (2001).

- [42] Dunn, J. J. and Studier, F. W. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *Journal of molecular biology* **166**(4), 477–535 (1983).
- [43] Jeng, S. T., Gardner, J. F., and Gumpert, R. I. Transcription termination by bacteriophage T7 RNA polymerase at rho-independent terminators. *The Journal of biological chemistry* **265**(7), 3823–30 (1990).
- [44] Macdonald, L. E., Zhou, Y., and McAllister, W. T. Termination and slippage by bacteriophage T7 RNA polymerase. *Journal of molecular biology* **232**(4), 1030–47 (1993).
- [45] Yu, J. Y., DeRuiter, S. L., and Turner, D. L. RNA interference by expression of short-interfering RNAs and hairpin RNAs in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* **99**(9), 6047–52 (2002).
- [46] Kim, J. and Winfree, E. Synthetic in vitro transcriptional oscillators. *Molecular systems biology* **7**, 465 (2011).
- [47] Hamdorf, M., Muckenfuss, H., Tschulena, U., Pleschka, S., Sanzenbacher, R., Cichutek, K., and Flory, E. An inducible T7 RNA polymerase-dependent plasmid system. *Molecular biotechnology* **33**(1), 13–21 (2006).
- [48] Peng, Y., Lu, J. X., and Shen, X. F. shRNA driven by Pol II/T7 dual-promoter system effectively induce cell-specific RNA interference in mammalian cells. *Biochemical and biophysical research communications* **360**(2), 496–500 (2007).
- [49] Wang, T. H., Yu, S. H., and Au, L. C. Facilitated in vivo synthesis of ribonucleic acid and protein via T7 RNA polymerase. *Analytical biochemistry* **375**(1), 97–104 (2008).
- [50] Vermeulen, A., Behlen, L., Reynolds, A., Wolfson, A., Marshall, W. S., Karpilow,

- J., and Khvorova, A. The contributions of dsRNA structure to Dicer specificity and efficiency. *RNA* **11**(5), 674–82 (2005).
- [51] Elroy-Stein, O., Fuerst, T. R., and Moss, B. Cap-independent translation of mRNA conferred by encephalomyocarditis virus 5' sequence improves the performance of the vaccinia virus/bacteriophage T7 hybrid expression system. *Proceedings of the National Academy of Sciences of the United States of America* **86**(16), 6126–30 (1989).
- [52] Castanotto, D., Li, H., and Rossi, J. J. Functional siRNA expression from transfected PCR products. *RNA* **8**(11), 1454–60 (2002).
- [53] Zheng, L., Liu, J., Batalov, S., Zhou, D., Orth, A., Ding, S., and Schultz, P. G. An approach to genomewide screens of expressed small interfering RNAs in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* **101**(1), 135–40 (2004).
- [54] Hornung, V., Ellegast, J., Kim, S., Brzozka, K., Jung, A., Kato, H., Poeck, H., Akira, S., Conzelmann, K. K., Schlee, M., Endres, S., and Hartmann, G. 5'-triphosphate RNA is the ligand for RIG-I. *Science* **314**(5801), 994–7 (2006).
- [55] Pichlmair, A., Schulz, O., Tan, C. P., Naslund, T. I., Liljestrom, P., Weber, F., and Reis e Sousa, C. RIG-I-mediated antiviral responses to single-stranded RNA bearing 5'-phosphates. *Science* **314**(5801), 997–1001 (2006).
- [56] Kim, D. H., Longo, M., Han, Y., Lundberg, P., Cantin, E., and Rossi, J. J. Interferon induction by siRNAs and ssRNAs synthesized by phage polymerase. *Nature biotechnology* **22**(3), 321–5 (2004).
- [57] Marques, J. T., Devosse, T., Wang, D., Zamanian-Daryoush, M., Serbinowski, P., Hartmann, R., Fujita, T., Behlke, M. A., and Williams, B. R. A structural basis for discriminating between self and nonself double-stranded RNAs in mammalian cells. *Nature biotechnology* **24**(5), 559–65 (2006).

- [58] Schlee, M., Roth, A., Hornung, V., Hagmann, C. A., Wimmenauer, V., Barchet, W., Coch, C., Janke, M., Mihailovic, A., Wardle, G., Juranek, S., Kato, H., Kawai, T., Poeck, H., Fitzgerald, K. A., Takeuchi, O., Akira, S., Tuschl, T., Latz, E., Ludwig, J., and Hartmann, G. Recognition of 5' triphosphate by RIG-I helicase requires short blunt double-stranded RNA as contained in panhandle of negative-strand virus. *Immunity* **31**(1), 25–34 (2009).
- [59] Gondai, T., Yamaguchi, K., Miyano-Kurosaki, N., Habu, Y., and Takaku, H. Short-hairpin RNAs synthesized by T7 phage polymerase do not induce interferon. *Nucleic acids research* **36**(3), e18 (2008).
- [60] Sorek, R., Kunin, V., and Hugenholtz, P. Crispr—a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nature reviews. Microbiology* **6**(3), 181–6 (2008).
- [61] Horvath, P. and Barrangou, R. CRISPR/Cas, the immune system of bacteria and archaea. *Science* **327**(5962), 167–70 (2010).
- [62] Andersson, A. F. and Banfield, J. F. Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* **320**(5879), 1047–50 (2008).
- [63] Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., and Horvath, P. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**(5819), 1709–12 (2007).
- [64] Garneau, J. E., Dupuis, M. E., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadan, A. H., and Moineau, S. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**(7320), 67–71 (2010).
- [65] Brouns, S. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J., Snijders, A. P., Dickman, M. J., Makarova, K. S., Koonin, E. V., and van der Oost, J. Small

- CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**(5891), 960–4 (2008).
- [66] Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., Eckert, M. R., Vogel, J., and Charpentier, E. CRISPR RNA maturation by trans-encoded small RNA and host factor *rnase III*. *Nature* **471**(7340), 602–7 (2011).
- [67] Wiedenheft, B., Sternberg, S. H., and Doudna, J. A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **482**(7385), 331–8 (2012).
- [68] Hale, C. R., Zhao, P., Olson, S., Duff, M. O., Graveley, B. R., Wells, L., Terns, R. M., and Terns, M. P. RNA-guided RNA cleavage by a *crispr* RNA-cas protein complex. *Cell* **139**(5), 945–56 (2009).
- [69] Hale, C. R., Majumdar, S., Elmore, J., Pfister, N., Compton, M., Olson, S., Resch, A. M., Glover, C. V., r., Graveley, B. R., Terns, R. M., and Terns, M. P. Essential features and rational design of *crispr* RNAs that function with the cas ramp module complex to cleave RNAs. *Molecular cell* **45**(3), 292–302 (2012).
- [70] Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**(6096), 816–21 (2012).

Chapter 5

Sensitive multiplexed northern blots via hybridization chain reaction (HCR)

5.1 Introduction

In Chapter 2, we presented a mechanism for conditional RNAi activation based on scRNAs. Despite being able to design a system that is functional *in vitro*, the mechanism did not lead to gene silencing in tissue culture. One concern was the degradation of scRNAs in cells. In Chapter 3, we examined the fate of transfected or expressed scRNAs using northern hybridization. Since the mechanism is comprised of three different scRNAs, three separate northern blots are needed (one per scRNA). For expressed scRNAs, when no signal was detected on the blot, a concern we had was whether the scRNAs are being expressed, or whether the expression level is too low to be detected. Detection of low expression levels might also be a concern for conditional shRNA transcription presented in Chapter 4; this has not been studied. To solve these issues, our aim was to develop a sensitive multiplexed northern blot hybridization assay. For gene expression studies, a sensitive multiplexed assay would allow blotting a control reference gene as well as multiple genes of interest in one single experiment.

5.1.1 Northern blot

Northern blot is a hybridization-based technique which was developed as a variation of an older method used to detect DNA, the Southern blot [1]. The purpose of this method is to identify the presence and size of a specific RNA within a sample. The northern blot protocol includes the following main steps: RNA isolation, electrophoretic separation of RNA under denaturing conditions, RNA transfer and crosslinking to a membrane, hybridization of the membrane with a probe for the gene of interest and visualization of signal [2, 3]. The location of the signal obtained by the probe indicates the size of the RNA while the intensity of the signal provides information about the quantity of bound RNA.

Northern blot is not typically used for quantitative analysis and is considered to be semi-quantitative, providing information regarding relative RNA expression within a sample or across samples [3, 4]. To compare RNA between different samples, at least two probes must be used. The first probe targets the gene of interest and the second probe is used as a normalizing control, targeting a second gene that should be present in equal amounts across all samples. Though it is possible to hybridize the sample and the control probe concurrently [5] (if the two targets have distinct size), this is not commonly practiced. Instead, after hybridization and visualization of the first probe, the membrane is re-stripped of the bound probe and hybridized again to detect a second target (e.g., the control). Stripping and re-probing of the membrane can be difficult and reduce sensitivity [2].

5.1.2 Northern blot probes and detection strategies

Northern blot is a versatile technique; the length, base chemistry, and detection moiety can all be changed [2]. Short oligonucleotide probes [6, 7] can be used as well as full- [5] or partial-length cDNA [8]. Because hybridization is based on base pairing, alternative bases to DNA and RNA can be used, depending on probe length. For example, short probes can be made of 2'-OMe, locked nucleic acids (LNA) [9, 10], etc. For detection, both radioactive and non-radioactive probes can be used. Labeling strategies include end-

labeling [11, 12] as well as uniform probe labeling. Uniform labeling can be achieved using random-primer [13], nick translation [14, 15] and transcription reactions [16–18], all of which are rapid methods that provide high-density labeling. Due to its high sensitivity, radioactive labeling with ^{32}P is most commonly used. However, the health hazard, safety measurements and special training associated with radiation use as well as short half-life of the probe are a great disadvantage. Non-isotopic probes include the use of fluorescently labeled nucleotides or haptens (e.g digoxigenin or biotin) [19–22] which are incorporated into the probe. While fluorescent probes can be visualized directly, haptens are detected by an antibody conjugated to alkaline phosphatase (AP) or horseradish peroxidase (HRP) via chemiluminescent detection [6, 8, 23]. Similarly, both AP and HRP can be directly conjugated to the probe and assayed by chemiluminescence [21].

5.1.3 Northern blot in comparison to other methods

While northern blot is widely used throughout the field, new techniques offer advantages over it. Mainly, most newer techniques are considered to be more sensitive, and they also tend to be less sensitive to RNA degradation. Real-time PCR, nuclease protection assays, and fluorescent *in situ* hybridization (FISH) allow examination of multiple genes at once, and microarrays are high-throughput. FISH does not require the isolation of RNA, and in addition, provides information about RNA localization within the cell or tissue. However, no single method provides comprehensive information and so a thorough analysis of gene expression often requires the use of multiple techniques. The only method that provides information regarding both sequence and length is the northern blot. For this reason, northern blot is still widely used for RNA detection and to validate results obtained with other methods. It also provides information about the RNA condition (e.g., degradation) and it can distinguish between splice variants [2, 3].

As previously mentioned, the need to strip and re-probe a northern blot membrane in order to detect multiple targets is a limitation of the technique. To the best of our

knowledge, only one attempt has been made to develop a multiplexed northern blot assay. Hoeltke et al. [24] have developed a colorimetric multiplexed northern blot protocol. Differentially labeled probes have been used against two different targets. Probe labeling was done with digoxigenin or biotin. Each label is detected by alkaline phosphatase conjugate and three different naphthol-AS³-phosphate/diazonium salt combinations as substrates to AP. While hybridization to the different probes can be carried out simultaneously, label detection is carried out sequentially. Between each detection step, a heat/EDTA treatment is carried out to inactivate the formerly bound AP. The results are visualized as blue and red bands; if the target is bound by more than one probe, the signal results in a mixed color. The great advantage of this method is that it removes the need to strip and re-probe a membrane for detection of more than one target. However, it is still necessary to perform the detection reaction consecutively. Unfortunately, the dyes used in order to detect multiple colors at once are less sensitive than other colorimetric detection dyes; no other dye combination with improved sensitivity has been reported, thus making this method less applicable. Fluorescently labeled northern blot probes can be generated, however, they are considered to be less sensitive than other detection methods and are therefore not widely used [21].

In this chapter, we explore the use of a hybridization chain reaction (HCR) in order to simultaneously detect multiple targets on a northern blot.

5.1.4 Hybridization chain reaction (HCR)

In an HCR system [25], two complementary hairpins are designed to be kinetically trapped in a monomer state. Each hairpin has two single-stranded regions: a toehold and a loop. The complement to the toehold of one hairpin is sequestered in the loop of the other hairpin (and vice versa), kinetically trapping the hairpins. Upon the presence of an initiator, a cascade of reactions begins in which the hairpins form a nicked double-stranded polymer. In Figure 5.1(a) hairpins H1 and H2 are kinetically trapped. Upon the presence of initiator

I in Figure 5.1(b) the toehold ‘a’ of hairpin H1 base-pairs with its complementary target ‘a*’ in initiator I. Next, the stem ‘b’ of hairpin H1 binds to ‘b*’ in the initiator, the complex I·H1 is formed. Following the binding of H1 to the initiator I the loop ‘c’ and half of the stem ‘b*’ of hairpin H1 are exposed as single-stranded regions. In the next step (Figure 5.1(c)), toehold ‘c*’ of hairpin H2 binds to ‘c’ in the complex I·H1, followed by binding of the stems ‘b-b*’ and in the formation of a complex I·H1·H2. Once hairpin H2 is bound and open, a single-stranded region ‘a* -b*’ is exposed. This region is identical to the original initiator sequence and thus a new H1 hairpin can add to the polymer followed by another H2 hairpin, leading to the formation of a long alternating H1·H2 polymer.

Our lab has recently utilized HCR to develop a multiplexed fluorescent *in situ* hybridization method [26]. Gene probes were designed to include a single-stranded “tail” region which is an HCR initiator sequence. Orthogonal HCR polymers were designed, each initiated by a different initiator sequence. The HCR polymers are made fluorescent by attaching fluorescent dyes to the hairpin monomers. Multiple fluorescent polymers can be grown off the same mRNA by using multiple probes (with the same initiator sequence) against a target gene. Using this method, five mRNAs were detected in a single experiment within a single biological sample. In order to detect other mRNAs, all that is needed is to design a new probe complementary to a gene of interest.

In this chapter, HCR is utilized to improve the northern blot method. A great disadvantage of a northern blot is that in order to detect multiple genes, current practice requires stripping and re-probing the blot. This is time consuming, difficult and may reduce sensitivity of the blot. By using multiplexed fluorescent HCR as a detection moiety of probe hybridization, we can detect four targets simultaneously; this is demonstrated for both mRNA targets and miRNA targets. Fluorescence detection is not as sensitive as the standard radio-isotope ^{32}P . By using a radioactive HCR polymer, we demonstrate miRNA detection of 0.01 femtomol, a 5-fold improvement relative to the current literature [10].

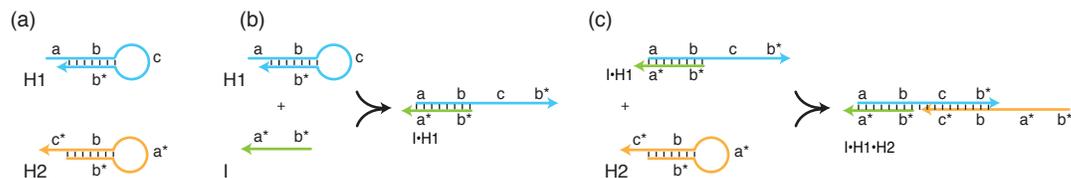


Figure 5.1: **HCR polymer formations schematic.** Letters marked with * are complementary to the corresponding unmarked letter (e.g a is complementary to a*). (a) Hairpins H1 and H2 are kinetically trapped in monomer state. (b) In the presence of initiator I hairpin H1 binds to the initiator and undergoes a conformation change, exposing the loop stem, forming complex I·H1. (c) The exposed loop and stem of hairpin H1 enable hairpin H2 to bind, forming a complex I·H1·H2 and exposing its loop and stem. The open form of H2 is identical to the initiator, enabling to another H1 hairpin to bind. A hybridization chain reaction occurs which forms a nicked double-stranded alternating H1 and H2 hairpins.

5.2 HCR as a detection method for northern blots

An HCR polymer can be used to detect and amplify the signal of a northern blot probe. Figure 5.2 depicts an overview of the northern blot procedure with HCR amplification as the signal detection step. To use HCR for detection, the probe design needs to be altered from that of “standard” probes. Traditional probes have a region complementary to the target RNA as well as a detection moiety such as a radioactive or hapten label. For HCR detection, the detection moiety of the probe is changed into a nucleotide sequence which is an HCR initiator. Once the probe is hybridized to its target, an HCR initiator “tail” is left unbound to the membrane (Figure 5.2, probe hybridization). After unbound probe is washed, HCR amplifier hairpins are added to the blot and an HCR polymer is grown off of each target-bound probe. Each amplifier hairpin is labeled, therefore each polymer contains multiple labels resulting in signal amplification (Figure 5.2, HCR amplification). The unbound hairpins are then washed and the signal is detected accordingly.

The HCR amplifiers are orthogonal. By using a different initiator per target-specific probe, a different polymer can be grown off multiple types of targets (Figure 5.2, red, blue and yellow polymers). The signal can be further amplified by growing multiple HCR

polymers per target. This can be achieved by using multiple probes per target, each with the same initiator sequence. For detection, any molecule used for standard northern blot probes can be used. Because HCR systems are orthogonal, multiplexing is inherent. Thus, as long as the probes are selective enough, multiple targets can be detected simultaneously. For multiplexing, distinct fluorescent labels should be used for each HCR amplifier.

5.3 Results

5.3.1 HCR “dot blot”

We used a dot blot to show that HCR can be used to detect a target mRNA in a blot format. To that end, we used a dot-blot approach, where an mRNA of interest is spotted in known amounts onto a membrane and detected using fluorescent HCR. Choi et al. [26] have previously used probes designed to target enhanced green fluorescent protein (EGFP) for *in situ* hybridization using HCR amplification. One of these probes was used for the dot blot since it was pre-validated and a plasmid for *in vitro* transcription of EGFP was present in the lab. This probe is 81 nucleotides long with a 50-nucleotide gene specific region, a 5-base spacer and a 26-nucleotide HCR initiator. Mild blotting conditions were used to promote probe binding since no hybridization to other targets was expected. Blotting conditions were adapted from *mirVana* miRNA isolation kit (Ambion).

Figure 5.3 demonstrates that HCR polymers can indeed be grown off an mRNA-bound probe attached to a positively charged nylon membrane. Figure 5.3(a) suggests a sensitivity estimated at three femtomoles. A second blot was performed using smaller amounts of spotted mRNA to see whether the sensitivity achieved in Figure 5.3(a) is due to full consumption of amplifiers or due to a sensitivity limit. A slight improvement in signal was observed when less RNA was spotted (Figure 5.3(b)). Taken together, both blots suggest a detection limit on the order of one to three femtomoles. This is less sensitive than commercial buffers such as ULTRAhyb Ultrasensitive Hybridization Buffer (Ambion), which

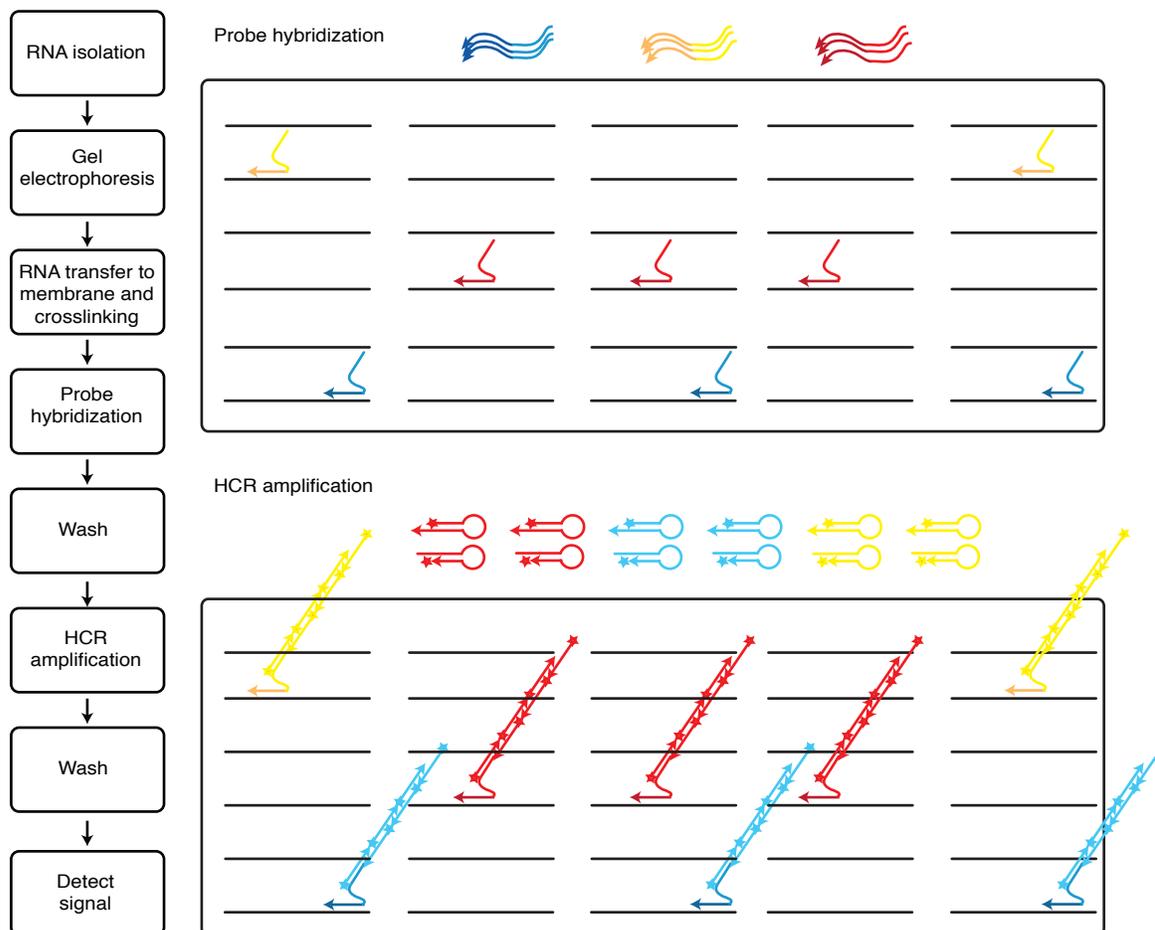


Figure 5.2: **Northern blot detection via HCR schematic.** Each probe contains an HCR initiator sequence. A different HCR polymer can be grown off each target, according to the initiator sequence. The two colors of each probe symbol the target binding sequence and the initiator sequence. Blue hairpins can form a polymer off blue initiator sequences etc. Arrows on probes and hairpins represent the 3' end. Stars represent the detection moiety.

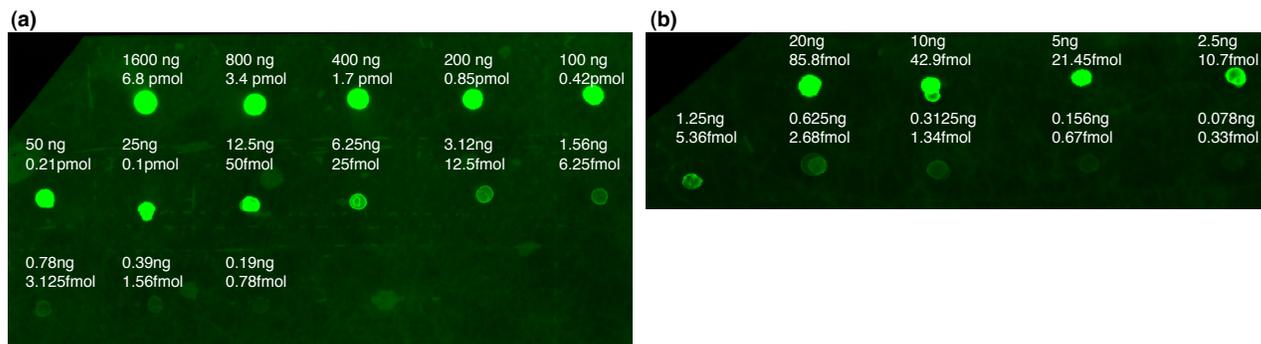


Figure 5.3: **EGFP "dot blot" detection via HCR.** Serial dilutions of *in vitro* transcribed EGFP mRNA were spotted onto a Nytran SPC (Whatman) membrane. RNA was cross-linked to the membrane by baking for 2 hours at 80°C. Blotting was done in mild conditions according to the *mirVana* miRNA blotting procedure (see Materials and methods). The EGFP probes (see Table 5.2.) were used at a 2nM concentration. Probe detection was carried out using a 10nM (each) amplifier solution in hybridization buffer. Amplifiers were snap cooled prior to hybridization and amplification step was carried out over night at room temperature in the dark (system A1, see Table 5.3). The amount of moles per spot were estimated based on a EGFP mRNA molecular weight of 223069.33 gr/mole. (a) Two-fold dilution series starting with 1600 nanogram EGFP mRNA. (b) Two-fold dilution series starting with 20 nanogram EGFP mRNA.

advertise the detection of 10,000 molecules (attomolar range). This sensitivity is based on the use of very pure and highly radioactive probes with prolonged exposure times (days) and 100% efficiency in probe hybridization to the target — conditions which are not standardly achieved. It is possible that the use of a different buffer such as ULTRAhyb might improve the hybridization conditions.

5.3.2 mRNA blots

5.3.2.1 Specificity and blotting conditions

To check for specificity of the assay (probe binding and non-specific HCR amplification) and to examine blotting conditions, destabilized enhanced green fluorescent protein (d2EGFP) was chosen as the detection target. This gene was chosen mainly due to its visible phenotype

in tissue culture and due to the availability of cell lines with and without d2EGFP. In addition, in our lab we have multiple probes against EGFP which have been validated *in situ*. These probes also match up with the d2EGFP sequence. Probes are 81 nucleotides in total with a 50-nucleotide gene specific region, a 5-nucleotide spacer and a 26-nucleotide HCR initiator sequence.

mRNA blots from total RNA were attempted using the same conditions for the dot blot. These conditions were not stringent enough and led to non-specific probe binding, most likely to ribosomal RNA (293A vs. 293 d2EGFP lanes in Figure 5.4(a)). Blotting was also attempted with the same hybridization and amplification buffer conditions that were used for *in situ* HCR amplification [26]. These conditions were not compatible for blotting and resulted in no signal (including for *in vitro* transcribed mRNA; data not shown). Finally, probe hybridization and HCR amplification were attempted using ULTRAhyb hybridization buffer (Ambion). Prior to RNA transfer to the membrane, the gel was stained with SYBRGold to verify the integrity of the RNA (Figure 5.4(b)). Staining of ribosomal RNA demonstrates that the RNA has not been degraded; the amount of *in vitro* transcribed mRNA (30ng) is not visible by SYBRGold staining in these conditions (control mRNA lane, Figure 5.4(b)). Figure 5.4(c) demonstrates that an HCR signal is obtained only for lanes containing either *in vitro* transcribed EGFP mRNA (control mRNA lane) or total RNA from d2EGFP expressing cells but not for total RNA from non-expressing 293A cells. The lack of signal in total RNA from 293A cells is due to the lack of d2EGFP mRNA and not due to mRNA degradation as evident by Figure 5.4(b). Thus, these conditions are suitable for specific mRNA detection and HCR amplification.

5.3.2.2 Multiplexed detection of endogenous targets

So far we have shown that HCR can detect over-expressed d2EGFP mRNA. We now demonstrate HCR's ability to detect endogenous levels of mRNA expression using glyceraldehyde-3-phosphate dehydrogenase (GAPDH). We chose GAPDH due to its relatively high expres-

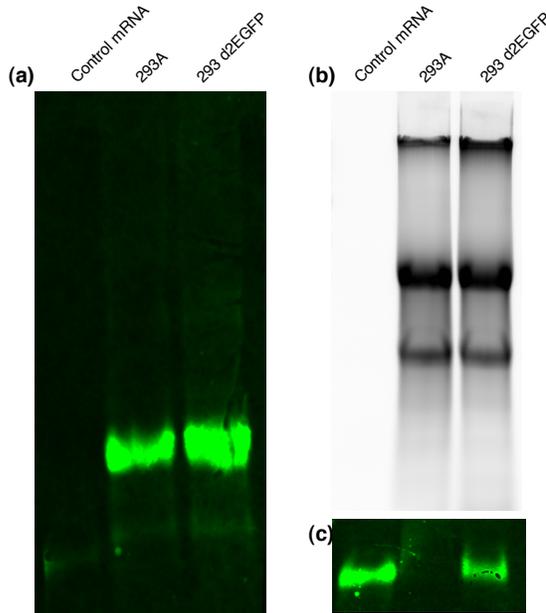


Figure 5.4: **Specific d2EGFP detection via HCR.** *In vitro* transcribed EGFP mRNA and total RNA extracted from 293A or 293 d2EGFP cells were separated on 1% denaturing agarose and transferred onto a Nytran SPC (Whatman) membrane. Following cross-linking and pre-hybridization a probe against d2EGFP (see Table 5.2) was added and hybridized overnight. The gel was washed three times before HCR amplification. The signal was amplified overnight using 14.3nM of each amplifier hairpin (system A3, see Table 5.3) labeled with Alexa 532. The membrane was washed 3 times prior to imaging. (a) Alexa 532 signal from blotted membrane. Blotting was done in mild conditions according to the *mirVana* miRNA blotting procedure (see Materials and methods section). 5.7nM of EGFP probe were used for hybridization. (b) SYBRGold staining of electrophorated RNA prior to transfer demonstrating the RNA is intact (30ng control mRNA and 22.5 μ g total RNA). (c) Alexa 532 signal from blotted membrane. Blotting conditions: pre-hybridization and hybridization with 9.17nM probe in ULTRAhyb. Two 5 minute washes at room temperature in low stringency buffer followed by a 5 minute wash in high stringency buffer at 55°C (NorthernMax kit, Ambion). After amplification the high stringency wash was done at 45°C.

sion level and because it is commonly used as a control in northern blot experiments. To demonstrate that the detected GAPDH mRNA runs corresponding to its size, a separate HCR system was used to detect a single-stranded RNA ladder (NEB). Probes are 81 nucleotides long with a 50-nucleotide gene-specific region, a 5-nucleotide spacer and a 26-nucleotide HCR initiator sequence. An HCR system labeled with Alexa 488 was used to detect an ssRNA ladder probe and an HCR system labeled with Alexa 647 was used to detect GAPDH probes. The expected length of GAPDH mRNA is 1401 nucleotides (NCBI accession NM_002046); a band corresponding to a size between 1000–2000 nucleotides is present in the Alexa 647 channel in Figure 5.5(a), suggesting that this is GAPDH. Both total RNA as well as total mRNA were used for this assay; total mRNA was used to control for non-specific binding to ribosomal RNA since it should be mostly depleted of ribosomal RNA. Figure 5.5(b) demonstrates that the 18S unit of the ribosomal RNA runs similar in size (1869 nucleotides, NCBI accession NR_003286) to the 2000nt marker, further supporting that the observed detected band is GAPDH and not ribosomal RNA.

To further validate that the observed signal is due to GAPDH detection and not ribosomal RNA (rRNA), a probe specific for 18S rRNA [27] was used in combination with probes against GAPDH and a single-stranded RNA (ssRNA) ladder. The 18S rRNA probe has a 20-nucleotide binding region whereas the other probes have a 50-nucleotide binding region. All probes have a binding region to the sequence of interest as well as an initiator for HCR. GAPDH, 18S rRNA and an ssRNA ladder were all detected using three orthogonal HCR systems; each system was labeled with a different fluorophore. Figure 5.6 demonstrates that a signal obtained for GAPDH (red) indeed runs lower than the 18S rRNA (green signal). In addition, it is observed that while the MicroPoly(A) purist kit does a good job of removing most of the 18S rRNA, some still remains as is observed by the green band in the total mRNA lane.

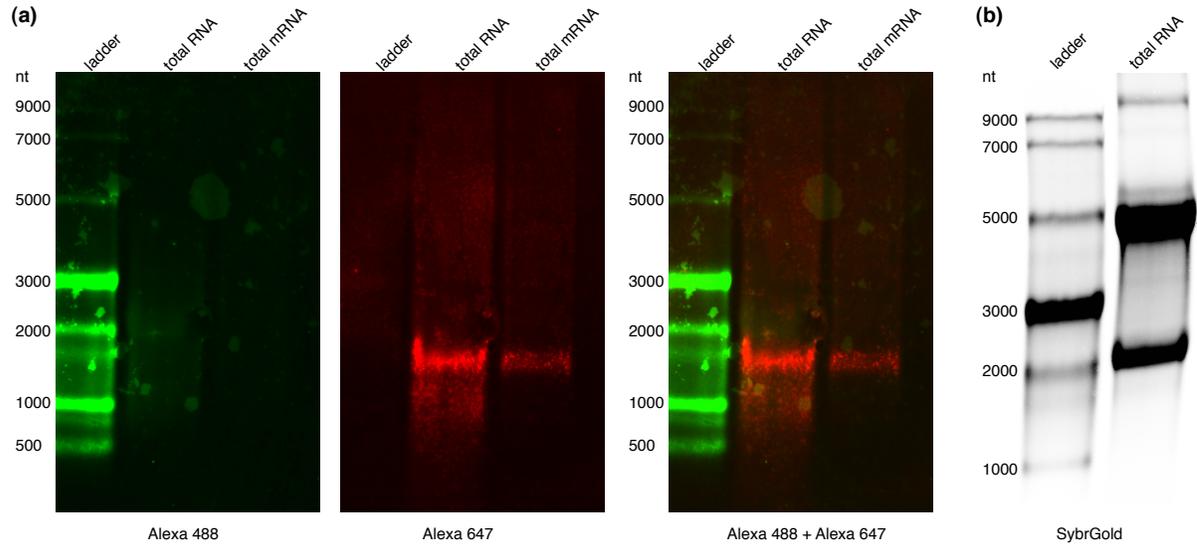


Figure 5.5: Multiplexed GAPDH and ssRNA ladder detection via HCR. Blotting was performed according to the mRNA blotting procedure (see Materials and methods). (a) Total RNA ($34.5\mu\text{g}$, Trizol extraction) or total mRNA (940ng , MicroPoly(A) purist (Ambion)) extracted from 293A cells were separated on 1% denaturing agarose and transferred onto a BrightStar-PLUS SPC (Whatman) membrane. As a size marker, $1.25\mu\text{g}$ ssRNA ladder (NEB) was used. Following cross-linking and pre-hybridization, five probes against GAPDH (each with an initiator for HCR system A1) were added to a 1.8nM final concentration (each), to detect the ladder, 16.3nM probe with an initiator for HCR system A2 was used. HCR amplification was performed at 20nM of each amplifier hairpin. Alexa 488 amplifiers bind to probes against ssRNA ladder, Alexa 647 amplifiers bind to probes against GAPDH. (b) SYBRGold staining of total RNA and ssRNA ladder demonstrates the RNA is intact and the size of ribosomal RNA bands. For a list of probes and HCR amplifier systems see Tables 5.2 and 5.3.

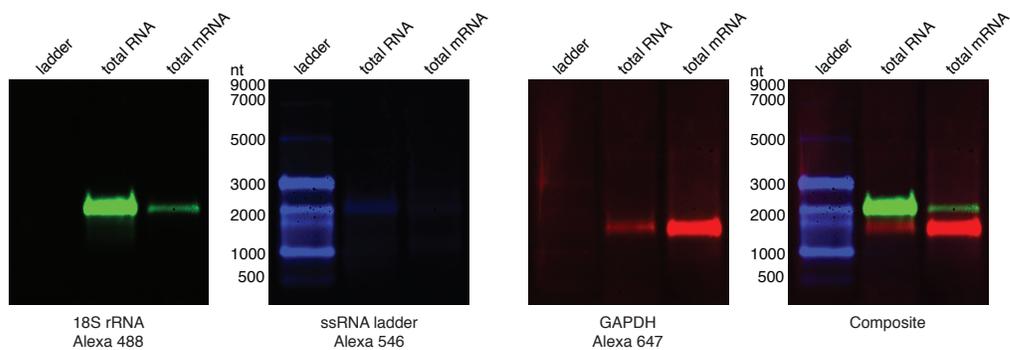


Figure 5.6: **Multiplexed detection of GAPDH, 18S rRNA and ssRNA ladder via HCR.** Total RNA ($8\mu\text{g}$, *mirVana* extraction) or total mRNA (920ng , MicroPoly(A) purist (Ambion)) extracted from 293A cells were separated on 1% denaturing agarose and transferred onto a BrightStar Plus (Ambion) membrane. As a size marker $1\mu\text{g}$, ssRNA ladder (NEB) was used. Blotting was performed according to the mRNA blotting procedure (see Materials and methods section). Five probes against GAPDH (each with an initiator for HCR system A1) were added to a 8.5nM final concentration (each). To detect the ladder, 8.5nM probe with an initiator for HCR system A2 was used. To detect 18S rRNA, 10nM probe with an initiator for HCR system A6 was used. HCR amplification was performed at 28.5nM of each amplifier hairpin. Alexa 488 amplifiers bind to the 18S rRNA probe, Alexa 546 amplifiers bind to the ssRNA ladder probe, Alexa 647 amplifiers bind to the GAPDH probes. For a list of probes and HCR amplifier systems see Tables 5.2 and 5.3.

5.3.3 miRNA blots

Our next goal was to use HCR to detect small RNAs in the range of 19–60 nucleotides (such as siRNAs, miRNAs, piwi-interacting RNAs (piRNAs) [28–30] and our own scRNAs). The detection of small RNAs is more challenging than that of longer mRNAs. Their abundance is typically low and combined with the loss of small RNAs during the extraction process, the blotting procedure requires increased sensitivity. Additionally, due to their short sequence, only one probe can be used per target which may reduce the amount of labels per target. Due to these challenges, a sensitive multiplexed assay using HCR is advantageous.

5.3.3.1 Blotting conditions for miRNAs

To examine miRNA detection using HCR, we chose hsa-miR-16a as the miRNA target, U6 small nuclear 1 (RNU6-1) as an internal control, and a microRNA marker (NEB) as a size marker. The U6 gene is 106 nucleotides long and is often used as a small-RNA loading control. Probes for the miRNA and the ladder are the full complement to the target. For the ladder, the complement to the 17 nucleotides that are present in all three bands was used as the probe. Due to initial difficulties in detecting microRNAs (data not shown), two additional controls were used: a synthetic miR16a target was used to check that the probe and blotting conditions can bind the miR-16a target, and in addition, the synthetic target was spiked into the sample prior to RNA extraction to check for the ability to recover miRNAs.

Two blotting conditions were used, one at room temperature as suggested by the *mirVana* protocol (for a detailed protocol see Materials and methods section; Figure 5.7(a)) and one at 55°C probe hybridization and 45°C HCR amplification as observed suitable for *in situ* [26] and mRNA blotting (Figure 5.7(b)). Figure 5.7 demonstrates that blotting at room temperature leads to non-specific HCR amplification as is evident both for the U6 probe and the miR-16a probe, while blotting in elevated temperatures results in specific detection (Figure 5.7(a) vs. Figure 5.7(b), blue and red channels). The miR-16a probe is

able to amplify its target. However, a signal was not observed for total RNA or samples enriched for small RNAs. The low signal obtained from the spiked sample suggests that the extraction procedure may result in partial loss of miRNAs or that the cross-linking process is not sufficient. Nevertheless, the possibility that miR-16a is not expressed, or expressed at non-detectable levels cannot be ruled out.

Real-time PCR was used to verify the expression level of miR-16a in HEK293A cells. The expression of miR-16a is estimated to be 4.5 ± 0.08 ¹-fold lower than that of U6. Once it was verified that miR-16a is present in the HEK293 cell line, we used a FAM-labeled LNA probe (Exiqon) against miR-16a to examine the northern blot procedure. LNA probes are commonly used as probes in northern blot detection and are considered to have increased sensitivity and specificity compared to DNA and RNA probes [9, 31]. The labeled LNA probe and a synthetic target were used as controls in the blot. Blotting conditions according to the *mirVana* protocol did not yield any signal, including for the control groups (data not shown). Given that using ULTRAhyb for the mRNA blots significantly improved the blotting conditions, we tried blotting with ULTRAhyb-Oligo (Ambion), which is designed for short probes. Figure 5.8(a) demonstrates that while the blotting conditions are sufficient for detection of the synthetic target and probe controls, they are not suitable for detection of endogenous levels of miR-16a.

Next, we examined the cross-linking conditions. Both short-wave UV light and baking cross-link RNA to the membrane through interactions with the RNA bases. Cross-linking of bases within the RNA target, especially for short targets, reduces their availability to interact with the probe. 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) cross-linking is thought to work via the 5' terminal phosphate of the small RNA and results in an immobilized RNA which has more free bases to interact with a probe. The use of EDC has improved the detection of small RNAs by a minimum of 20-fold [32]. By altering the cross-linking method to EDC, miR-16a was detected in total RNA, total RNA spiked with

¹For calculation of the error see Materials and methods section.

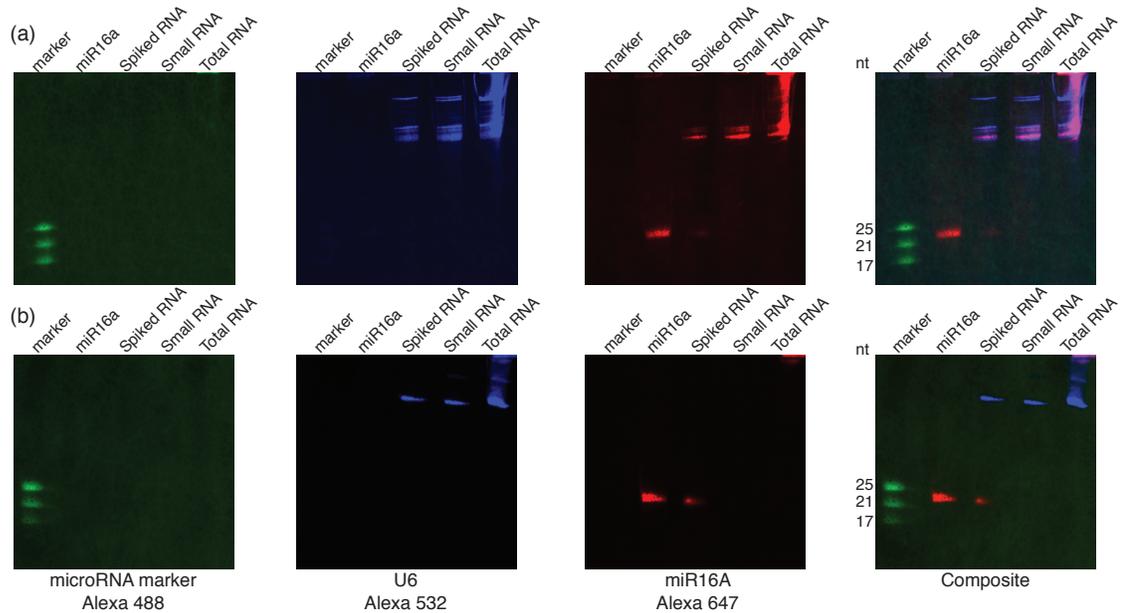


Figure 5.7: Multiplexed detection of synthetic miR-16a target, endogenous U6 RNA and microRNA ladder via HCR. Extraction of small RNAs was done from HEK293 d2EGFP cells according to the *mirVana* enrichment procedure for small RNAs (Ambion). For the sample spiked with a synthetic target, the target was added prior to extraction. 60ng microRNA marker, 0.5pmol synthetic miR-16a target, 720ng of small RNA extraction with a spiked synthetic marker, 975ng of small RNA extraction and 22.5 μ g total RNA extracted with ZR RNA MiniPrep (Zymo Research) were separated on a 15% denaturing polyacrylamide gel. Blotting was carried out according to the *mirVana* miRNA blotting procedure (see Materials and methods section). Probes were used at a 5nM concentration each. Amplification was done at 10nM. (a) Hybridization, wash and HCR amplification at room temperature. (b) Hybridization and wash at 55°C, HCR amplification and wash at 45°C. Probes: microRNA ladder A2 initiator, miR-16a A1 initiator, U6 A5 initiator. Amplifiers: A1 Alexa 647, A2 Alexa 488, A5 Alexa 546. For a list of probes and HCR amplifier systems see Tables 5.2 and 5.3.

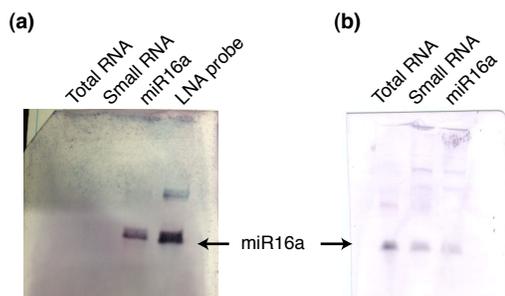


Figure 5.8: **Detection of endogenous miR-16a target using a FAM-labeled LNA probe.** Extraction of total RNA or small RNAs was done from HEK293A or HEK293 d2EGFP cells according to the relevant *mirVana* protocol (Ambion). Detection of the probe was done using 1.5mg/ml alkaline phosphatase anti-fluorescein antibody (Vector Laboratories) followed by BCIP/NBT chromogenic detection. (a) 15 μ g total RNA, 1.6 μ g small RNAs, 0.5pmol synthetic miR-16a and 5pmol FAM-labeled LNA probe were blotted onto a positively charged nylon membrane. Cross-linking was done by baking for 30 minutes at 80°C. Probe concentration for hybridization was 1.25nM. (b) 34 μ g total RNA, 18.5 μ g total RNA spiked with 0.1pmol synthetic miR-16a (after extraction), 0.1pmol synthetic miR-16a were blotted onto a positively charged nylon membrane. Cross-linking was done using EDC. Probe concentration for hybridization was 3.3nM.

a synthetic miR-16a target and a synthetic miR-16a target using a FAM-labeled LNA probe (Figure 5.8(b)). Taken together, these results suggest that the use of ULTRAhyb-Oligo in combination with EDC cross-linking are good conditions for detection of endogenous miRNAs.

5.3.3.2 Multiple HCR-initiator probes

The low abundance of miRNAs makes them a difficult target to detect. While the signal of mRNAs can be amplified by the use of multiple probes per target (thus growing multiple HCR polymers per target), the short length of small RNA targets allows use of only one probe per target. In order to try and increase the signal generated per probe, multiple initiators can be attached to an HCR probe, allowing one probe to facilitate the polymerization of multiple polymers. Figure 5.9 depicts this strategy using double-initiator (DI)

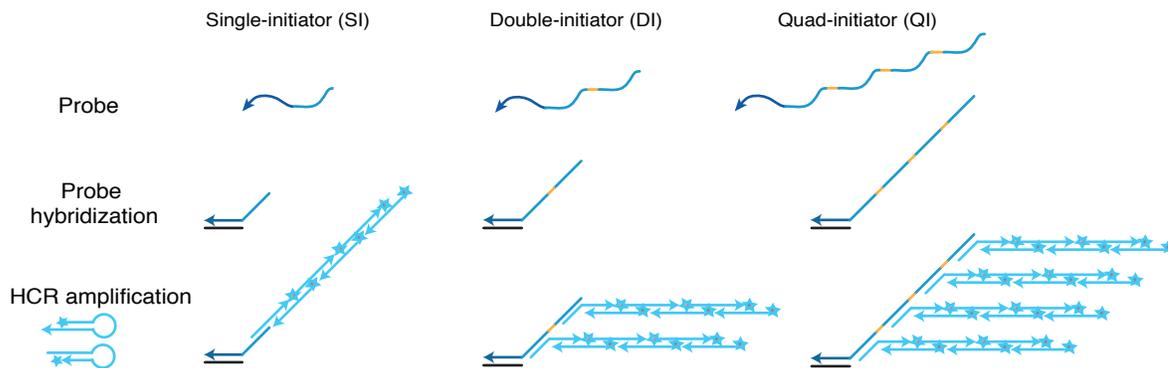


Figure 5.9: **HCR-based detection using probes with multiple initiators schematic.** Target (black), probe region complementary to the target (dark blue), initiator sequence (turquoise), spacer sequence between initiators (orange), amplifier hairpins (light blue) and detection moiety (star).

and quad-initiator (QI) probes. Figure 5.10 demonstrates that increasing the number of initiators increases the observed signal as expected. It is yet to be determined whether this relationship is linear. Compared to the signal of a single initiator (SI probe), a DI probe increased the signal by 1.84 ± 0.56 and a QI probe increased the signal by 3.07 ± 0.98 . Signal increase represents the average increase in signal across three independent experiments; errors represent the standard error of the mean. Amplification was done at 60nM for each amplifier hairpin for SI, DI and QI probes. To verify that this regime is in excess, an additional sample with 120nM amplifier hairpins was included with a QI probe. The signal did not significantly change by doubling the amount of amplifier suggesting that 60nM is a sufficient concentration.

5.3.3.3 Multiplexed miRNA detection using HCR

By using EDC cross-linking in combination with ULTRAhyb-Oligo for probe hybridization, a DI probe and HCR amplification, it is possible to detect endogenous levels of miR-16a in 293A total RNA (Figure 5.11). We next examined whether it is possible to detect two

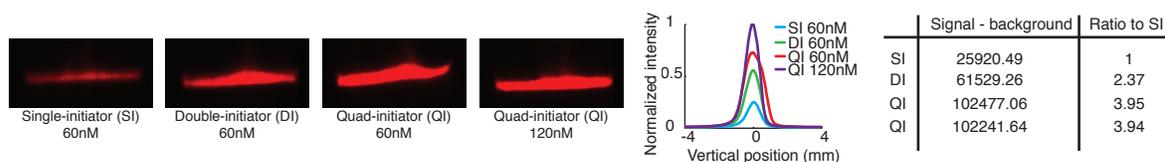


Figure 5.10: **HCR-based detection using probes with single, double and quad initiators.** One pmol synthetic miR-16a target were blotted using 20nM of single-, double-, or quad-initiator probes for HCR system A1. Amplification was carried out at 60nM or 120nM (indicated in the figure) with Alexa 546-labeled amplifiers. The histogram depicts the normalized intensity of each band. For a detailed method see miRNA blotting procedure with ULTRAhyb-Oligo in the Materials and methods section. For a list of probes and HCR amplifier systems see Tables 5.2 and 5.3.

different microRNAs simultaneously. The challenge in this case is that both miRNAs would be present at the same location on the membrane. To this end, we have chosen miR-21 as the second miRNA target. Based on data obtained in small RNA library sequencing, we expect this microRNA to be expressed at very low levels in HEK cells and expressed at similar levels to miR-16a in the glioblastoma cell line U87MG [33]. The use of RNA extracted from both cell lines can be used as a positive and negative control on a blot.

Figure 5.12 demonstrates that a microRNA ladder, U6 RNA, miR-16a and miR-21 can all be detected simultaneously in a northern blot procedure; these targets could be detected from as little as 1 μ g total RNA (Figure 5.12(c) and (d)). In order to detect all four targets, the microRNA ladder, as well as one probe, had to use the same label due to technical reasons related to lab equipment. The amplifier systems were different and only the labeling (Alexa dye) is identical. Detection of two microRNAs is not limited to small (endogenous) amounts of target; 0.5pmol (each) of miR-16a and miR-21 can be detected using multiplexed blotting (Figure 5.13).

Based on small-RNA library sequencing data [33], we expected miR-16a and miR-21 to be expressed at similar levels in U87MG cells. However, detection of miR-16a was surprisingly low (Figure 5.12). This is not likely due to the blotting procedure, as is evident from Figure 5.13. It might be due to a variation in the glioblastoma cell line used

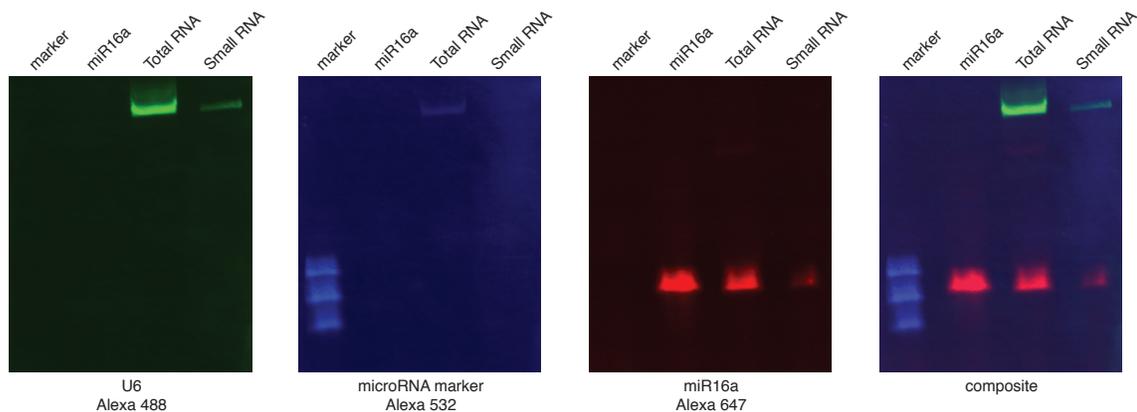


Figure 5.11: **Multiplexed HCR-based detection of endogenous miR-16a.** 12ng microRNA marker, 0.1pmol synthetic miR-16a, 25.8 μ g total RNA and 364ng extracted RNA enriched for small RNAs were used for blotting. Hybridize with miR-16a-DI probe, U6 short SI probe, microRNA marker probe DI at 5.7nM. After hybridization and HCR amplification the membrane was washed twice in 2 \times SSC 0.1% SDS at room temperature for 5 minutes followed by a 5-minute wash in 0.1 \times SSC 0.1% SDS at 37 $^{\circ}$ C. For a list of probes and HCR amplifier systems see Tables 5.2 and 5.3.

or due to bias in small RNA library sequencing [34, 35]. The expression of each microRNA relative to U6 was examined using qRT-PCR in both cell lines. Preliminary data show that expression of miR-16a in U87MG cells is indeed lower than expression of miR-21 and in the HEK293A cell line, miR-21 is expressed at lower levels than miR-16a (data not shown).

5.3.3.4 Detection limit of HCR-based miRNA northern blot detection

Kim et al. [10] have recently reported that by using digoxigenin-labeled LNA probes in combination with EDC cross-linking, detection of 0.05 femtomol miRNA was achieved; this is comparable to radioactive labeling. We next examined whether HCR detection can compare in sensitivity. Using serial dilutions of synthetic miR-16a, HCR detection with a QI probe using Alexa 647-labeled amplifiers has a sensitivity limit of 0.1 femtomol (Figure 5.14(a)). Radioactive labeling with 32 P is most commonly used due to its high sensitivity. Whereas using a DI probe with 32 P labeled HCR amplifiers maintains a sensitivity of 0.1

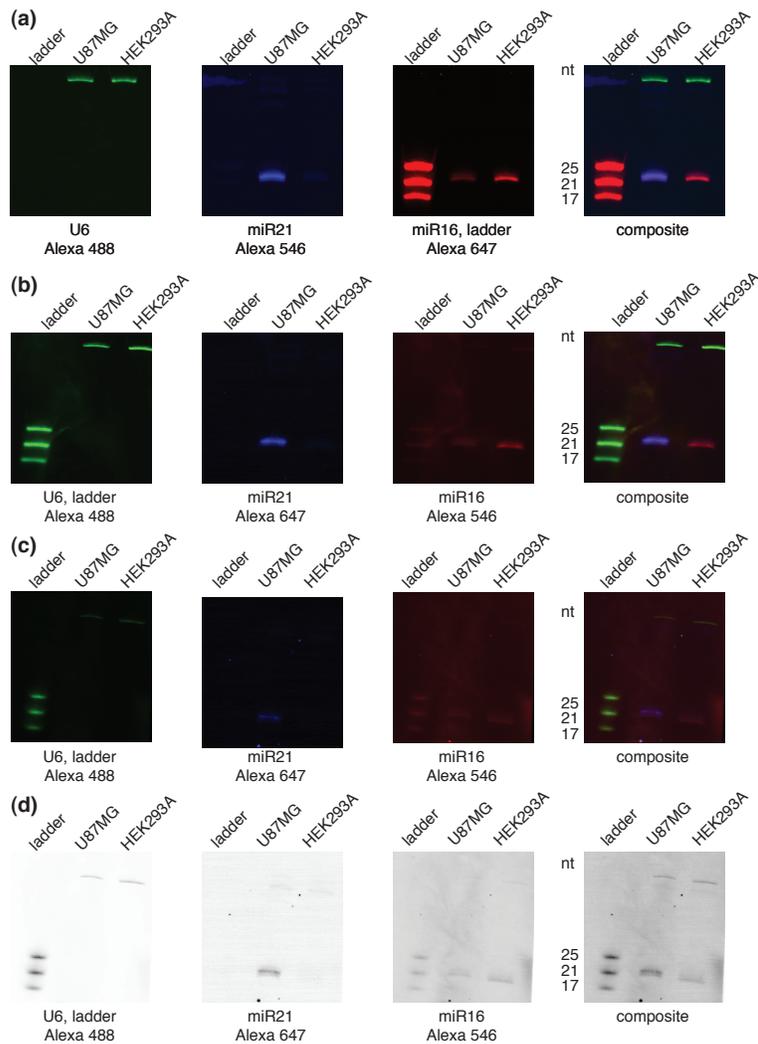


Figure 5.12: **Multiplexed HCR-based detection of endogenous miR-16a, miR-21 and U6.** All blots were performed using 20nM probes and 60nM of each amplifier hairpin. For detailed methods see miRNA blotting procedure with ULTRAhyb-Oligo in the Materials and methods section. Probes: ladder DI A2, U6_{short} SI A5, miR-16a QI A1, miR-21 QI A3. (a) 5µg total RNA, 24ng microRNA marker. Amplifiers: A1 Alexa 647, A2 Alexa 647, A3 Alexa 546, A5 Alexa 488. (b) 3µg total RNA, 18ng microRNA marker. Amplifiers: A1 Alexa 546, A2 Alexa 488, A3 Alexa 647, A5 Alexa 488. (c) 1µg total RNA, 18ng microRNA marker. Slightly less than 20nM probe for the microRNA ladder were used. Amplifiers: same as panel C. (d) Greyscale representation of panel (c). For a list of probes and HCR amplifier systems see Tables 5.2 and 5.3.

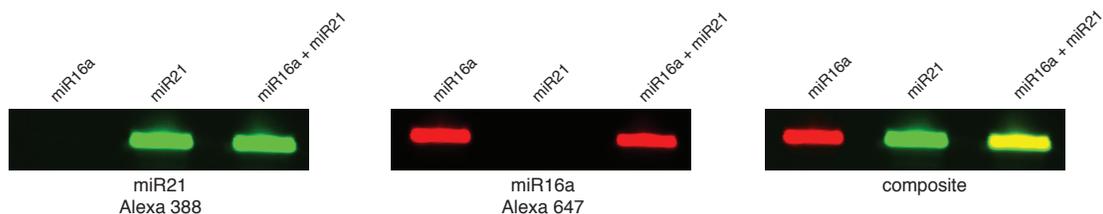


Figure 5.13: **Multiplexed HCR-based detection of miR-16a and miR-21.** 0.5pmol of synthetic microRNA targets were blotted using 20nM probes and 50nM of each amplifier hairpin, for a detailed method see miRNA blotting procedure with ULTRAhyb-Oligo in the Materials and methods section. Probes: miR-16a DI A1, miR-21 DI A3. Amplifiers: A1 Alexa 647, A3 Alexa 488. For a list of probes and HCR amplifier systems see Tables 5.2 and 5.3.



Figure 5.14: **Sensitivity of HCR-based detection.** Serial dilutions of 5' phosphorylated synthetic miR-16a targets blotted using ULTRAhyb-Oligo (see Materials and methods section). (a) 20nM QI A1 probe, 60nM A1 amplifiers labeled with Alexa 647. (b) 20nM DI A1 probe, 0.5×10^6 cpm/ml of each ^{32}P amplifier hairpin (10^6 cpm/ml total). (c) 13.5nM QI A1 probe, 0.5×10^6 cpm/ml of each ^{32}P amplifier hairpin. For a list of probes and HCR amplifier systems see Tables 5.2 and 5.3.

femtomol, by using radioactive HCR with a QI, probe the detection limit improves to 0.01 femtomol, a five-fold sensitivity increase compared to that reported in the literature (Figure 5.14(b) and (c)). The QI sensitivity limit is maintained when probe concentration is increased from 13.5nM to 20nM (data not shown).

5.3.3.5 Estimation of polymer length

Finally, we wanted to know how many hairpins (fluorophores) are present in each HCR polymer. To estimate the number of hairpins present, the signal obtained from one probe labeled with an Alexa dye was compared to the signal obtained from an HCR polymer

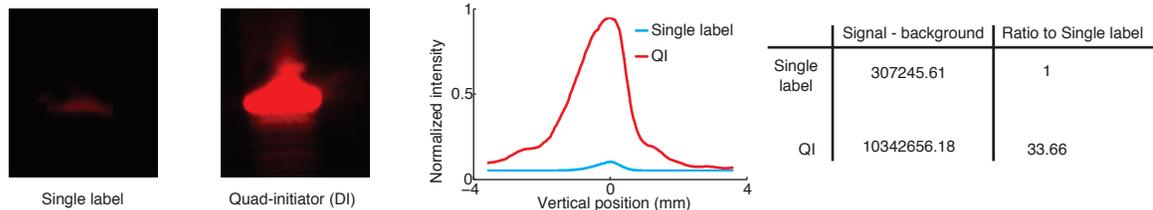


Figure 5.15: **Estimation of HCR polymer length.** 2pmol 5' phosphorylated synthetic miR-16a target were blotted using 20nM miR-16a antisense probe labeled with Alexa 647 or with a QI A1 HCR probe. HCR detection was carried out using 60nM A1 Alexa 647 labeled amplifiers. The blot of the single labeled probe was blot underwent the exact same procedure as the HCR blot without the addition of amplifiers. The histogram depicts the normalized intensity of each band produced. For a list of probes and HCR amplifier systems see Tables 5.2 and 5.3.

initiated by a QI probe. The signal obtained for the QI probe was roughly 33 times higher than that of a probe with one label. The signal obtained for a QI probe should be composed of four HCR polymers, estimating a polymer length of about eight hairpins per polymer (Figure 5.15). This estimate is in agreement with estimation of the polymer length of an SI probe against d2EGFP (data not shown).

5.4 Discussion

We have presented the use of a hybridization chain reaction as a method for northern blot detection. We have shown that HCR is suitable to detect both long targets such as mRNAs as well as short targets such as microRNAs. In HCR detection, the target probe contains an additional “tail” which serves as an HCR initiator. During the detection step, a polymer is grown off of each bound probe. Each amplifier hairpin is labeled with a detection moiety (fluorescent dye or radio-label) for signal readout. The HCR polymers are orthogonal using a different HCR initiator on each probe each target has a unique polymer attached to it.

The main advantage of HCR over current detection methods (e.g ^{32}P , biotin etc.) is that it allows for parallel multiplexing. By labeling each HCR hairpin (and therefore

polymer) with a specific fluorophore, each polymer is visualized distinctly and specifically with its own target. Thus, multiple targets can be probed and detected simultaneously on the same blot. Not only does this significantly save time, but it also reduces the risk and complications involved with stripping and re-probing blots. Many possibilities exist for the generation of orthogonal HCR polymers. The limitation in the number of targets that can be detected simultaneously is technical and depends on the spectral channels of the imaging device.

A limitation of fluorescent HCR is that fluorescence is typically less sensitive than radioactive labeling [21] with ^{32}P . In this work, we show a detection limit of about 0.1 femtomol miR-16a using a quad-initiator RNA HCR probe, whereas the reported literature suggest a detection limit of 0.05 femtomol using a ^{32}P or digoxigenin-labeled LNA probe [10]. However, if HCR labeling is switched to ^{32}P , then the detection limit of HCR becomes 0.01 femtomol, a 5-fold improvement relative to the literature.

Further improvements to sensitivity can also be made by changing the nucleotide composition of the probe to 2'-OMe or LNA. The main drawback is that these may not allow for multiple initiators; due to synthesis technicalities they cannot be transcribed, and are more expensive.

Alternatively, improvements to sensitivity can be made by increasing polymer length or increasing the number of polymers per target. Our work suggests that each polymer is composed of approximately eight hairpins, while gel studies for the same polymers suggest a mean length of 20 hairpins [26]. Further optimization of hybridization conditions may be needed to grow longer polymers. For long RNA targets, multiple probes can be used to grow multiple polymers per target and thus improve sensitivity. This strategy is not possible with short miRNA targets. To overcome this, we have designed probes with multiple initiators. This way, one probe can have multiple HCR-polymers attached to it therefore leading to an increase in signal. Indeed, we have shown that multiple-initiator probes lead to a stronger signal when compared to a single-initiator probe. A further

increase in signal may be obtained by optimizing the spacer length between the initiator sequences.

Other strategies based on nucleic acids structures have been used to obtain signal amplification. Mainly, the use of DNA dendrimers and branched DNA (bDNA) assays have been used as the amplification method [36–42]. These assays were mostly focused on signal amplification for ISH, microarrays, miRNA detection in microtiter plates and plasma. Although these strategies can potentially be used for signal amplification in blotting methods, to the best of our knowledge, this is not common practice.

The work presented here is not limited to northern blot detection of multiple targets. It may also be used to detect the same target mRNA redundantly in two channels using different HCR systems. This can assist in validating the specificity of probes. More broadly, HCR can be used as a detection method for many other hybridization-based techniques such as Southern blot, array formats and possibly for point-of-care detection. This work provides a stepping stone toward achieving sensitive, multiplexed detection in diverse *in vitro* settings using HCR.

5.5 Materials and methods

***in vitro* transcription.** pCS2plus-EGFP plasmid was a generous gift from Dr. Le Trinh. The plasmid was linearized with NotI prior to SP6 transcription. Transcription was carried out using AmpliScribe SP6 high yield transcription kit (Epicenter), according to the manufacturer. The transcribed RNA was purified using Qiagen's RNeasy mini kit according to the manufacturer, on-column DNaseI digestion was performed. RNA was quantified using A₂₆₀ absorbance on a NanoDrop8000 (Thermo Scientific).

Snap cool. Amplifier hairpins were snap cooled by heating them to 95°C for 90 seconds followed by a ≥ 30 minute incubation at room temperature in the dark.

Cell lines. HEK293A cells were purchased from Invitrogen (catalog #R705-07), HEK293 d2EGFP cells were a generous gift from Dr. Chase Beisel. The destabilized EGFP sequence comes from pd2EGFP-1 plasmid (Clontech, PT3205-5 catalog #6008-1). Cells were maintained at 37°C 5% CO₂ in DMEM (Invitrogen) supplemented with 10% fetal bovine serum (Invitrogen).

RNA extraction. Total RNA was extracted either with Trizol reagent (Ambion), ZR RNA MiniPrep (Zymo Research) or *mirVana* miRNA isolation kit (Ambion).

mRNA blotting procedure. Denaturing formaldehyde/MOPS agarose gel was prepared with 1×NorthernMax denaturing gel buffer (Ambion) according to the manufacturer. Prior to loading, total RNA was mixed 1:1 (vol/vol) with formamide and heated to 65°C for 15 minutes. The gel was run in 1×MOPS buffer (Ambion) at 55Volts for 2–3 hours. The gel was washed four times in water prior to transfer. A positively charged nylon membrane (BrightStar-Plus (Ambion) or Nytran-SPC (Whatman)) was pre-wet in water and then equilibrated 5 minutes in 20×SSC. RNA transfer to the membrane by “downward transfer” in 20×SSC for 3–4 hours. The RNA was cross-linked to the membrane by baking at 80°C for two hours between two 3MM Whatman filter sheets. The membrane was pre-hybridized at 65°C for at least 30 minutes in ULTRAhyb (Ambion). Probes were added to the pre-hybridization solution and hybridization was carried out overnight at 55°C. The blots were

washed twice in $2\times\text{SSC}$ 0.1% SDS for 5 minutes at room temperature followed by a 5-minute $0.1\times\text{SSC}$ 0.1%SDS wash at 55°C unless otherwise specified. The membrane was then pre-hybridized again in ULTRAhyb followed by addition of snap cooled HCR amplifier hairpins. Amplification was carried out overnight at 45°C in the dark. Prior to imaging, the blots were washed twice in $2\times\text{SSC}$ 0.1% SDS for 5 minutes at room temperature followed by a 5-minute $0.1\times\text{SSC}$ 0.1%SDS wash at 45°C unless otherwise specified.

mirVana miRNA blotting procedure. Pre-hybridization was carried out for at least an hour at 65°C in $6\times\text{SSC}$, $10\times\text{Denhardt's}$ solution, 0.2%SDS. Probes were hybridized to the membrane overnight at room temperature in $6\times\text{SSC}$, $5\times\text{Denhardt's}$ solution, 0.2% SDS unless otherwise specified. The membrane was washed at room temperature three times in $6\times\text{SSC}$, 0.2% SDS prior to HCR amplification unless otherwise specified. Amplification was carried out at room temperature overnight in the dark unless otherwise specified. After amplification the membrane was washed as indicated above.

miRNA blotting procedure with ULTRAhyb-Oligo. Samples were run on a 15% denaturing polyacrylamide gel at 300 volts for 25–30 minutes. When total RNA was used, the gel was pre-run prior to loading the RNA. For electrophoresis, the samples were mixed 1:1(vol/vol) with formamide and heated to 95°C for 5 minutes or to 65°C for 15 minutes. A positively charged nylon membrane (Roche) or NytranSPC (Whatman) was pre-wet in water and then equilibrated 5 minutes in $0.5\times\text{TBE}$ prior to transfer. Semi-dry transfer in $0.5\times\text{TBE}$ was used at 0.8mA to 2mA per square centimeter of gel for 45 minutes to 2 hours using a Panther semidry electroblotter (Owl separation systems). Unless otherwise specified, the RNA was cross-linked to the membrane using EDC (see below). Pre-hybridization, hybridization and HCR amplification were carried out in ULTRAhyb-Oligo (Ambion) at 37°C unless otherwise specified. The membrane was washed 2–3 times in $2\times\text{SSC}$ 0.5% SDS or $2\times\text{SSC}$ 0.1% SDS for 5–15 minutes after probe hybridization and after amplification.

BCIP/NBT chromogenic detection. Detection was done using the protocol and solu-

tions in the Biotin Chromogenic Detection kit (Thermo Scientific). Alternatively, blocking was done in 1X casein solution (Vector Laboratories) and washed in 1×PBS 1% Tween20.

EDC cross-linking. 245 μ l of 12.5M 1-methylimidazole (Alfa Aesar) were added to 9ml of RNase-free water. pH was adjusted to 8 using 1M HCl. 0.753gr of 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC; Sigma or Bio-Rad) were added to the solution and the volume was adjusted to 24ml with water. The EDC solution was used to saturate a 3MM Whatman chromatography paper. The membrane was placed on the saturated sheet with the RNA side facing up and both the membrane and the Whatman paper were wrapped in Saran. Cross-linking was done by a 1–2 hour incubation at 60°C. Following cross-linking, the membrane was washed with water to remove excess cross-linking solution.

Imaging. Membranes were imaged on an FLA-5100 imaging system (Fuji Photo Film). Radioactive membranes were exposed onto an image plate (Fujifilm type BAS-MS) and scanned using the IP-S mode at 600V. Fluorescent membranes were scanned using the following settings:

Dye	Excitation	Filter
Alexa 488	473 nm	BP 530 \pm 10nm
Alexa 532	532 nm	BP 570 \pm 10nm
Alexa 647	635 nm	LP 665 nm

Table 5.1: Excitation lasers and emission filters used.

Quantification. Multi Gauge ver3.0 (Fujifilm) software was used to quantify gels using the quantitative analysis with a profile feature.

5' end labeling. HCR amplifier hairpins were 5' end labeled with [γ -³²P] ATP (10mCi/ml, MP Biomedicals) using T4 polynucleotide kinase (New England Biolabs). Unincorporated [γ -³²P] ATP were removed by spin column chromatography using Illustra MicroSpin G-25 columns (GE Healthcare) according to the manufacturer. Counts were measured on a Beckman LS-5000TD Liquid Scintillation Counter.

qPCR analysis. RNA for qPCR was extracted using TaqMan MicroRNA Cells-to-Ct kit (Ambion) according to the manufacturer. cDNA was synthesized using TaqMan microRNA

reverse transcription kit (Ambion). Real-time PCR reaction was prepared using TaqMan Universal PCR Master Mix, No AmpErase UNG and carried out in a CFX96 (Bio-Rad). TaqMan primer pairs used were purchased from Applied Biosystems: hsa-miR-16 (catalog #000391), hsa-miR-21 (catalog #000397) and U6 snRNA (catalog #001973).

Expression of microRNA relative to U6 was done by simplifying the Pfaffl method [43] to include only one sample while correcting for qPCR efficiency of each reaction using the equation,

$$\text{Ratio} = \frac{U6}{\text{microRNA}} = \frac{E_{\text{microRNA}}^{C_t(\text{microRNA})}}{E_{U6}^{C_t(U6)}}. \quad (5.1)$$

Three biological replicates were included and three technical replicates per biological replicate (9 data points total).

The efficiency of the reaction was calculated using the following equation $E=10^{\frac{-1}{\text{slope}}}$ where the *slope* was calculated based on a serial 10-fold dilution of a cDNA sample. Reactions with *Ct* values higher than 36 were not used for the analysis. Each dilution series was analyzed in triplicate.

To account for the uncertainty in the *Ct* values on the Ratio (equation (5.1)), we use the simple propagation of errors. The uncertainty, σ_f of a quantity $f(x, y)$, that is a function of variables x and y with known uncertainties σ_x and σ_y , is given by,

$$\sigma_f = \sqrt{\left(\frac{\partial f}{\partial x}\sigma_x\right)^2 + \left(\frac{\partial f}{\partial y}\sigma_y\right)^2}. \quad (5.2)$$

Thus, the uncertainty in the Ratio (equation (5.1)), given the uncertainties $\sigma_{C_t(\text{microRNA})}$ and $\sigma_{C_t(U6)}$ is given by,

$$\sigma_{\text{Ratio}} = \text{Ratio} \sqrt{\left(\log(E_{\text{microRNA}})\sigma_{C_t(\text{microRNA})}\right)^2 + \left(\log(E_{U6})\sigma_{C_t(U6)}\right)^2}. \quad (5.3)$$

Probe	HCR system	Sequence
-------	------------	----------

EGFP SI	A1	CCGAAUACAAAGCAUCAACGACUAGA AAAAAGUUCUUCUGCUUGUCGGC CAUGAUAUAGACGUUGUGGCUG UUGUAGUUGU
EGFP SI	A3	GACUACUGAUAACUGGAUUGCCUAG AAUUUGUUCUUCUGCUUGUCGGC CAUGAUAUAGACGUUGUGGCUG UUGUAGUUGU
GAPDH 1 SI	A1	CCGAAUACAAAGCAUCAACGACUAGA AAAAAAAAAGAAGAUGCGGCUGA CUGUCGAACAGGAGGAGCAGAG AGCGAAGCGG
GAPDH 2 SI	A1	CCGAAUACAAAGCAUCAACGACUAGA AAAAAUCCGUUGACUCCGACCUU CACCUUCCCAUGGUGUCUGAG CGAUGUGGCU
GAPDH 3 SI	A1	CCGAAUACAAAGCAUCAACGACUAGA AAAAACCCGUUCUCAGCCUUGAC GGUGCCAUGGAAUUUGCCAUGG GUGGAAUCAU
GAPDH 4 SI	A1	CCGAAUACAAAGCAUCAACGACUAGA AAAAAUUCCACGAUACCAAAGUU GUCAUGGAUGACCUUGGCCAGG GGUGCUAAGC
GAPDH 5 SI	A1	CCGAAUACAAAGCAUCAACGACUAGA AAAAAUCGCUGUUGAAGUCAGAG GAGACCACCUUGGUGCUCAGUGU

		AGCCCAGGAU
18S rRNA SI	A6	CCACAUACCAUCAGACCAGACUAGAC AAAUACGGAACUACGACGGUAUC UG
Low range ssRNA ladder SI	A2	GACCCUAAGCAUACAUCGUCCUUCAU UUUUUCUCGACGAAGACUCCC
ssRNA ladder SI	A2	GACCCUAAGCAUACAUCGUCCUUCA UUUUUUUUUUUUUCCAAGACAU CUUCCAGUCGCUGGCGCUUGG GGUACCAUCAGCU
miRNA ladder SI	A2	GACCCUAAGCAUACAUCGUCCUUCAU UUUUUAUCUCAACCAGCCACUG
miRNA ladder DI	A2	GACCCUAAGCAUACAUCGUCCUUCAU UUUUUGACCCUAAGCAUACAUCGUCC UUCAUUUUUUUAUCUCAACCAGCC ACUG
hsa-miR-16a SI	A1	CCGAAUACAAAGCAUCAACGACUAGA AAAAACGCCAAUAUUUACGUGCU GCUA
hsa-miR-16a DI	A1	CCGAAUACAAAGCAUCAACGACUAGA AAAAACCGAAUACAAAGCAUCAACGA CUAGAAAAACGCCAAUAUUUACG UGCUGCUA
hsa-miR-16a QI	A1	CCGAAUACAAAGCAUCAACGACUAGA AAAAACCGAAUACAAAGCAUCAACGA CUAGAAAAACCGAAUACAAAGCAUC AACGACUAGAAAAACCGAAUACAAA

hsa-miR-21 QI	A3	<p>GCAUCAACGACUAGAAAAACGCCA AUAUUUACGUGCUGCUA GACUACUGAUAACUGGAUUGCCUUAG <i>AAUUUGACUACUGAUAACUGGAUUGC</i> CUUAGAAUUUGACUACUGAUAACUGG AUUGCCUUAGAAUUUGACUACUGAU ACUGGAUUGCCUUAGAAUUUUCAAC AUCAGUCUGAUAAGCUA</p>
U6 SI	A5	<p>UACGCCUAAGAAUCCGAACCCUAUG <i>AAAUACGUUCCAAUUUUAGUAUA</i> UGUGCUGCCGAAGCGA</p>
U6 short SI	A5	<p>UACGCCUAAGAAUCCGAACCCUAUG <i>AAAU</i>AUGUGCUGCCGAAGCGA</p>

Table 5.2: List of probes used. In bold is the target binding sequence, italicized is the 5nt spacer, the rest is the initiator(s) sequence. Sequences are listed 5' to 3'

HCR system	Hairpin	Sequence
A1	H1	UCUAGUCGUUGAUGCUUUGUAUUCGGCGACAGAUAAC CGAAUACAAAGCAUC /C9-dye-3'/
A1	H2	/5'-dye-C12/ CCGAAUACAAAGCAUCAACGACUAGAGAU GCUUUGUAUUCGGUUAUCUGUCG
A2	H1	AUGAAGGACGAUGUAUGCUUAGGGUCGACUCCAUAG ACCCUAAGCAUACAU /C9-dye-3'/
A2	H2	/5'-dye-C12/ GACCCUAAGCAUACAUCGUCCUUCAUAUG UAUGCUUAGGGUCUAUGGAAGUC
A3	H1	CUAAGGCAAUCCAGUUAUCAGUAGUCUGACACGACUG ACUACUGAUAAACUGG /C9-dye-3'/
A3	H2	/5'-dye-C12/ GACUACUGAUAAACUGGAUUGCCUAGCCA GUUAUCAGUAGUCAGUCGUGUCA
A5	H1	CAUAGGGUUCGGAUUCUAGGGCGUAGCAGCAUCAAU ACGCCCUAAGAAUCC /C9-dye-3'/
A5	H2	/5'-dye-C12/ UACGCCCUAAGAAUCCGAACCCUAUGGGA UUCUAGGGCGUAUUGAUGCUGC
A6	H1	GUCUAGUCUGGUCUGAUGGUAUGUGGACAAUCCUAGC CACAUACCAUCAGAC /C9-dye-3'/
A6	H2	/5'-dye-C12/ CCACAUACCAUCAGACCAGACUAGACGUCU GAUGGUAUGUGGCUAGGAUUGU

Table 5.3: List of HCR amplifiers used. Sequences are listed 5' to 3' .

Bibliography

- [1] Southern, E. M. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of molecular biology* **98**(3), 503–17 (1975).
- [2] Streit, S., Michalski, C. W., Erkan, M., Kleeff, J., and Friess, H. Northern blot analysis for detection and quantification of RNA in pancreatic cancer cells and tissues. *Nature protocols* **4**(1), 37–43 (2009).
- [3] Bohm-Hofstatter, H., Tschernutter, M., and Kunert, R. Comparison of hybridization methods and real-time PCR: their value in animal cell line characterization. *Applied microbiology and biotechnology* **87**(2), 419–25 (2010).
- [4] Reue, K. mRNA quantitation techniques: considerations for experimental design and application. *The Journal of nutrition* **128**(11), 2038–44 (1998).
- [5] Cohen, R. D., Castellani, L. W., Qiao, J. H., Van Lenten, B. J., Lusis, A. J., and Reue, K. Reduced aortic lesions and elevated high density lipoprotein levels in transgenic mice overexpressing mouse apolipoprotein A-IV. *The Journal of clinical investigation* **99**(8), 1906–16 (1997).
- [6] Trayhurn, P., Duncan, J. S., Nestor, A., Thomas, M. E., and Rayner, D. V. Chemiluminescent detection of mRNAs on northern blots with digoxigenin end-labeled oligonucleotides. *Analytical biochemistry* **222**(1), 224–30 (1994).
- [7] Augood, S. J., Ruth, J. L., and Emson, P. C. A rapid method of non radioactive northern blot analysis. *Nucleic acids research* **18**(14), 4291 (1990).
- [8] Trayhurn, P. Northern blotting. *The Proceedings of the Nutrition Society* **55**(1B), 583–9 (1996).
- [9] Valoczi, A., Hornyik, C., Varga, N., Burgyan, J., Kauppinen, S., and Havelda, Z.

- Sensitive and specific detection of microRNAs by northern blot analysis using LNA-modified oligonucleotide probes. *Nucleic acids research* **32**(22), e175 (2004).
- [10] Kim, S. W., Li, Z., Moore, P. S., Monaghan, A. P., Chang, Y., Nichols, M., and John, B. A sensitive non-radioactive northern blot method to detect small RNAs. *Nucleic acids research* **38**(7), e98 (2010).
- [11] Novogrodsky, A., Tal, M., Traub, A., and Hurwitz, J. The enzymatic phosphorylation of ribonucleic acid and deoxyribonucleic acid. ii. further properties of the 5'-hydroxyl polynucleotide kinase. *The Journal of biological chemistry* **241**(12), 2933–43 (1966).
- [12] Hilario, E. End labeling procedures: an overview. *Molecular biotechnology* **28**(1), 77–80 (2004).
- [13] Feinberg, A. P. and Vogelstein, B. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analytical biochemistry* **132**(1), 6–13 (1983).
- [14] Rigby, P. W., Dieckmann, M., Rhodes, C., and Berg, P. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. *Journal of molecular biology* **113**(1), 237–51 (1977).
- [15] Mathew, C. G. Radiolabeling of DNA by nick translation. *Methods in molecular biology* **2**, 257–61 (1985).
- [16] Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K., and Green, M. R. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic acids research* **12**(18), 7035–56 (1984).
- [17] Theissen, G., Richter, A., and Lukacs, N. Degree of biotinylation in nucleic acids estimated by a gel retardation assay. *Analytical biochemistry* **179**(1), 98–105 (1989).

- [18] Holtke, H. J. and Kessler, C. Non-radioactive labeling of RNA transcripts in vitro with the hapten digoxigenin (DIG); hybridization and ELISA-based detection. *Nucleic acids research* **18**(19), 5843–51 (1990).
- [19] Guesdon, J. L. Immunoenzymatic techniques applied to the specific detection of nucleic acids. a review. *Journal of immunological methods* **150**(1-2), 33–49 (1992).
- [20] Sato, M., Murao, K., Mizobuchi, M., and Takahara, J. Quantitative and sensitive northern blot hybridization using PCR-generated DNA probes labeled with digoxigenin by nick translation. *BioTechniques* **15**(5), 880–2 (1993).
- [21] During, K. Non-radioactive detection methods for nucleic acids separated by electrophoresis. *Journal of chromatography* **618**(1-2), 105–31 (1993).
- [22] Kessler, C. Non-radioactive analysis of biomolecules. *Journal of biotechnology* **35**(2-3), 165–89 (1994).
- [23] Shifman, M. I. and Stein, D. G. A reliable and sensitive method for non-radioactive northern blot analysis of nerve growth factor mRNA from brain tissues. *Journal of neuroscience methods* **59**(2), 205–8 (1995).
- [24] Hoeltke, H. J., Ettl, I., Finken, M., West, S., and Kunz, W. Multiple nucleic acid labeling and rainbow detection. *Analytical biochemistry* **207**(1), 24–31 (1992).
- [25] Dirks, R. M. and Pierce, N. A. Triggered amplification by hybridization chain reaction. *Proceedings of the National Academy of Sciences of the United States of America* **101**(43), 15275–8 (2004).
- [26] Choi, H. M., Chang, J. Y., Trinh le, A., Padilla, J. E., Fraser, S. E., and Pierce, N. A. Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nature biotechnology* **28**(11), 1208–12 (2010).

- [27] Deindl, E. 18s ribosomal RNA detection on northern blot employing a specific oligonucleotide. *BioTechniques* **31**(6), 1250, 1252 (2001).
- [28] Kim, V. N. Small RNAs: classification, biogenesis, and function. *Molecules and cells* **19**(1), 1–15 (2005).
- [29] Zamore, P. D. and Haley, B. Ribo-gnome: the big world of small RNAs. *Science* **309**(5740), 1519–24 (2005).
- [30] Carthew, R. W. Gene regulation by microRNAs. *Current opinion in genetics & development* **16**(2), 203–8 (2006).
- [31] Varallyay, E., Burgyan, J., and Havelda, Z. Detection of microRNAs by northern blot analyses using LNA probes. *Methods* **43**(2), 140–5 (2007).
- [32] Pall, G. S., Codony-Servat, C., Byrne, J., Ritchie, L., and Hamilton, A. Carbodiimide-mediated cross-linking of RNA to nylon membranes improves the detection of siRNA, miRNA and piRNA by northern blot. *Nucleic acids research* **35**(8), e60 (2007).
- [33] Landgraf, P., Rusu, M., Sheridan, R., Sewer, A., Iovino, N., Aravin, A., Pfeffer, S., Rice, A., Kamphorst, A. O., Landthaler, M., Lin, C., Succi, N. D., Hermida, L., Fulci, V., Chiaretti, S., Foa, R., Schliwka, J., Fuchs, U., Novosel, A., Muller, R. U., Schermer, B., Bissels, U., Inman, J., Phan, Q., Chien, M., Weir, D. B., Choksi, R., De Vita, G., Frezzetti, D., Trompeter, H. I., Hornung, V., Teng, G., Hartmann, G., Palkovits, M., Di Lauro, R., Wernet, P., Macino, G., Rogler, C. E., Nagle, J. W., Ju, J., Papavasiliou, F. N., Benzing, T., Lichter, P., Tam, W., Brownstein, M. J., Bosio, A., Borkhardt, A., Russo, J. J., Sander, C., Zavolan, M., and Tuschl, T. A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129**(7), 1401–14 (2007).
- [34] Hafner, M., Renwick, N., Brown, M., Mihailovic, A., Holoch, D., Lin, C., Pena, J. T., Nusbaum, J. D., Morozov, P., Ludwig, J., Ojo, T., Luo, S., Schroth, G., and Tuschl,

- T. RNA-ligase-dependent biases in miRNA representation in deep-sequenced small RNA cDNA libraries. *RNA* **17**(9), 1697–712 (2011).
- [35] Sorefan, K., Pais, H., Hall, A. E., Kozomara, A., Griffiths-Jones, S., Moulton, V., and Dalmay, T. Reducing ligation bias of small RNAs in libraries for next generation sequencing. *Silence* **3**(1), 4 (2012).
- [36] Urdea, M. S., Running, J. A., Horn, T., Clyne, J., Ku, L. L., and Warner, B. D. A novel method for the rapid detection of specific nucleotide sequences in crude biological samples without blotting or radioactivity; application to the analysis of hepatitis b virus in human serum. *Gene* **61**(3), 253–64 (1987).
- [37] Urdea, M. S., Wilber, J. C., Yeghiazarian, T., Todd, J. A., Kern, D. G., Fong, S. J., Besemer, D., Hoo, B., Sheridan, P. J., Kokka, R., and et al. Direct and quantitative detection of HIV-1 RNA in human plasma with a branched DNA signal amplification assay. *AIDS* **7 Suppl 2**, S11–4 (1993).
- [38] Nilsen, T. W., Grayzel, J., and Prenskey, W. Dendritic nucleic acid structures. *Journal of theoretical biology* **187**(2), 273–84 (1997).
- [39] Stears, R. L., Getts, R. C., and Gullans, S. R. A novel, sensitive detection system for high-density microarrays using dendrimer technology. *Physiological genomics* **3**(2), 93–9 (2000).
- [40] Mora, J. R. and Getts, R. C. Enzymatic microRNA detection in microtiter plates with DNA dendrimers. *BioTechniques* **41**(4), 420, 422, 424 (2006).
- [41] Yang, W., Maqsoodi, B., Ma, Y., Bui, S., Crawford, K. L., McMaster, G. K., Witney, F., and Luo, Y. Direct quantification of gene expression in homogenates of formalin-fixed, paraffin-embedded tissues. *BioTechniques* **40**(4), 481–6 (2006).

- [42] Tsongalis, G. J. Branched DNA technology in molecular diagnostics. *American journal of clinical pathology* **126**(3), 448–53 (2006).
- [43] Pfaffl, M. W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic acids research* **29**(9), e45 (2001).

Appendix A

Supplementary information for Chapter 2

A.1 Stepping gel

Each step of the conditional catalytic Dicer substrate formation mechanism was validated by comparing conversion to anneal on native polyacrylamide gel electrophoresis.

- Detection target X_{short} , hairpin A, hairpin B, and hairpin C are run individually as reactant size markers (lanes 1–4, Figure A.1).
- Step 1: Detection target X_{short} and hairpin A interact to form a band that corresponds to product $X_{short}\cdot A$ (lane 5, Figure A.1). As expected, this product migrates at about the same speed as the annealed product $X_{short}\cdot A$ (lane 6, Figure A.1).
- Step 2: Detection target X_{short} , hairpin A, and hairpin B interact to form a band that corresponds to product $X_{short}\cdot A\cdot B$ (lane 7, Figure A.1). As expected, this product migrates at about the same speed as the annealed product $X_{short}\cdot A\cdot B$ (lane 8, Figure A.1).
- Step 3: Detection target X_{short} , hairpin A, hairpin B, and hairpin C interact to form bands that correspond to products $X_{short}\cdot A$, $B\cdot C$ and $X_{short}\cdot A\cdot B\cdot C$ (lane 9, Figure

A.1). As expected, similar products are observed in the annealed reaction (lane 10, Figure A.1).

- In the absence of a detection target (OFF state), the hairpins are metastable and minimal B·C is formed (lane 11, Figure A.1). When all three hairpins are annealed, a product corresponding to B·C is formed (lane 12, Figure A.1).
- In the absence of a detection target, hairpins B and C are metastable (lane 13, Figure A.1). When the two hairpins are annealed, a product corresponding to B·C is formed (lane 14, Figure A.1). Lane 14 also serves as a size marker for the expected final product B·C.
- Not all of B and C are consumed in the anneal, possibly because the basepairs in the hairpins form more easily than intermolecular basepairs.

A.2 Quantification of ON and OFF states

Three independent experiments were carried out to examine the ON and OFF states of conditional Dicer substrate formation (Figure A.2). The amount of Dicer substrate B·C formed was quantified relative to detection target X_{short} . The mechanism is turned ON by either a short synthetic target or DsRed mRNA (lanes X_{short} and X, respectively, in Figure A.2). Minimal B·C is produced in the OFF states. The OFF state was examined in the absence of a detection target, in the presence of an off-target mRNA GAPDH or in the presence of the silencing target d2EGFP mRNA (lanes “no target”, Z and Y, respectively in Figure A.2).

A.3 Quantification of catalytic property

Three independent experiments were carried out to examine the catalytic property of conditional Dicer substrate formation. The amount of Dicer substrate B·C formed was quantified

relative to $1\times$ detection target X_{short} . Minimal B·C is produced in the absence of a detection target, whereas as little as $0.1\times X_{short}$ is sufficient to convert roughly 50% of hairpins B and C. Conversion with $0.3\times X_{short}$ is equivalent to conversion with $1\times X_{short}$ (Figure A.3).

A.4 *In vitro* Dicer assay

Validation that only the final product, B·C, of the conditional catalytic DsiRNA formation mechanism is processed by Dicer. Each step of the reaction was subjected to incubation in Dicer reaction conditions in the presence '+' lanes or absence '-' lanes of Dicer. For a detailed description of each step see section A.1. Complex B·C is formed in the presence of detection target X_{short} and hairpin A (lanes 13, 14, Figure A.4(a)), or when hairpins B and C are pre-annealed (lanes 15, 16, Figure A.4(a)). Minimal B·C is formed in the absence of X_{short} (lanes 7, 8, Figure A.4(a)). Native and denaturing polyacrylamide gel electrophoresis demonstrate that the B·C complex is processed by Dicer to form an siRNA (lanes 8, 14, 16, Figure A.4(a) and (b), respectively). X_{short} ·A is not fully denatured under these conditions as evidenced by lanes 9 through 14 of Figure A.4(b).

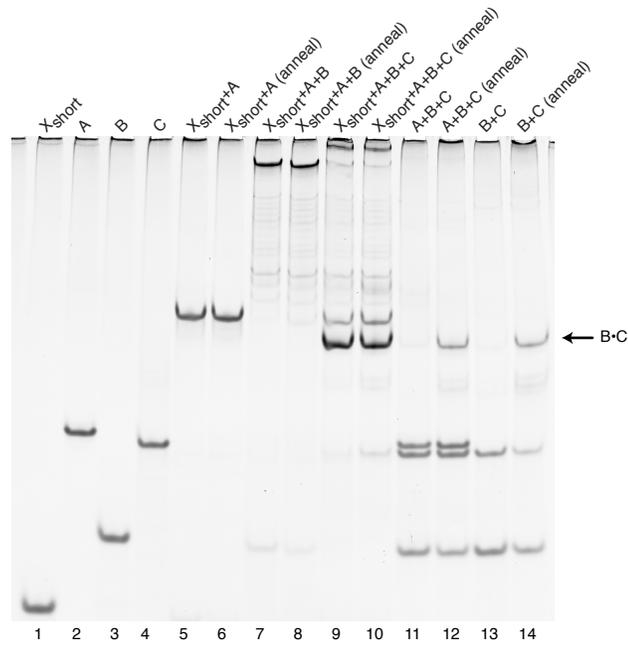


Figure A.1: **Stepping gel for the conditional catalytic DsiRNA formation mechanism.** Native polyacrylamide gel electrophoresis demonstrating each step of the reaction depicted in Figure 2.1. A $0.5\mu\text{M}$ reaction master mix was prepared for each duplicate. The master mix was then split into two separate reactions: 2-hour incubation at 37°C and an anneal. Reactions were run on 20% polyacrylamide gel at 200 volts for 9.5 hours in $1\times\text{TBE}$ followed by SYBRGold staining.

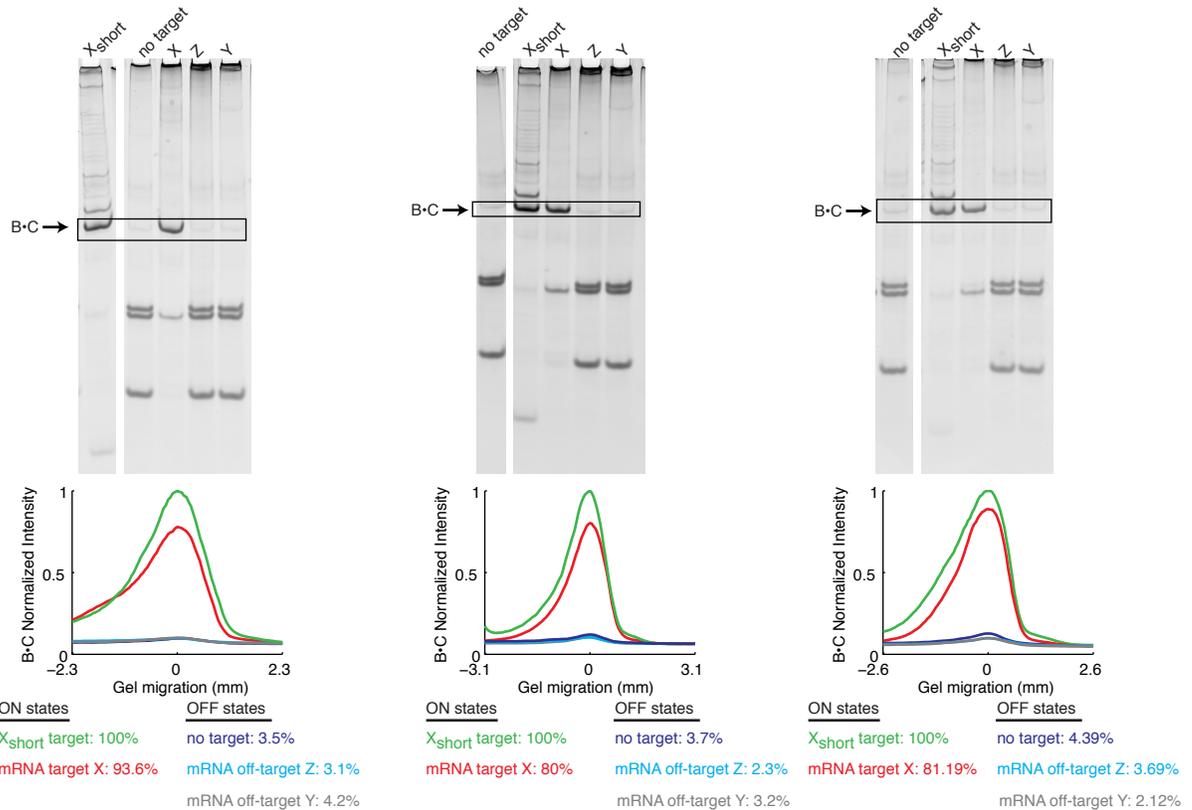


Figure A.2: **Quantification of B·C final product formation in ON and OFF states.** ON states (X_{short}, DsRed mRNA target X) and OFF states (no target, GAPDH mRNA target Z and silencing d2EGFP target Y) of the mechanism. Plots represent the normalized intensity of B·C in each lane across three separate experiments. Quantification was done as described in methods section of Chapter 2. A 0.5 μM reaction master mix containing A, B and C was prepared. The master mix was then split into the separate reactions and targets were added for a 2-hour incubation at 37°C. Reactions were run on 20% polyacrylamide gel at 200 volts for 10.5 hours in 1×TBE followed by SYBRGold staining. Images were edited to remove irrelevant lanes, missing lanes are represented as a white space.

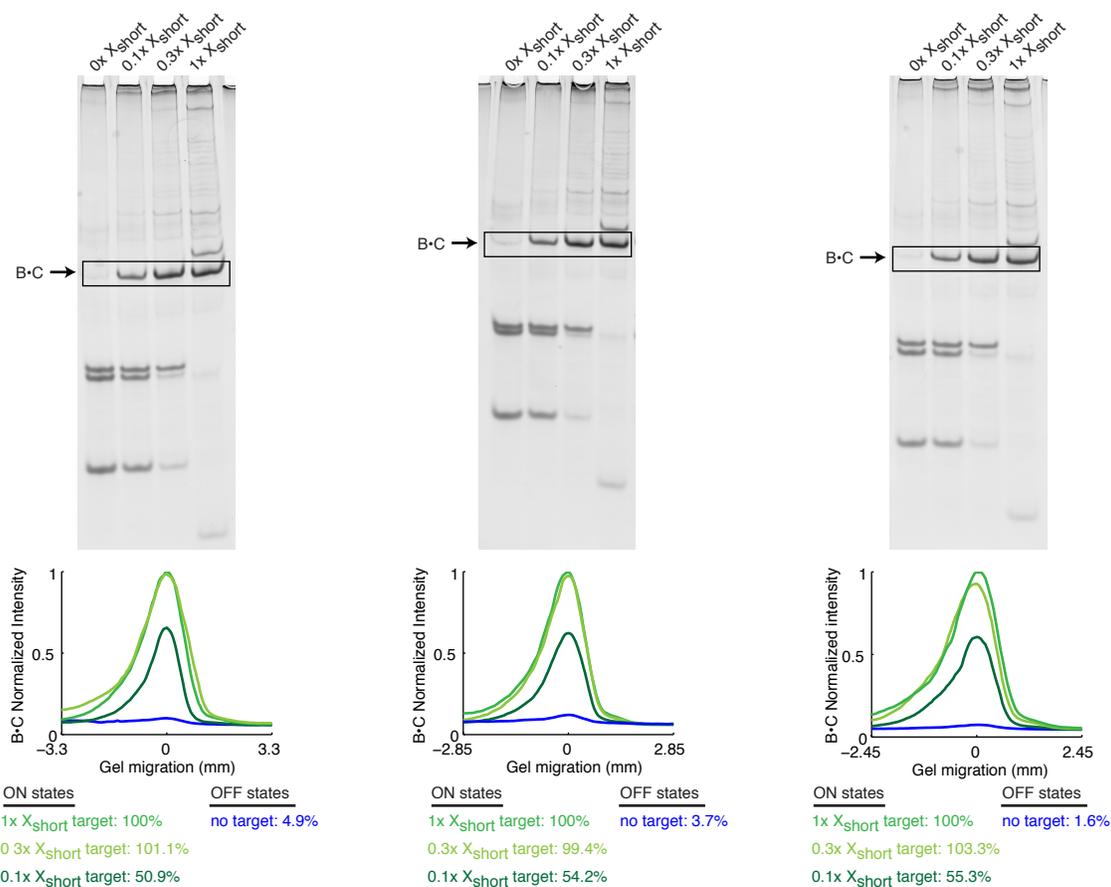


Figure A.3: **Quantification of catalytic B·C final product formation.** Sub-stoichiometric amounts of target X_{short} generate above stoichiometric amounts of B·C. Plots represent the normalized intensity of B·C in each lane across three separate experiments. Quantification was done as described in methods section of Chapter 2. A $0.5\mu\text{M}$ reaction master mix containing A, B and C was prepared. The master mix was then split into the separate reactions and targets were added for a 2-hour incubation at 37°C . Reactions were run on 20% polyacrylamide gel at 200 volts for 10.5 hours in $1\times\text{TBE}$ followed by SYBRGold staining.

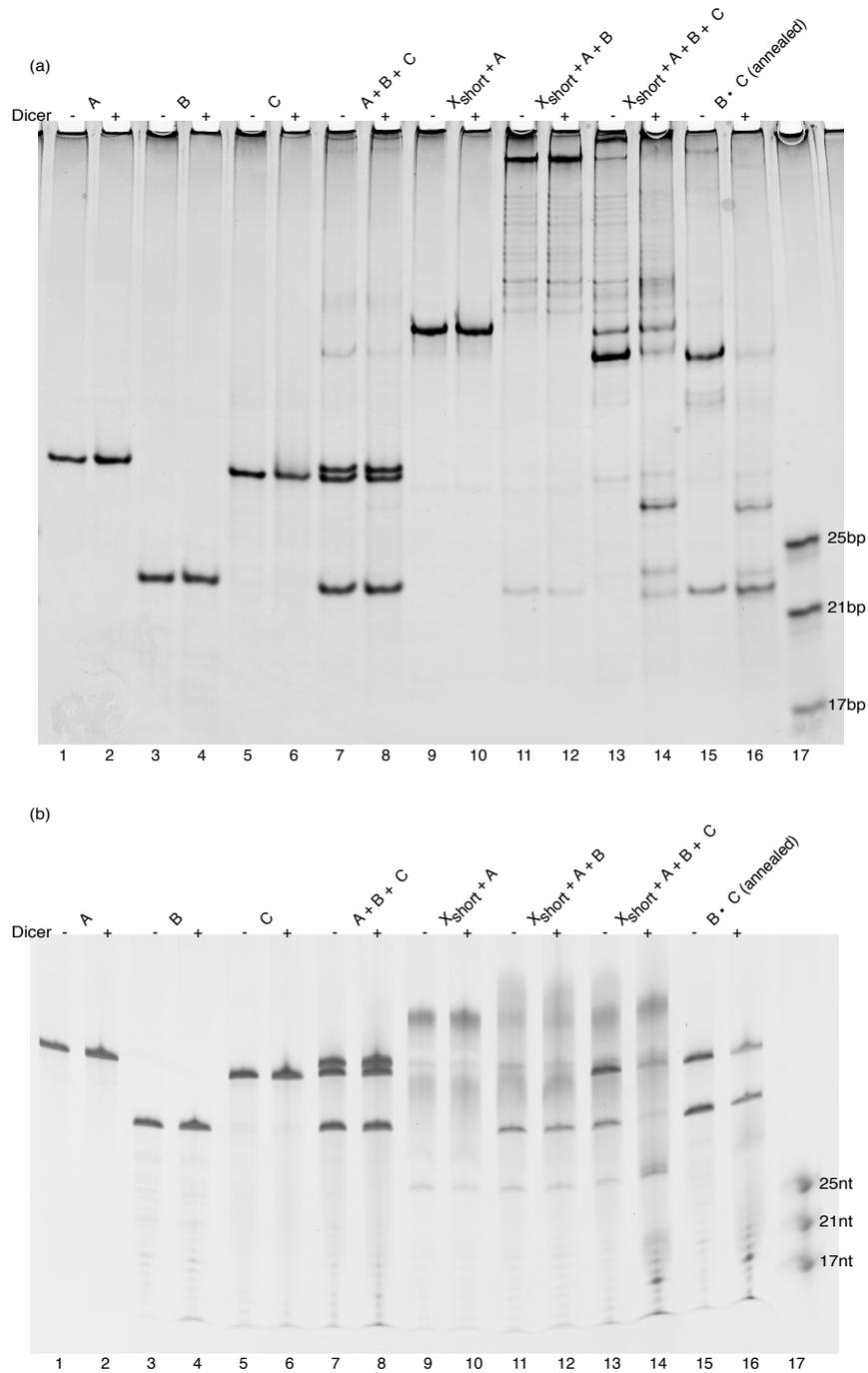


Figure A.4: **Dicer processing gel for the conditional catalytic DsiRNA formation mechanism.** $0.5\mu\text{M}$ reaction master mix was prepared for each duplicate. The master mix was then split into two separate reactions, with and without Dicer. Reactions were incubated for 2 hours at 37°C . (a) Native polyacrylamide gel electrophoresis. $4\mu\text{l}$ of each reaction were run on 20% polyacrylamide gel at 200 volts for 10.5 hours in $1\times\text{TBE}$ followed by SYBRGold staining. (b) Denaturing polyacrylamide gel electrophoresis. 14% denaturing polycacrylamide gel (with formamide) was pre-run at 15W $1\times\text{TBE}$ for 1 hour. $3\mu\text{l}$ of each reaction were run at 15W for 1 hour in $1\times\text{TBE}$ followed by SYBRGold staining.

Appendix B

Supplementary information for Chapter 3

B.1 Catalytic Dicer substrate formation mechanism based on 5'-toehold hairpins

Prior to designing the conditional Dicer substrate formation mechanism presented in Chapter 2, several mechanisms using a 5'-toehold based hairpins were explored. Figure B.1(a) depicts a schematic representation of such a mechanism (M1). *In vitro* Dicer cleavage assays of strands corresponding to mechanism M1 demonstrate that hairpin C, as well as complexes $X_{short} \cdot A$ and $X_{short} \cdot A \cdot B$, get degraded by Dicer (Figure B.1(b), no Dicer '−' lanes vs. Dicer '+' lanes). Some degradation is also observed for hairpins A and B. Using 2'-OMe chemical modifications (hairpins A2, B2 and C2) abrogate this pattern such that only the final product B2·C2 is processed (Figure B.1(c), no Dicer '−' lanes vs. Dicer '+' lanes). Using a mechanism with hairpins that is based on 5'-toeholds results in premature exposure of the antisense strand of the siRNA. In complex $X_{short} \cdot A \cdot B$ a single-stranded region of hairpin B is exposed containing the domains 'z*-y*-x*' which comprise the antisense (Figure B.1(a)); this might affect the performance of the mechanism. For this reason, this current version of a 5'-toehold mechanism is not suitable for *in vivo* application.

Strand	Sequence
M1.A2	<u>CUCGAUCUCGAACUCGUGGCUGGUCAGCUUGCCGUACACGA</u> <u>GUUCG</u>
M1.B2	<u>CGAACUCGUGUACGGCAAGCUGACCGAGACUUCAGGGUCAG</u> <u>CUUGCCGUACA</u>
M1.C2	<u>UACGGCAAGCUGACCCUGAAGUCUCGGUCAGCUUGCCGUAC</u> <u>ACGAGACUUCAGGGUCAGC</u>

Table B.1: List of M1 sequences modified with 2'-OMe. 2'-OMe modifications are underlined.

Appendix C

Supplementary information for Chapter 4

C.1 Stepping gel

Each step of the conditional shRNA transcription mechanism was validated by comparing conversion to anneal on native polyacrylamide gel electrophoresis. Both DNA and RNA X_{short} targets were used (Figure C.1 (a) and (b), respectively). The order of the lanes is identical across both gels.

- Detection target X_{short} , hairpin A, and hairpin B are run individually as reactant size markers (lanes 1–3, Figure C.1).
- Step 1: Detection target X_{short} and hairpin A interact to form a band that corresponds to product $X_{short}\cdot A$ (lane 4, Figure C.1). As expected, this product migrates at about the same speed as the annealed product $X_{short}\cdot A$ (lane 5, Figure C.1).
- Step 2: Detection target X_{short} , hairpin A, and hairpin B interact to form a band that corresponds to product $X_{short}\cdot A\cdot B$ (lane 6, Figure C.1). As expected, this product migrates at about the same speed as the annealed product $X_{short}\cdot A\cdot B$ (lane 7, Figure C.1).

- In the absence of a detection target, the hairpins are metastable and minimal A·B is formed (lane 8, Figure C.1). When the two hairpins are annealed, a product corresponding to A·B is formed (lane 9, Figure C.1).

C.1.1 Conditional shRNA transcription and Dicer gel

Transcription products for each step of the conditional shRNA transcription mechanism were examined by radioactive *in vitro* transcription (odd lanes (‘-’ Dicer) in Figure C.2). An *in vitro* Dicer cleavage assay demonstrates that the products are cleaved to generate the expected siRNA (even lanes (‘+’ Dicer) in Figure C.2).

- The disrupted T7 promoter is inactive. Minimal transcription is observed from either hairpin A or B (lanes 1 and 3, respectively, Figure C.2).
- Step 1: Some transcription is observed from $X_{short}\cdot A$ (lane 5, Figure C.2). The product is larger than the expected shRNA and is minimally present in the presence of hairpin B (lane 5 vs. lane 9, Figure C.2).
- Step 2: The major transcription product of $X_{short}\cdot A\cdot B$ corresponds to an shRNA in size (lane, 9 Figure C.2). As expected, this product is cleaved by Dicer to generate an siRNA (lane 10, Figure C.2).
- In the absence of a detection target (OFF state), the hairpins are metastable and minimal shRNA is transcribed (lane 7, Figure C.2).

C.1.2 Quantification of ON-to-OFF ratios

The ON and OFF ratios of conditional shRNA transcription were determined by quantifying the shRNA transcription product. As expected, minimal transcription is observed in the absence of a detection target X_{short} (Figure C.3).

C.2 Design of a triggered shRNA transcription mechanism using the H1 promoter

Promoter choice

The second promoter considered for conditional shRNA transcription was the mammalian H1 promoter; discussed below are its advantages and disadvantages. The H1 promoter is a mammalian RNA polymerase III (Pol III) promoter that naturally drives the expression of small RNAs and is commonly used to express shRNAs [1–4]. It is relatively small (100bp), and has well-defined transcription and termination sites [5, 6]. Transcription is terminated when Pol III encounters a run of four or five thymidines; we have therefore added a poly-T termination sequence after the hairpin (domain ‘t’ in Figure 4.1). Another advantage of the H1 promoter is that it is endogenously present in cells (unlike T7 RNA polymerase). However, this raises the possibility of interfering with endogenous pathways which is a disadvantage. Other concerns were the promoter size and the lack of *in vitro* assays for transcription. Despite being short, relative to other promoters, 99bp are significantly longer than molecules previously engineered in the lab and we were not sure whether we would be able to engineer them to undergo toehold-mediated branch migration.

Design

The design follows the same logic operation as the T7-promoter-based system described in Chapter 4. However, whereas the T7-promoter-based design is metastable, the H1-promoter-based system is designed to be stable. Meaning, the secondary structure of the starting components (hairpins A and B) is the thermodynamic minimum and therefore it is not favorable for hairpins A and B to interact. We also added cooperative and stronger binding to this design. Segment ‘a*’ was added to the toehold of hairpin B, this segment also binds the detection target X. Therefore, both hairpin A and hairpin B bind to the target, further stabilizing their interaction (Figure C.4). Sequence design and thermodynamic analysis was performed as discussed in Chapter 4. The transcribed shRNA is designed to

down-regulate d2EGFP.

Hairpin construction

A challenge present in this design, in comparison to the T7 promoter design, is the size of the hairpins. Despite using a relatively short promoter, the promoter itself is 99 nucleotides long. In addition, the shRNA and target detection sequences need to be added. The total length of hairpin A is 290 nucleotides, and of hairpin B 256 nucleotides. In order to construct these hairpins, each hairpin was ordered as four segments which were annealed together by heating to 95°C for 2 minutes followed by a controlled gradual cooling at -1°C per minute to 23°C in a PCR block in 1× T4 DNA ligase buffer (NEB). After the anneal, T4 DNA ligase was added to the reaction and ligation was carried out overnight at 16°C followed by a denaturing PAGE purification and ethanol precipitation. The full length of the hairpins was verified by using fluorescently labeled probes that bind to the 5′ or 3′ end of each hairpin (data not shown). Helper strands (strands that bind to the hairpin to assist in binding the probe) were used to facilitate binding to the 3′ end of hairpin A and 5′ end of hairpin B, helper strands were used.

***In vitro* studies**

Each step of the conditional shRNA transcription mechanism was validated by comparing conversion to anneal on agarose gel electrophoresis.

- Detection target X_{short} , hairpin A, and hairpin B were run individually as reactant size markers; X_{short} cannot be visualized under these conditions (lanes 1–3, Figure C.5).
- Step 1: Detection target X_{short} and hairpin A interact to form a band that corresponds to product $X_{short}\cdot A$ (lane 4, Figure C.5). As expected, this product migrates at about the same speed as the annealed product $X_{short}\cdot A$ (lane 5, Figure C.5). Hairpin B cannot bind to X_{short} without cooperativity from hairpin A (lanes 6,7 Figure C.5).

- Step 2: Detection target X_{short} , hairpin A, and hairpin B interact to form a band that corresponds to product $X_{short} \cdot A \cdot B$ (lane 8, Figure C.5). As expected, this product migrates at about the same speed as the annealed product $X_{short} \cdot A \cdot B$ (lane 9, Figure C.5).
- In the absence of a detection target (OFF state), the hairpins are stable and minimal $A \cdot B$ is formed. As expected from stable hairpins, when annealed, minimal product corresponding to $A \cdot B$ is formed (lanes 9 and 10, respectively, Figure C.5). It is yet to be determined whether the observed higher products are due to A or B dimers or to the product $A \cdot B$.

The components of this mechanism are significantly larger than previously studied hairpins in the lab, requiring a branch migration of over 100 bases. Not only were we able to design a system composed of large components that performs well (good ON and OFF states), but this is also a first demonstration of long branch migration *in vitro*.

To the best of our knowledge, no *in vitro* Pol III transcription system is available. To obtain Pol III transcription *in vitro*, we tried to use HeLaScribe Nuclear Extract *in vitro* Transcription system (Promega). However, no transcription was observed. The kit is specifically designed to use with Pol II promoters and therefore may not contain Pol III or may not have appropriate conditions for Pol III transcriptions. Due to the presence of RNA Pol III in cells we chose to proceed to *in vivo* studies.

***In vivo* studies**

The H1 promoter is being transcribed by the mammalian RNA Pol III; therefore, transfection of our mechanism is expected to down-regulate d2EGFP. As described in Chapter 4, the two parts of the mechanism are transcription template formation and shRNA transcription. Our first goal was to examine whether the final product can be transcribed and lead to RNAi-mediated d2EGFP knockdown. This was tested by transfection of an annealed final product ($X_{short} \cdot A \cdot B$) into d2EGFP-expressing cells. As controls, plasmid

transfection and a linear PCR fragment of the H1 promoter driving the expression of an shRNA targeting the same d2EGFP region (PCR amplification off a plasmid, see Materials and methods) were used. To facilitate the transfection of short linear DNA fragments, pOri-e1 plasmid (generous gift from Fred Tan) was used as carrier DNA. The selected shRNA target down-regulates d2EGFP as can be seen by the pSilencer transfection in Figure C.6. The annealed mechanism does not lead to d2EGFP knockdown (Figure C.6, Annealed). Transfection of a short linear DNA fragment (Figure C.6, PCR) is able to knockdown d2EGFP about 70%, suggesting that the transfection conditions are suitable. This suggests that the annealed structure is not being transcribed, possibly due to steric hindrance making the H1 promoter inaccessible to Pol III. A spacer sequence between the target binding site and the H1 promoter might be needed in order to fit RNA Pol III and to achieve transcription of the $X_{short} \cdot A \cdot B$ template. Further studies are needed in order to achieve transcription of an annealed mechanism in tissue culture. Once resolved, the full mechanism (transcription template formation followed by shRNA transcription) can be studied. Optimization will most likely be needed in determining the appropriate dimensions for promoter segments ‘p1’ and ‘p2’. Indeed, preliminary data of hairpin A transfection into d2EGFP-expressing cells suggests that the current segmentation may result in d2EGFP knockdown (data not shown).

Conclusion

We have designed a stable conditional shRNA transcription system based on the mammalian H1 promoter. The use of the H1 promoter requires that the system components (hairpins) be significantly larger than previously studied. Using large hairpins we were able to engineer conditional hybridization cascades that performed according to design. The hairpins did not interact in the absence of a detection target, while the presence of a detection target resulted in the expected final product formation. Due to the lack of an established Pol III *in vitro* transcription system we were unable to test transcriptional activity of our mechanism. Transfection of short PCR fragments containing a d2EGFP

shRNA under the control of the H1 promoter resulted in approximately 70% d2EGFP knockdown, whereas an annealed reaction had no effect on d2EGFP expression. Additional work is required to obtain correct dimensions that enable RNA Pol III to transcribe the H1 promoter while bound to a template, as well as make the H1 promoter inactive in its initial, partially unpaired, state of each hairpin.

In comparison to the T7-promoter-based system presented in Chapter 4, the study of this mechanism is more challenging. The need for large hairpins requires a difficult ligation of several templates making small modifications to the system cumbersome. The lack of an *in vitro* transcription system is also a drawback, making it difficult to decouple transcription and transfection. The endogenous presence of RNA Pol III in mammalian cells is both an advantage and a disadvantage. The advantage is that a special cell line does not need to be created to study the system, while the disadvantage is that our mechanism might interfere with endogenous processes making it more difficult to interpret results. However, from a practical standpoint, the use of the H1-promoter-based system is more appropriate both for studying gene function via conditional RNAi and for therapeutic applications.

C.3 Materials and methods

Strand sequences. The hairpins and detection target for this mechanism are DNA. The H1 promoter sequence used (Table C.1) was adapted from pSilencer3.1 plasmid (Ambion). The d2EGFP shRNA target sequence (Table C.1) was adapted from Beisel et al.[7] (corresponding to bases 128–150). The shRNA 3' overhang is a UU incorporated from the transcription termination signal. The target used as the input for this design is a random sequence. See Table C.2 for full sequences.

Cloning of pSilencer d2EGFP shRNA. The shRNA coding sequence corresponding to the same d2EGFP region as the mechanism (see above) was cloned into pSilencer3.1 H1 Hygro (Ambion) according to the manufacturer using the recommended loop sequence.

Domain	Size
shRNA target	GAC CCT GAA GTT CAT CTG CAC C
H1 promoter	AAT TCA TAT TTG CAT GTC GCT ATG TGT TCT GGG AAA TCA CCA TAA ACG TGA AAT GTC TTT GGA TTT GGG AAT CTT ATA AGT TCT GTA TGA GAC CAC TCG

Table C.1: List of strands for H1-promoter-based shRNA transcripton.

Strand	Size
X	GCT ATA ACG CAT AAT CAC CTC ATA ACA GTT CAA TCT CCC
A	GGG AGA TTG AAC TGT TAT GAG GTG ATT ATA ATT CAT ATT TGC ATG TCG CTA TGT GTT CTG GGA AAT CAC CAT AAA CGT GAA ATG TCT TTG GAT TTG GGA ATC TTA TAA GTT CTG TAT GAG ACC ACT CGG GTG CAG ATG AAC TTC AGG GTC CTC ACA GAC CCT GAA GTT CAT CTG CAC CCG AGT GGT CTC ATA CAG AAC TTA TAA GAT TCC CAA ATC CAA AGA CAT TTC ACG TTT ATG GTG ATT TCC CAG AAC ACA TAG CGA CAT GCA AAA TAA TCA CCT CAT AAC AGT TC
B	TTT GCA TGT CGC TAT GTG TTC TGG GAA ATC ACC ATA AAC GTG AAA TGT CTT TGG ATT TGG GAA TCT TAT AAG TTC TGT ATG AGA CCA CTC GAA AAA AGG TGC AGA TGA ACT TCA GGG TCT GTG AGG ACC CTG AAG TTC ATC TGC ACC CGA GTG GTC TCA TAC AGA ACT TAT AAG ATT CCC AAA TCC AAA GAC ATT TCA CGT TTA TGG TGA TTT CCC AGA ACA CAT AGC GAC ATG CAA ATA TGA ATT GCG TTA TAG C

Table C.2: List of strands for H1-promoter-based shRNA transcripton.

Briefly, shRNA template oligonucleotides were diluted to $1\mu\text{g}/\mu\text{l}$ and annealed in DNA annealing solution by heating to 90°C for 3 minutes and incubating for one hour at 37°C . The annealed template was then diluted to $8\text{ng}/\mu\text{l}$ and ligated into the pSilencer vector. Clones were verified by sequencing.

Strand	Sequence
5'	GAT CCG CCC TGA AGT TCA TCT GCA CTT CAA GAG AGT GCA GAT GAA CTT CAG GGT CTT TTT TGG AAA
3'	AGC TTT TCC AAA AAA GAC CCT GAA GTT CAT CTG CAC TCT CTT GAA GTG CAG ATG AAC TTC AGG GCG

Note: The shRNA sequence should read 'GACCC...' but instead is 'GCC...' (missing A in position 130 of d2EGFP). This does not affect the silencing ability of the shRNA.

Table C.3: DNA templates for cloning d2EGFP shRNA into pSilencer plasmid.

Linear DNA template from pSilencer d2EGFP shRNA. A short PCR fragment containing the H1 promoter and d2EGFP shRNA was PCR amplified off the pSilencer d2EGFP shRNA plasmid using the following primers and conditions:

5' primer, m13(-20) forward: 5'-GTA AAA CGA CGG CCA GT-3'

3' primer, pSilencer sequencing rev: 5'-GAG TTA GCT CAC TCA TTA GGC -3'

PCR conditions: annealing temperature of 55°C, protocol according to AccuPrime PFX supermix (Invitrogen). The PCR product was either gel purified from an agarose gel using the QIAquick gel extraction kit (Qiagen) or using the QIAquick PCR purification kit according to the manufacturer.

Bibliography

- [1] Brummelkamp, T. R., Bernards, R., and Agami, R. A system for stable expression of short interfering RNAs in mammalian cells. *Science* **296**(5567), 550–3 (2002).
- [2] Paddison, P. J., Caudy, A. A., Bernstein, E., Hannon, G. J., and Conklin, D. S. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes & development* **16**(8), 948–58 (2002).
- [3] Paul, C. P., Good, P. D., Winer, I., and Engelke, D. R. Effective expression of small interfering RNA in human cells. *Nature biotechnology* **20**(5), 505–8 (2002).
- [4] Paule, M. R. and White, R. J. Survey and summary: transcription by RNA polymerases I and III. *Nucleic acids research* **28**(6), 1283–98 (2000).
- [5] Myslinski, E., Ame, J. C., Krol, A., and Carbon, P. An unusually compact external promoter for RNA polymerase III transcription of the human H1 RNA gene. *Nucleic acids research* **29**(12), 2502–9 (2001).
- [6] Gunnery, S., Ma, Y., and Mathews, M. B. Termination sequence requirements vary among genes transcribed by RNA polymerase iii. *Journal of molecular biology* **286**(3), 745–57 (1999).
- [7] Beisel, C. L., Bayer, T. S., Hoff, K. G., and Smolke, C. D. Model-guided design of ligand-regulated RNAi for programmable control of gene expression. *Molecular systems biology* **4**, 224 (2008).

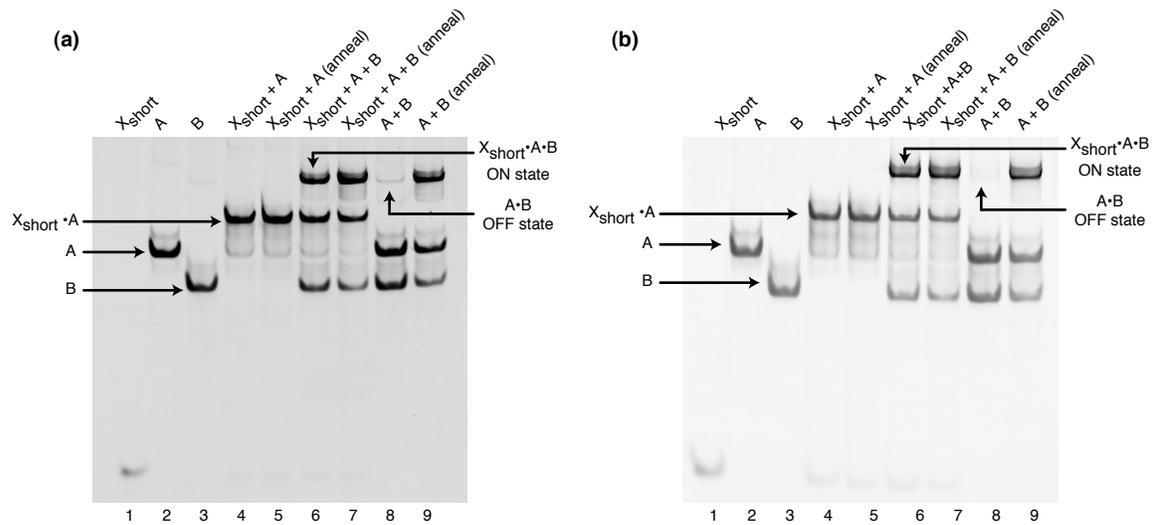


Figure C.1: **Conditional shRNA transcription template formation.** Native polyacrylamide gel electrophoresis demonstrating each step of the reaction depicted in Figure 4.2. (a) DNA detection target X_{short} . (b) RNA detection target X_{short} . A $0.5\mu\text{M}$ reaction master mix was prepared for each duplicate. The master mix was then split into two separate reactions: 2-hour incubation at 37°C and an anneal. Reactions were run on 10% polyacrylamide gel at 200 volts for 1.5 hours in $1\times\text{TBE}$ followed by SYBRGold staining.

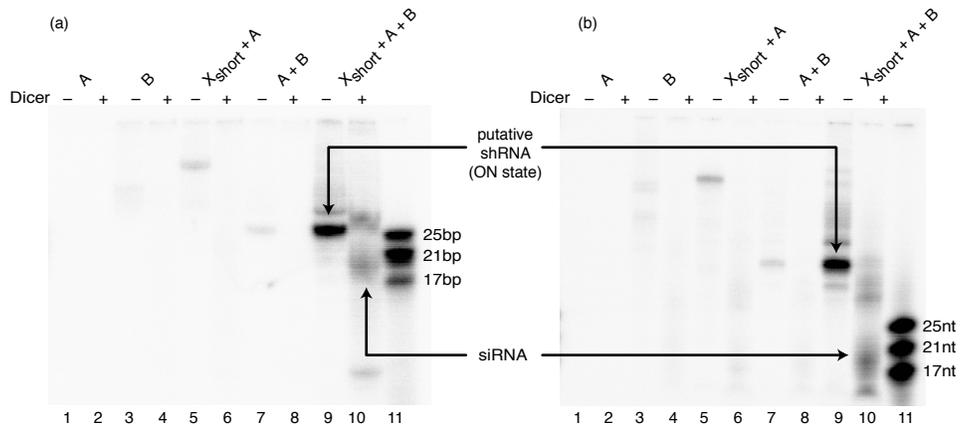


Figure C.2: **Conditional shRNA transcription and Dicer processing.** Gel electrophoresis demonstrating the transcription product of each step of the reaction depicted in Figure 4.1 of Chapter 4. Odd lanes (–) were not subjected to Dicer processing, while even lanes (+) were cleaved by Dicer. Each reaction was split into two, one half was run on a native gel and the other on a denaturing gel. (a) 20% native polyacrylamide gel electrophoresis, 250 volts for four hours. (b) 15% denaturing polyacrylamide gel electrophoresis, 500 volts for 1.5 hours.

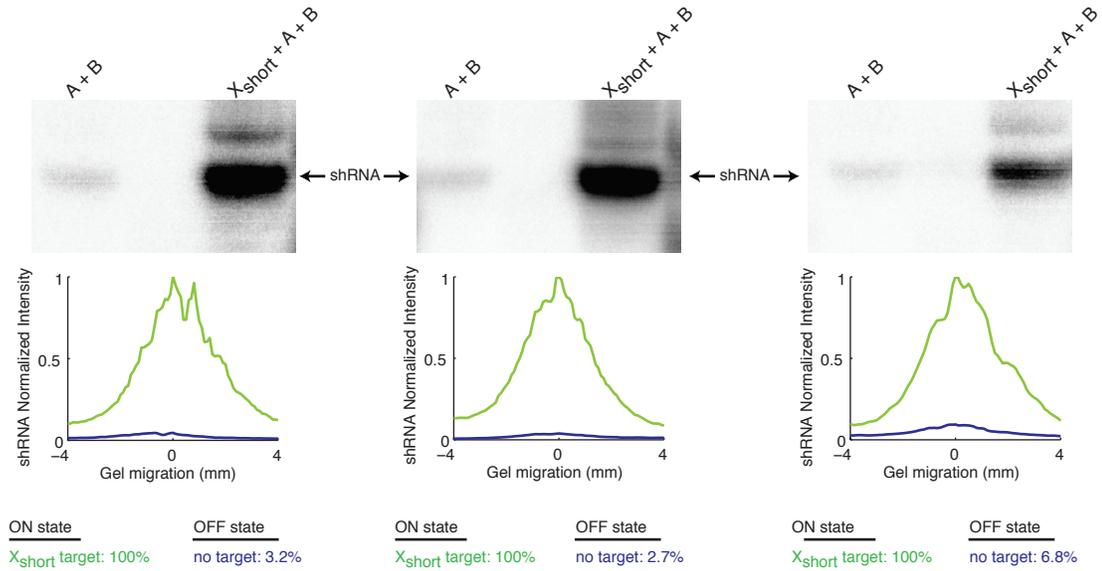


Figure C.3: **Quantification of transcribed shRNA product.** Radioactive *in vitro* transcription in the presence or absence of detection target X_{short} . The shRNA band in three separate experiments was quantified and the band intensity was plotted. For methods see the Materials and methods section in Chapter 4.

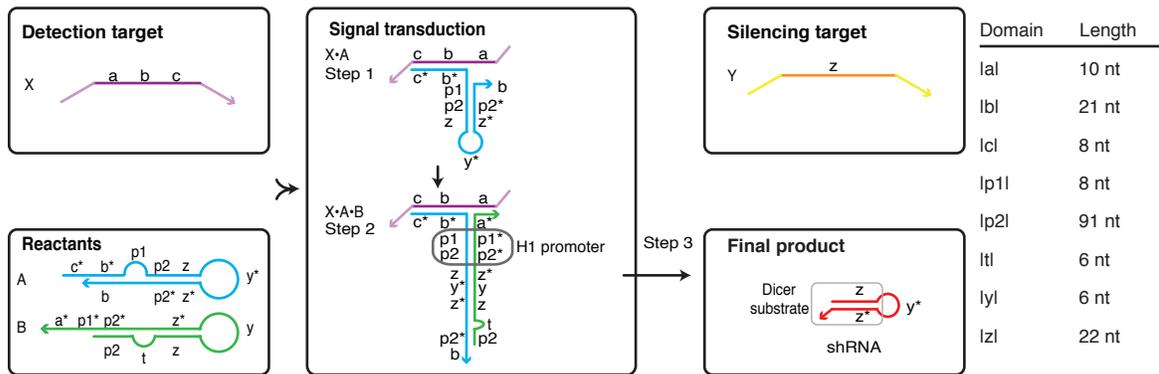


Figure C.4: **Conditional shRNA transcription mechanism based on H1 promoter schematic.** The detection target X is recognized by A (step 1), leading to the binding and opening of B (step 2), which in turn forms an intact promoter and drives the transcription of an shRNA against the silencing target Y (step 3).

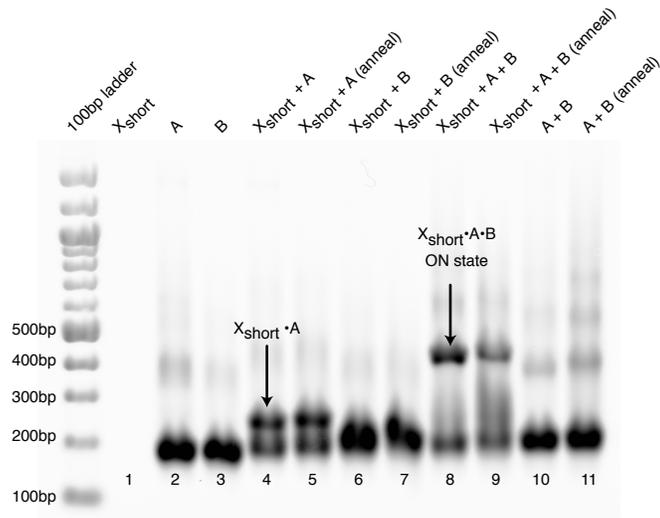


Figure C.5: **Conditional shRNA transcription template formation with H1 promoter.** Agarose gel electrophoresis demonstrating each step of the reaction depicted in Figure C.4. $0.5\mu\text{M}$ reactions were incubated for 2.75 hours at 37°C or annealed by heating to 95°C for 5 minutes followed by gradual cooling ($-1^\circ\text{C}/\text{minute}$) to 23°C . Reactions were run on 1% agarose gel, SYBRGold was pre-added to the loading dye.

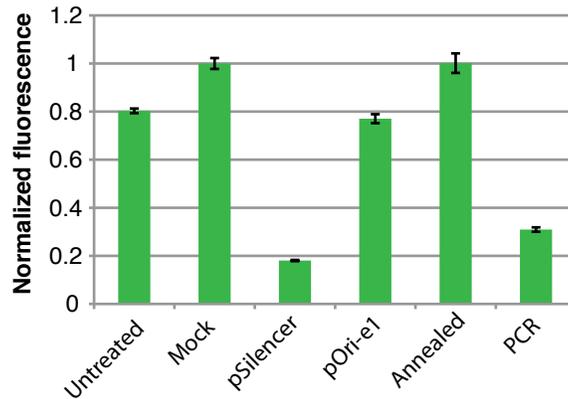


Figure C.6: **Transfection of an annealed H1-promoter-based mechanism does not lead to d2EGFP knockdown.** HEK293 d2EGFP cells were transfected with pSilencer d2EGFP plasmid (pSilencer), pOri-e1 (pOri), pOri-e1 with an annealed mechanism (Annealed) or pOri-e1 with a linear DNA template containing the H1 promoter and shRNA PCR amplified from pSilencer (PCR). Transfection was done in a 24-well plate format according to the Lipofectamine2000 protocol (Invitrogen) using 800ng plasmid DNA and 180ng linear DNA. 44 hours post transfection, cells were analyzed on a flow cytometer. Transfections were normalized relative to a mock transfection.