

Signal Transduction with Hybridization Chain Reactions

Thesis by

Jonathan Ben-Zion Sternberg

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2013

(Defended May 22nd, 2013)

© 2013

Jonathan Ben-Zion Sternberg

All Rights Reserved

Acknowledgements

I would like to thank my advisor, Professor Niles Pierce, for enlightening scientific conversations, as well as for providing me with the intellectual community, resources, and most importantly, freedom necessary to carry out the exciting work presented in this thesis. I would also like to thank the great group of people I have had the luck to know during my time in the Pierce Lab. Thank you, Dr. Victor Beck, Ma'ayan Schwarzkopf, Dr. Jeff Vieregg, Dr. Tobias Heinen, Lisa Hochrein, Dr. Harry Choi, and the rest of the lab members. You have all helped me in a variety of ways. For help with proofreading my thesis, I thank Melinda Kirk.

I would also like to thank the wonderful committee members with whom I have had the luck to work. Thank you, Professor David Tirrell, Professor Scott Fraser, and Professor Michael Elowitz; your insights and support throughout these last few years have been instrumental.

I owe a great deal of thanks to my family as well. Mom and Dad (Ima and Aba), if it were not for your continuous love and support for more than three decades, I would not be where I am today. Thank you for instilling in me the values that allowed me to pursue scientific work at Caltech and the motivation to work hard and to strive to achieve my personal best. My brother, Ariel, thank you for your love and support, and for being the amazing friend that you are to me.

Last, but not least, I would like to thank my wife. Kayla, my dear, thank you for believing in me, for standing by me, and for being the wonderful, warm, and caring person that you are. Thank you for being my best friend (and writing editor) for the last 11 years of my life. Most importantly, thank you for giving me the two most beautiful gifts a

person can ask for: the gift of being loved by another, and the gift of children. Liam and Rafael, you have given me the most meaningful moments of my life. Debbie, Jeff, Sarah, Dan, and Avi, I could not have asked for a warmer, more supporting family than the one I have had the luck to acquire by marriage. I love you all.

Abstract

Some of the most exciting developments in the field of nucleic acid engineering include the utilization of synthetic nucleic acid molecular devices as gene regulators, as disease marker detectors, and most recently, as therapeutic agents. The common thread between these technologies is their reliance on the detection of specific nucleic acid input markers to generate some desirable output, such as a change in the copy number of an mRNA (for gene regulation), a change in the emitted light intensity (for some diagnostics), and a change in cell state within an organism (for therapeutics). The research presented in this thesis likewise focuses on engineering molecular tools that detect specific nucleic acid inputs, and respond with useful outputs.

Four contributions to the field of nucleic acid engineering are presented: (1) the construction of a single nucleotide polymorphism (SNP) detector based on the mechanism of hybridization chain reaction (HCR); (2) the utilization of a single-stranded oligonucleotide molecular Scavenger as a means of enhancing HCR selectivity; (3) the implementation of Quenched HCR, a technique that facilitates transduction of a nucleic acid chemical input into an optical (light) output, and (4) the engineering of conditional probes that function as sequence transducers, receiving target signal as input and providing a sequence of choice as output. These programmable molecular systems are conceptually well-suited for performing wash-free, highly selective rapid genotyping and expression profiling *in vitro*, *in situ*, and potentially in living cells.

Contents

Acknowledgements	iii
Abstract	v
List of Figures	ix
List of Tables	x
1 Introduction	1
References	4
2 Isothermal Enzyme-Free Detection of SNPs via Hybridization Chain Reaction	9
2.1 Introduction	9
2.2 Detection of SNPs with Hybridization Chain Reaction	11
2.3 Kinetic Control of HCR via ΔG Tuning	13
2.4 Experimental Verification of Kinetic Control of HCR	15
2.5 Effect of Mutation Location on SNP Discrimination	17
2.6 Detection of SNP Cancer Markers	19
2.7 Improving HCR Discrimination with Scavenger	21
2.8 SNP Profiling via HCR Multiplexing	24
2.9 Detection of SNPs in Long RNA	25
2.10 Conclusion	28
References	29

3	Transducing Sequence to Light with Quenched HCR	35
3.1	Introduction	35
3.2	Design of 2-Hairpin and 4-Hairpin Quenched HCR	37
3.3	Analysis of 2-Hairpin and 4-Hairpin Quenched HCR	38
3.4	Multiplexing Quenched HCR	40
3.5	Conclusion	43
	References	44
4	Sequence Transduction with Conditional Probes	46
4.1	Introduction	46
4.2	Design of Conditional Probes	49
4.3	Experimental Verification of Conditional Probe Function	50
4.4	Multiplexing via Conditional Probe and Quenched HCR	52
4.5	Detection of Long RNA	52
4.6	Conclusion	52
	References	56
	Appendices	57
A	Appendix to Chapter 2	58
A.1	Methods	58
A.2	Strand Sequences	60
B	Appendix to Chapter 3	64
B.1	Methods	64
B.2	Background-subtracted Data of Quenched HCR Multiplexing	66
B.3	Sensitivity of Quenched HCR	67
B.4	RNA-based Quenched HCR	68
B.5	Fast Target Detection with Quenched HCR	68
B.6	Strand Sequences	70

C Appendix to Chapter 4	71
C.1 Methods	71
C.2 Strand Sequences	71

List of Figures

2.1	SNP Detection	12
2.2	Kinetically Discrimination of SNPs via HCR at a Time Scale of Choice	16
2.3	Effect of Mutation Location on SNP Discrimination	18
2.4	Detection of SNP Cancer Markers	20
2.5	Scavenger- and HCR-Mediated SNP Detection	22
2.6	Generality of Scavenger- and HCR-Mediated Selectivity	23
2.7	SNP Profiling via HCR Multiplexing	26
2.8	Detection of SNPs in Long RNA	27
3.1	Design of Quenched HCR	39
3.2	End-labeled 2-Hairpin-Periodicity Quenched HCR vs. 4-Hairpin-Periodicity Quenched HCR	41
3.3	Multiplex Detection of Cancer Gene Markers	42
4.1	Illustration of Conditional Probe	48
4.2	Sequence Transduction with Conditional Probes	51
4.3	Conditional Probe Multiplexing	53
4.4	Conditional Probe Mediated Sequence Transduction Triggered by Long RNA	53
B.1	Background-Subtracted Data of Multiplexing with Quenched HCR	66
B.2	Sensitivity Analysis of Quenched HCR	67
B.3	RNA-based Quenched HCR	68
B.4	Fast Target Detection with Quenched HCR	69

List of Tables

A.1	Strands Utilized in Figure 2.1	61
A.2	Strands Utilized in Figure 2.2	61
A.3	Strands Utilized in Figure 2.3	61
A.4	Strands Utilized in Figure 2.4	62
A.5	Strands Utilized in Figure 2.5.	62
A.6	Strands Utilized in Figure 2.6.	63
A.7	Strands Utilized in Figure 2.7	63
A.8	Strands Utilized in Figure 2.8.	63
B.1	Strands Utilized in Chapter 3	70
B.2	Additional Strands Utilized in Appendix B	70
C.1	Strands Utilized in Chapter 4	72

Chapter 1

Introduction

Breakthroughs in the field of nucleic acid engineering over the last two decades include *in vitro* genotyping technologies [1–3], antisense translation inhibition [4–6], and gene therapy [7–9]. Recently, dynamic nucleic acid systems engineering has been demonstrated with the implementation of DNA walkers that traverse space in a programmed manner [10–12], self-assembly reactions, in which predefined structures self-assemble from nucleic acid strands [13, 14], and triggered self-assembly reactions, in which the presence of a specific molecule is required to initiate a molecular self-assembly reaction [11, 15, 16].

To date, few dynamic nucleic acid systems have been developed into useful technologies. Nonetheless, the technologies into which such systems have been integrated, such as molecular-beacon-based real-time PCR, and RNA genotyping in zebrafish [17, 18], suggest that dynamic nucleic acid systems are likely to find increasing application.

Dirks and Pierce introduced the hybridization chain reaction (HCR) mechanism in 2004 [15], demonstrating that metastable nucleic acid molecules can be engineered to self-assemble into polymers in the presence of a nucleic acid trigger, i.e., a target sequence of choice. This isothermal, enzyme-free triggered polymerization technology has found wide use; examples include multiplexed imaging of mRNA expression in zebrafish [18], miRNA detection on a graphene oxide surface [19], and highly sensitive DNA detection via formation of DNAzyme nanowires [20]. This thesis focuses on engineering molecular nucleic acid tools that operate in concert with HCR. These nucleic acid tools selectively detect RNA in bulk, and could potentially be employed in genotyping applications *in situ* and in living cells, as well as in diagnostic applications.

Owing mostly to advances in genomics [21–25], the field of diagnostics has undergone a shift from the macro to the micro. Indeed, clinicians who have formerly had to rely on gross physical changes in patients (e.g., reported symptoms, changes in appearance, cell morphology changes as revealed in biopsies and, more recently, MRI) to diagnose diseases, are transitioning to tools that enable molecular-driven diagnoses [26–31]. Molecular diagnostic tools could revolutionize the practice of medicine by: (1) replacing traditional diagnostic techniques that rely on cell culture and therefore often require weeks [32] with techniques that provide answers within hours or even less [33], (2) reducing the adverse health effects associated with current diagnostic methods [32, 34–37], and (3) minimizing expenditure associated with incorrect diagnosis [38].

The field of molecular diagnostics holds much promise, but it also poses an array of challenges. For a molecular diagnostic technique to find mass appeal, it must be (1) fast, (2) cheap, (3) easy to use, (4) portable, (5) robust, (6) selective (i.e., not yield false positives), (7) sensitive (i.e., not yield false negatives), and (8) suitable to interaction with biological specimens.

A technology that addresses all of these challenges is not presently available. Real-time reverse-transcription PCR (qRT-PCR), for example, relies on an expensive machine and requires technical expertise and clean RNA samples containing neither DNA nor nucleases [39]. Hence, qRT-PCR does not fully satisfy criteria 2–5. Similarly, nucleic acid sequencing technologies are slow, expensive, technically challenging, and unsuitable for point-of-care diagnostics (due to their dependence on large machines). Sequencing technologies, therefore, do not fully satisfy criteria 1–4 [40, 41]. By comparison, the tools described in this thesis (HCR, Scavenger, Quenched HCR, and Conditional Probe) satisfy criteria 1–6 and could potentially satisfy criteria 7–8 as well.

In Chapter 2, we provide two additional contributions to HCR. First, we demonstrate that in addition to being an isothermal, enzyme-free method of nucleic acid detection, HCR is also highly selective. Second, we demonstrate that Scavenger, a single-stranded oligonucleotide, enhances HCR’s selectivity by competitively inhibiting off-target triggering of HCR. We demonstrate that Scavenger improves the selectivity of HCR in a variety of

cases, including the (G→A) substitution, which is the hardest SNP to detect in RNA-RNA hybridization reactions [42, 43]. To the best of our knowledge, selective detection of the G→A SNP, solely via a hybridization-based detection scheme in which both the target and the detector are RNA oligonucleotides, has not been accomplished previously.

In Chapter 3 we demonstrate that Quenched HCR is an effective method of transducing hybridization reactions to light emissions. Because HCR polymer formation is transduced to light, cumbersome gel assays, which have been our primary means of monitoring HCR *in vitro* reactions to date, can be replaced with real-time bulk fluorescence assays [44].

Lastly, in Chapter 4 we describe conditional probes that execute the logical operation: if target sequence A is present, expose target sequence B. In contrast to the transducer of Seelig et al. [45], our Conditional Probe is unimolecular, remains bound to target, and can therefore provide spatial information *in situ* [18] as well as in substrate-based assays. Our conditional probes are conceptually similar to those of Shimron et al. [20]; used in combination with Quenched HCR, we achieve discrimination ratios up to an order of magnitude higher than those previously demonstrated.

Taken together, the tools discussed in this thesis have the potential to form a fast, portable, easy-to-use diagnostic framework. Further, these tools could form the basis for new wash-free methods for expression profiling in bulk samples, in fixed biological samples, and possibly within living cells.

References

- [1] Saiki, R. K. *et al.* Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science (New York, N.Y.)* **239**, 487–91 (1988).
- [2] Higuchi, R., Fockler, C., Dollinger, G. & Watson, R. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Biotechnology Nature Publishing Company* **11**, 1026–30 (1993).
- [3] Lipshutz, R. J., Fodor, S. P., Gingeras, T. R. & Lockhart, D. J. High density synthetic oligonucleotide arrays. *Nature Genetics* **21**, 20–24 (1999).
- [4] Stephenson, M. L. & Zamecnik, P. C. Inhibition of Rous sarcoma viral RNA translation by a specific oligodeoxyribonucleotide. *Proceedings of the National Academy of Sciences of the United States of America* **75**, 285–8 (1978).
- [5] Toth, P. P. Emerging LDL therapies: Mipomersen-antisense oligonucleotide therapy in the management of hypercholesterolemia. *Journal of Clinical Lipidology* **7**, S6–S10 (2013).
- [6] Bhattacharjee, G. *et al.* Inhibition of Vascular Permeability by Antisense-Mediated Inhibition of Plasma Kallikrein and Coagulation Factor 12. *Nucleic Acid Therapeutics* (2013).
- [7] Morgan, R. A. *et al.* Cancer regression in patients after transfer of genetically engineered lymphocytes. *Science (New York, N.Y.)* **314**, 126–9 (2006).
- [8] Bainbridge, J. W. B. *et al.* Effect of gene therapy on visual function in Leber’s congenital amaurosis. *The New England Journal of Medicine* **358**, 2231–9 (2008).
- [9] Cartier, N. *et al.* Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science (New York, N.Y.)* **326**, 818–23 (2009).
- [10] Yin, P., Turberfield, A. J. & Reif, J. H. Designs of autonomous unidirectional walking DNA devices. *DNA Computing* **3384**, 410–425 (2005).

- [11] Yin, P., Choi, H. M. T., Calvert, C. R. & Pierce, N. a. Programming biomolecular self-assembly pathways. *Nature* **451**, 318–22 (2008).
- [12] Omabegho, T., Sha, R. & Seeman, N. C. A bipedal DNA Brownian motor with coordinated legs. *Science* **324**, 67–71 (2009).
- [13] Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006).
- [14] Yin, P. *et al.* Programming DNA tube circumferences. *Science* **321**, 824–826 (2008).
- [15] Dirks, R. M. & Pierce, N. A. Triggered amplification by hybridization chain reaction. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 15275–8 (2004).
- [16] Li, B., Ellington, A. D. & Chen, X. Rational, modular adaptation of enzyme-free DNA circuits to multiple detection methods. *Nucleic Acids Research* **39**, e110 (2011).
- [17] Manganelli, R., Tyagi, S. & Smith, I. Real Time PCR Using Molecular Beacons : A New Tool to Identify Point Mutations and to Analyze Gene Expression in Mycobacterium tuberculosis. *Methods in Molecular Medicine* **54**, 295–310 (2001).
- [18] Choi, H. M. T. *et al.* Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nature Biotechnology* **28**, 1208–1212 (2010).
- [19] Yang, L., Liu, C., Ren, W. & Li, Z. Graphene surface-anchored fluorescence sensor for sensitive detection of microRNA coupled with enzyme-free signal amplification of hybridization chain reaction. *ACS Applied Materials & Interfaces* **4**, 6450–3 (2012).
- [20] Shimron, S., Wang, F., Orbach, R. & Willner, I. Amplified detection of DNA through the enzyme-free autonomous assembly of hemin/G-quadruplex DNzyme nanowires. *Analytical Chemistry* **84**, 1042–8 (2012).
- [21] Adams, M. D. *et al.* Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* **252**, 1651–1656 (1991).

- [22] Venter, J. C. GENOMICS: Shotgun Sequencing of the Human Genome. *Science* **280**, 1540–1542 (1998).
- [23] Ronaghi, M., Uhlén, M. & Nyrén, P. A sequencing method based on real-time pyrophosphate. *Science* **281**, 363, 365 (1998).
- [24] Rothberg, J. M. & Leamon, J. H. The development and impact of 454 sequencing. *Nature Biotechnology* **26**, 1117–24 (2008).
- [25] Li, R., Li, Y., Kristiansen, K. & Wang, J. SOAP: short oligonucleotide alignment program. *Bioinformatics (Oxford, England)* **24**, 713–4 (2008).
- [26] Perez, E. A., Pusztai, L. & van De Vijver, M. Improving patient care through molecular diagnostics. *Seminars in Oncology* **31**, 14–20 (2004).
- [27] Andreeff, M. *et al.* HOX expression patterns identify a common signature for favorable AML. *Leukemia* **22**, 2041–2047 (2008).
- [28] Brizova, H., Kalinova, M., Krskova, L., Mrhalova, M. & Kodet, R. Quantitative measurement of cyclin D1 mRNA, a potent diagnostic tool to separate mantle cell lymphoma from other B-cell lymphoproliferative disorders. *Diagnostic Molecular Pathology* **17**, 39–50 (2008).
- [29] Wang, L. *et al.* Expression of the uptake drug transporter hOCT1 is an important clinical determinant of the response to imatinib in chronic myeloid leukemia. *Clinical Pharmacology and Therapeutics* **83**, 258–64 (2008).
- [30] Tiveljung-Lindell, A. *et al.* Development and implementation of a molecular diagnostic platform for daily rapid detection of 15 respiratory viruses. *Journal of Medical Virology* **81**, 167–175 (2009).
- [31] Dawson, S.-J. *et al.* Analysis of Circulating Tumor DNA to Monitor Metastatic Breast Cancer. *New England Journal of Medicine* **368**, 130313140010009 (2013).
- [32] McNerney, R. *et al.* Tuberculosis diagnostics and biomarkers: needs, challenges, recent

- advances, and opportunities. *The Journal of Infectious Diseases* **205 Suppl**, S147–58 (2012).
- [33] Weile, J. & Knabbe, C. Current applications and future trends of molecular diagnostics in clinical bacteriology. *Analytical and Bioanalytical Chemistry* **394**, 731–742 (2009).
- [34] Watson, R. L. *et al.* Antimicrobial Use for Pediatric Upper Respiratory Infections: Reported Practice, Actual Practice, and Parent Beliefs. *PEDIATRICS* **104**, 1251–1257 (1999).
- [35] Bazalová, Z., Benes, J. & Bronská, E. [A study of antibiotic treatment of upper respiratory infections in primary care]. *Klinická Mikrobiologie a Infekční Lékarství* **14**, 24–9 (2008).
- [36] Alonso-Martínez, J. L., Sánchez, F. J. A. & Echezarreta, M. a. U. Delay and misdiagnosis in sub-massive and non-massive acute pulmonary embolism. *European Journal of Internal Medicine* **21**, 278–82 (2010).
- [37] Newman-Toker, D. E. & Pronovost, P. J. Diagnostic errors—the next frontier for patient safety. *JAMA : The Journal of the American Medical Association* **301**, 1060–2 (2009).
- [38] Vakil, N., Rhew, D., Soll, A. & Ofman, J. J. The cost-effectiveness of diagnostic testing strategies for *Helicobacter pylori*. *The American Journal of Gastroenterology* **95**, 1691–8 (2000).
- [39] Bustin, S. A. & Nolan, T. Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *Journal of Biomolecular Techniques : JBT* **15**, 155–66 (2004).
- [40] Metzker, M. L. Sequencing technologies - the next generation. *Nature Reviews Genetics* **11**, 31–46 (2010).
- [41] Kircher, M. & Kelso, J. High-throughput DNA sequencing—concepts and limitations. *BioEssays* **32**, 524–36 (2010).
- [42] Serra, M. J. & Turner, D. H. Predicting thermodynamic properties of RNA. *Methods in Enzymology* **259**, 242–61 (1995).

- [43] Mathews, D. H., Sabina, J., Zuker, M. & Turner, D. H. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *Journal of Molecular Biology* **288**, 911–40 (1999).
- [44] Chang, B.-Y. Smartphone-based Chemistry Instrumentation: Digitization of Colorimetric Measurements. *Bulletin of the Korean Chemical Society* **33**, 549–552 (2012).
- [45] Seelig, G., Soloveichik, D., Zhang, D. Y. & Winfree, E. Enzyme-free nucleic acid logic circuits. *Science* **314**, 1585–8 (2006).

Chapter 2

Isothermal Enzyme-Free Detection of SNPs via Hybridization Chain Reaction

2.1 Introduction

Single nucleotide polymorphisms (SNPs) are associated with cancer [1–5], meningitis [6, 7], hepatitis C [8], and a variety of other diseases. SNPs confer antibiotic resistance [9, 10] and can be used as pharmacogenetic [11–14] and forensic markers [15]. Owing to their importance as disease markers, a plethora of techniques for SNP detection have been developed [16–20]. Yet, due to limitations in speed, cost, ease of use, portability, robustness, selectivity, or sensitivity, SNP detection technologies have mostly not been integrated into the clinic [23], with the exception of real-time PCR, which has recently found minimal use at the clinic [21, 22].

High-throughput, next-generation sequencing technologies [24], for example, rely on large, expensive instrumentation whose use in a point-of-care setup is not currently practical. Similarly, microarray data analysis requires a large machine and relies on costly chips that render this technique cost-prohibitive for non-high-throughput applications. Moreover, microarrays have limited selectivity, as they often contain probes that hybridize to multiple genes [25], and limited sensitivity, as they do not reliably detect low-abundance genes [25, 26].

In contrast to microarrays, real-time reverse-transcription PCR (qRT-PCR) reliably dis-

criminate between SNPs [17], and can detect single copy mRNAs [27]. For these reasons, the use of qRT-PCR as a clinical diagnostic has increased in the last decade. At the same time, qRT-PCR has some shortcomings. First, its high sensitivity can be a hindrance, as even trace amounts of genetic DNA are amplified, thus providing false positives [28]. Similarly, sample impurities, as well as amplicon and target secondary structure can result in false negatives [29, 30]. Moreover, to detect and quantify mRNA, qRT-PCR relies on a proxy (cDNA) that is generated by Reverse Transcriptase. The use of this enzyme introduces three additional limitations to qRT-PCR: (1) due to variance in secondary structures in various RNA targets as well as in PCR primers, the formation of cDNA from an RNA template does not follow a fixed mathematical transformation, thus rendering qRT-PCR non-quantitative; (2) qRT-PCR is limited to locations and conditions that ensure the stability of Reverse Transcriptase; and (3) the optical signal created during qRT-PCR is physically separated from RNA, such that this technique offers no spatial information concerning the transcript whose presence and quantity the technique gauges. Lastly, qRT-PCR relies on a thermocycler, which in turn limits both the speed of the method and the use of this method in point-of-care setups.

Another technique commonly used to detect SNPs involves the use of molecular beacons [31]. While molecular beacons selectively detect SNPs [16, 17], their sensitivity is limited, as the maximum ratio of beacons that can be hybridized to targets is 1:1.

A technique that was demonstrated to achieve both high specificity and high sensitivity utilizes nicking enzymes in order to obtain signal amplification (NESA) [32]. However, to achieve detection independently of target sequence, the recognition site of the nicking enzyme is introduced to the target sequence via Rolling Circle Amplification. This, in turn, requires DNA ligase, DNA polymerase, and a heat block, in addition to a nicking enzyme and other PCR reagents. The use of multiple enzymes limits this detection technique both in buffer constitution and in permissible temperatures. Additionally, this detection scheme can only be employed where the enzymes' stability is maintained.

This chapter assesses the merits of SNP detection via Hybridization Chain Reaction (HCR) [33] as an alternative to, or supplemental tool to be used with, the aforementioned

techniques.

2.2 Detection of SNPs with Hybridization Chain Reaction

In a Hybridization Chain Reaction (HCR) [33], nucleic acid hairpins (h1 and h2) undergo isothermal, non-enzymatic self-assembly into polymers in the presence of target (T), a nucleic acid sequence with which the hairpins are designed to hybridize (Figure 2.1). HCR is initiated by a toehold-mediated branch migration of h1 in the presence of T, and it proceeds by an alternating addition of h2 and h1 to the living end of polymer (Figure 2.1, panel A). Due to the high energy of an h1·h2 intermediate, both hairpin species maintain their metastable hairpin conformation in the absence of T [34].

HCR is therefore characterized by polymers in the presence of T and monomers in the absence of T. While this duality of states was demonstrated by Dirks and Pierce in 2004 [33], neither the selectivity of HCR nor kinetic control over the rate of HCR polymerization were previously addressed. Our studies focused on the detection of single nucleotide polymorphisms (SNPs), so as to demonstrate HCR selectivity in the presence of mutations that are expected to be the hardest to detect, due to small $\Delta\Delta G$ values between the detector-target energetics and the detector-off-target energetics [35]. As demonstrated in this chapter, the detection of SNPs requires fine control over the kinetics of HCR.

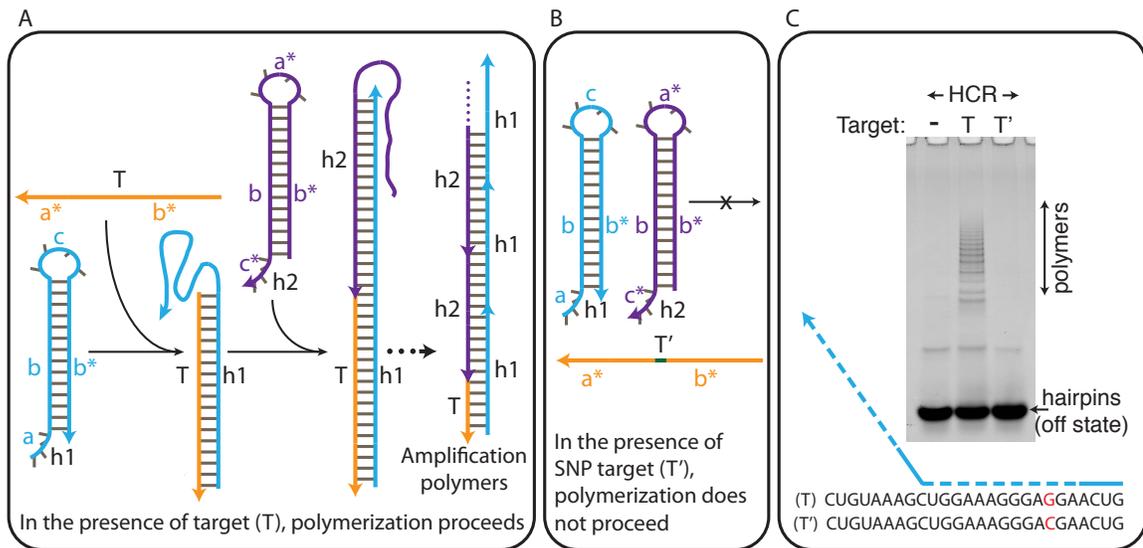


Figure 2.1: SNP detection. (A) In the presence of an oligonucleotide (T) that is complementary to the toehold and half of the stem of hairpin 1 (h1), the metastable hairpins form polymers. (B) In the presence of target (T') that has a single mismatch (illustrated in green) with respect to T, the hairpins retain their closed structure. (C) Gel electrophoresis of HCR-mediated SNP detection. The hairpins polymerize in the presence of T (lane 2), but remain closed in the absence of T, as well as in the presence of T'. h1 is illustrated in solid lines (toehold and loop) and dashed lines (stem), and is shown in duplex with T along the sequence window at which these two strands are expected to hybridize.

2.3 Kinetic Control of HCR via ΔG Tuning

Because the nucleic acids utilized in HCR are metastable, the fraction of hairpins predicted to form polymers at thermodynamic equilibrium is essentially independent of target (either cognate or SNP) [34]. For this reason, HCR-mediated SNP discrimination at thermodynamic equilibrium is not predicted to be feasible. Hence, we proceed to check the feasibility of pre-equilibrium SNP detection.

We recall that HCR is initiated by the formation of either a $T \cdot h1$ or a $T' \cdot h1$ duplex. Once a duplex is formed, however, the rate at which hairpins are incorporated into living polymers in the subsequent steps of HCR in a T -containing test tube is about equal to the rate at which hairpins are incorporated into polymers in a T' -containing test tube. Hence, the rate of formation of $T \cdot h1$ and $T' \cdot h1$ at short times, i.e., while the majority of T' remains undetected, is expected to govern the duration of time during which SNP discrimination is expected to be near its optimum¹. Once this time frame has passed, a non-marginal fraction of the T' targets have turned into $T' \cdot h1$ duplexes, and conversion (the fraction of hairpins that are assembled into polymers out of total hairpins) in a T' -test tube begins to approach the conversion that occurs in a T -test tube. In other words, discrimination is lost as test tubes transition from the short time scale that characterizes T polymerization to the long time scale that characterizes T' polymerization.

Even though SNP discrimination is a transient phenomenon, we can control when SNP discrimination begins to appear (as HCR polymers in a T -containing test tube generate signal above background), and when discrimination is lost (as T' polymerization approaches T polymerization). In practice, polymerization in the absence of targets also contributes to loss of discrimination, however, target-free polymerization occurs in a longer time scale than that of T' -triggered polymerization. Controlling HCR kinetics is achieved via a principle that we refer to as ΔG tuning.

In 2009 Zhang et al. demonstrated that for short toeholds, the rate of toehold-mediated branch migration varies exponentially with the ΔG of toehold hybridization: $k \sim e^{-\frac{\Delta G}{RT}}$ [36]. Accordingly, we employ the engineering principle of ΔG tuning as a means of controlling the

¹The energetics of $T \cdot h1$ formation are predicted to have less favorable energetics than all subsequent hairpin addition steps in HCR, because the initial step yields one less base-stack than subsequent steps.

rate of formation of HCR polymers; the more favorable the ΔG per polymerization step, the faster HCR proceeds, and the earlier SNP discrimination is both obtained, as T-containing test tube polymers become detectable, and lost, as T'-containing test tube polymers are formed. The selectivity of SNP discrimination is expected to depend on $\Delta\Delta G$: the difference in free energy between T·h1 formation and T'·h1 formation. The energetics of addition of h1 to polymers that end with h2 depends on the energetics of h1 addition to T, since the main difference between these hybridization reactions is the formation of one extra base-stack in the former. Hence, tuning the energetics of h1 addition to T, results in a similar tuning of h1 addition to polymers that end with h2. Our ΔG tuning approach, therefore, focused on the free-energy of addition of h1 to T, and of h2 to T·h1. On a molecular level, ΔG tuning is obtained by controlling the nucleotide make-up of the toeholds of the two hairpins and/or their lengths (Figure 2.2).

2.4 Experimental Verification of Kinetic Control of HCR

To validate SNP detection via HCR, three different HCR systems were designed [37] according to the ΔG tuning principle, and their conversion as a function of time was monitored in the presence of T, T', and T'' (perfect target, a target that differs from T by an SNP, and a target that differs from T by two nucleotide substitutions, respectively). The HCR systems were designed to have characteristic time constants of discrimination of ~ 10 hours, ~ 1 hour and a few minutes (and are thus labeled slow, medium, and fast, respectively) in test tubes containing $1\mu\text{M}$ of each molecule. As the data presented in Figure 2.2 shows, ΔG tuning of hairpin toeholds appears to be a viable way of controlling the characteristic time constant of discrimination. Our results indicate that HCR systems can be designed to discriminate between T and T' at a time scale of choice.

In addition to discriminating SNPs in a time scale of choice, the fast system is expected to lose its discrimination between T and T'' faster than the medium speed system, which, in turn, is expected to lose its discrimination between T and T'' faster than the slow system. This expectation is supported by the data in Figure 2.2.

Further, the extent to which hairpins maintain their metastable structure (i.e., avoid polymerization) in the absence of target is assumed to depend on the ΔG of addition of the toehold of h1 to its reverse complement (loop of h2); the more negative the ΔG of hybridization between h1's toehold and h2's loop, the faster the rate of an h1-h2 formation, and the more target-free polymers are observed. As can be seen in lane 16 of the bottom panel of Figure 2.2, the data is in agreement with this hypothesis as well.

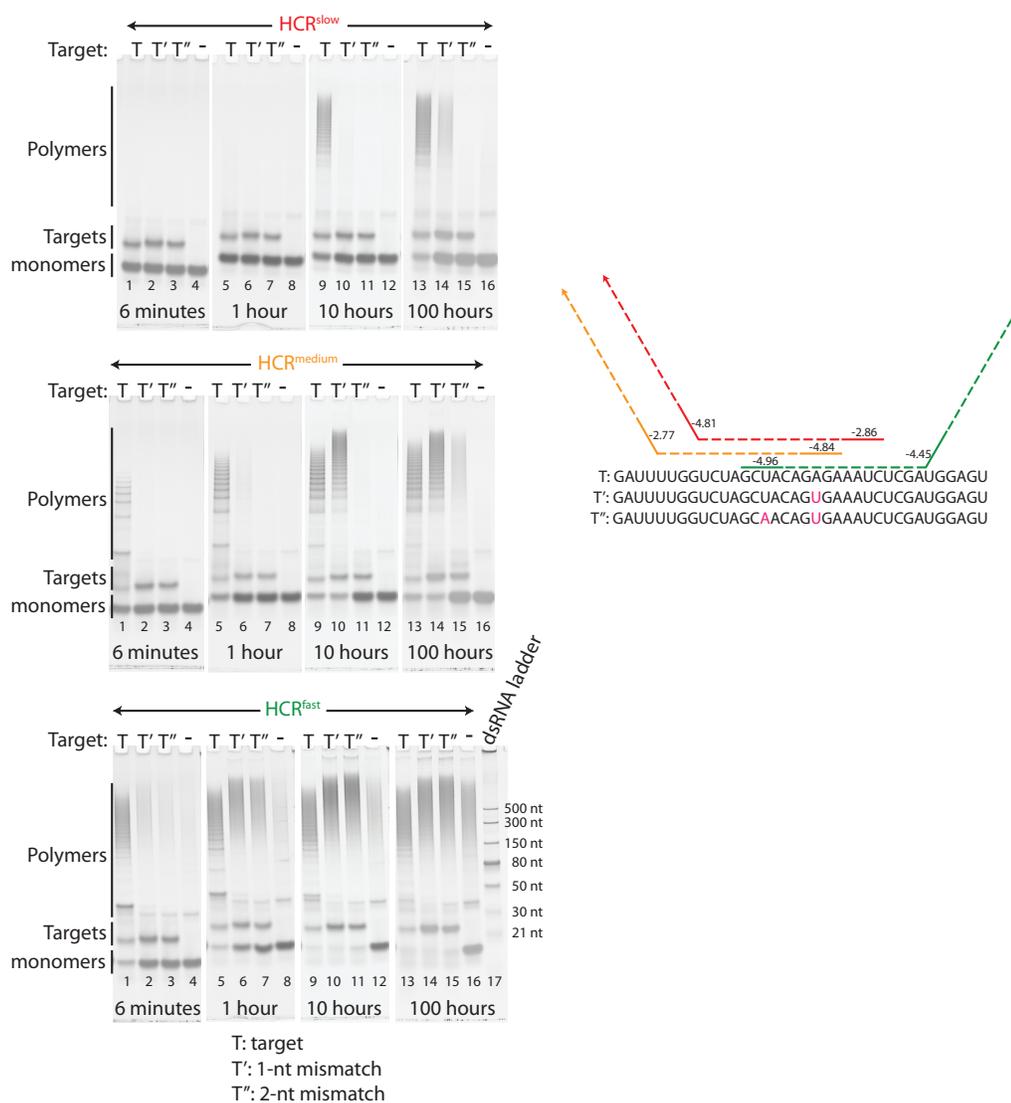


Figure 2.2: Kinetic discrimination of SNPs via HCR at a time scale of choice. The conversion of three HCR systems (labeled slow, medium, and fast, denoting systems that were designed to optimize SNP detection at the time scale of ~ 10 hours, ~ 1 hour, and minutes, respectively) is monitored as a function of time. The fraction of hairpins that are assembled into polymers increases as a function of time with either T, T', or T'' (a target that complements h1, an off-target that has an SNP when hybridized to h1, and an off-target that has two substitutions when hybridized to h1), as well as in a target-free manner. The right panel illustrates the target-nucleotides with which h1 of each of the three hairpins systems hybridizes. The red system is HCR^{slow} ; the toehold and loop of its h1 constituent are illustrated in solid red and its stem portion is illustrated with a dashed red line. $\text{HCR}^{\text{medium}}$ and HCR^{fast} appear in orange and green, respectively. The numbers that appear next to the toehold and loop domains of each of the three systems' h1, are NUPACK [37–40]-calculated ΔG values (in units of kcal/mol) of duplex formation between the 4-nt toehold sequence and its reverse complement, and between the 4-nt loop sequence and its reverse complement.

2.5 Effect of Mutation Location on SNP Discrimination

In the previous section, we verified that HCR discriminates between SNPs pre-equilibrium. Next, we propose a set of experiments to address the question of where an HCR system should hybridize with a T/T' target pair to achieve an optimal discrimination ratio. We recall that HCR is initiated with a toehold-mediated branch migration of h1 in the presence of T or some off-target. Since branch migration is initiated at the toehold of h1 and is terminated in the stem portion of h1 that is closest to h1's loop, the location along h1 at which the branch has to migrate through an SNP could possibly affect the discrimination ratio that the HCR detector achieves.

The following set of experiments was conducted in order to address this question. We designed two seed sequences lacking secondary structure² and containing nucleotide triplicate sequences that repeat themselves four times in the target set (Figure 2.3). Both of these seed sequences were mutated to eight daughter sequences that lack secondary structure and vary from their respective parent seed sequence by an SNP. The SNP type, e.g., C→G in panel A, Figure 2.3, as well as the nearest neighbors [41] of each of the SNP loci, were conserved in four data sets that were produced from two seed sequences and sixteen daughter sequences. We then analyzed conversion as a function of mutation location at different time points.

The results of this study (Figure 2.3) suggest that SNPs located in the first half (with respect to toehold) of the stem of h1 provide discrimination ratios that are as good as, or better than, those obtained by SNPs located at either the toehold of h1 or the second half of the stem of h1. Hence, we tend to favor positioning the mutation in the first half of the stem (relative to the toehold) in an effort to increase the initial chances of success when designing new highly selective HCR systems.

²Seed sequences as well as mutated sequences were designed to be secondary-structure free so as to not skew the results of this study in favor of mutations that reside in locations along the target-set that are rich in secondary structure, as such locations are expected to form kinetic traps to the branch migration that h1 undergoes.

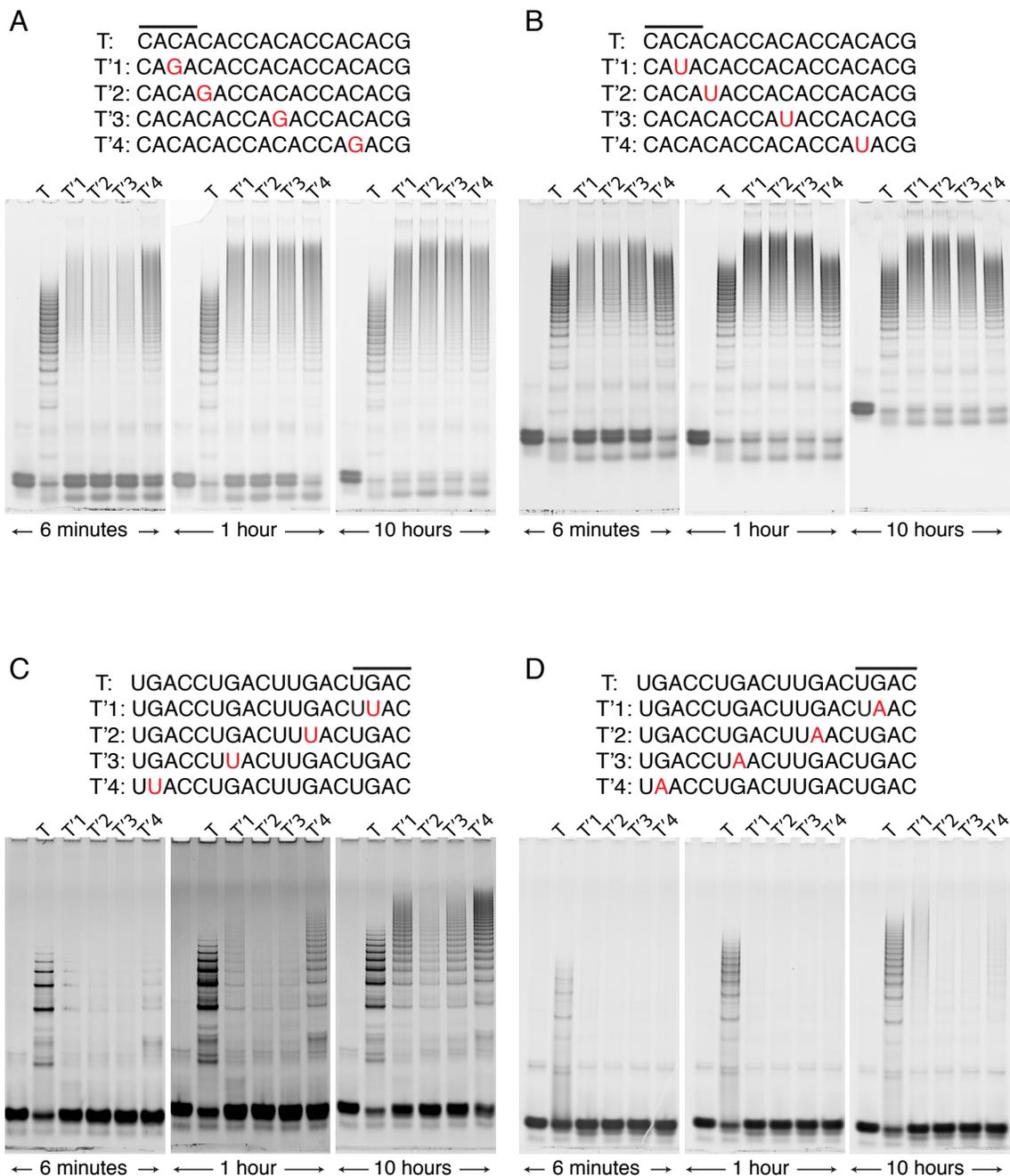


Figure 2.3: Effect of mutation location on SNP discrimination. Seeking to determine the location(s) that optimize discrimination, HCR systems were fixed (one system was incubated with the targets that appear in panels A and B, and a different HCR system was incubated with the targets that appear in panels C and D); the SNP type, e.g., C→G in panel A, as well as the nearest neighbors of the SNP were, likewise, kept fixed in each of the four panels. The target sequences with which h1 toeholds hybridize are denoted in black lines. SNPs located in the first half (with respect to toehold) of the stem of h1 provide discrimination ratios that are as good as, or better, than those obtained by SNPs located in either the toehold of h1 or the second half of the stem of h1. For example, the data collected from T'4 in panel A reveals that a C→G mutation with nearest neighbors A/A should not be placed too near the loop of h1.

2.6 Detection of SNP Cancer Markers

We now turn to the challenge of obtaining high levels of discrimination for an arbitrary SNP after the passage of a time period of choice, focusing on SNP cancer markers. To this end, we engineered HCR systems to detect three common SNP cancer markers, without placing any requirements on either the sequences of these targets or their secondary structures. Specifically, HCR systems were designed to detect the pervasive cancer markers BRAF 1799T→A [42–48]³, JAK2 1849G→T [2, 49, 50]⁴ and PTEN 388C→G [51, 52]⁵. Further, to demonstrate that high discrimination ratios can be obtained at a time scale of choice in a hairpin concentration of choice, we engineered the HCR systems to reach high levels of discrimination after a 1-hour incubation period at 37°C of 1 μ M hairpins with 1 μ M targets.

Three phenomena are highlighted in the results of this set of experiments (Figure 2.4). First, high selectivity is obtained for each of the three cancer markers. Second, false negatives are avoided by generating two distinct fluorescent signals for the wild-type and mutant sequences. Third, the use of spectrally distinct fluorophores facilitates the use of multiple HCR systems in a single test tube.

³These publications suggest that ~50% of melanoma patients carry a mutation in their BRAF gene, and among these BRAF mutations, ~80-90% are the BRAF 1799T→A single nucleotide polymorphism.

⁴These publications suggest that above ~90% of polycythemia vera patients, ~50% of essential thrombocythemia patients, and ~50% of primary myelofibrosis patients have the JAK2 1849G→T mutation. In some occasions these disorders develop into malignant cancers.

⁵These publications suggest that ~20% of ovarian cancer patients have some mutation in their PTEN gene, and among these mutations, ~6% are PTEN 388C→G.

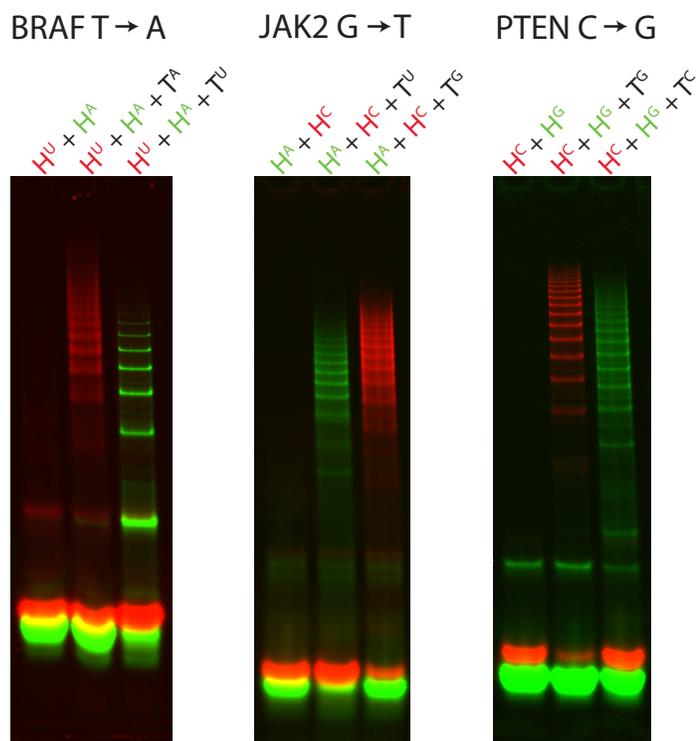


Figure 2.4: Detection of SNP cancer markers. For each of three RNA cancer markers, SNP detection is obtained with two hairpin systems; one is designed to polymerize in the presence of the mutant but not in the wild-type form of the gene, and the other is designed to polymerize in the presence of the wild-type version of the gene but not in the presence of the mutant form of the gene. We denote HCR systems with a superscript that defines the nucleotide that h1 contains at the SNP locus. For example, to detect the cancer marker in the BRAF T→A SNP (T^A), H^U is utilized. High selectivity is obtained in each of the three cancer markers. False negatives are avoided by generating two distinct fluorescent signals, one for the wild-type marker of the gene, and another for the cancer marker. Lastly, 2-color multiplexing is successfully achieved.

2.7 Improving HCR Discrimination with Scavenger

The discrimination ratios obtained with HCR-mediated SNP detection depend on $\Delta\Delta G$ of the first step of HCR (the free-energy difference between $T \cdot h1$ and $T' \cdot h1$). We therefore expect that HCR will provide poor discrimination ratios in the presence of SNPs that provide small $\Delta\Delta G$ values. Among such SNPs, the hardest to detect in RNA-RNA hybridization is $G \rightarrow A$ because the detector-target ($h1^U \cdot T^A$) energetics are almost isoenergetic to the detector-off-target ($h1^U \cdot T^G$) energetics [41, 53], thereby leading to poor discrimination ratios in the absence of Scavenger (Figure 2.5, compare lanes 2 and 3).

We wish to establish HCR as a tool that obtains high discrimination ratios for all possible SNPs. To achieve this goal, we introduce Scavenger as a competitive inhibitor of the nearly isoenergetic off-target. Scavenger is a single-stranded oligonucleotide that hybridizes with T' and has an SNP when hybridized to T (Figure 2.5, panels A and B, respectively). The energetics and concentration of Scavenger are tuned such that it preferentially forms a duplex with T' ($T' \cdot S$) relatively to duplex $T \cdot S$. Hence, when Scavenger and HCR are mixed together, polymerization with T' is impeded relative to polymerization with T .

This conceptual approach is especially compelling in the case of the most-challenging SNP: $G \rightarrow A$. For this case, H^U has poor selectivity when $T \equiv T^A$ and $T' \equiv T^G$, because $h1^U \cdot T^G$ is nearly isoenergetic to $h1^U \cdot T^A$ thereby yielding a small $\Delta\Delta G$. By contrast, Scavenger (S^C) binds strongly to T^G but rejects T^A , yielding a typical SNP $\Delta\Delta G$. Hence, crucially, Scavenger can be selective for T' even though HCR is not selective for T , thereby yielding HCR selectivity via competitive inhibition.

To demonstrate the utility of Scavenger, we begin with the difficult $G \rightarrow A$ substitution. HCR provides little to no selectivity for the $G \rightarrow A$ substitution (Figure 2.5, compare lanes 2–3). When the same HCR system is incubated with Scavenger, however, high discrimination is obtained (Figure 2.5, lanes 4–5). To the best of our knowledge, highly selective detection of the $G \rightarrow A$ SNP in an RNA-RNA hybridization assay has not been previously demonstrated.

To establish the generality of Scavenger, we sought to demonstrate its utility in improving discrimination with additional SNP constitutions. To this end, HCR systems designed

G→A SNP

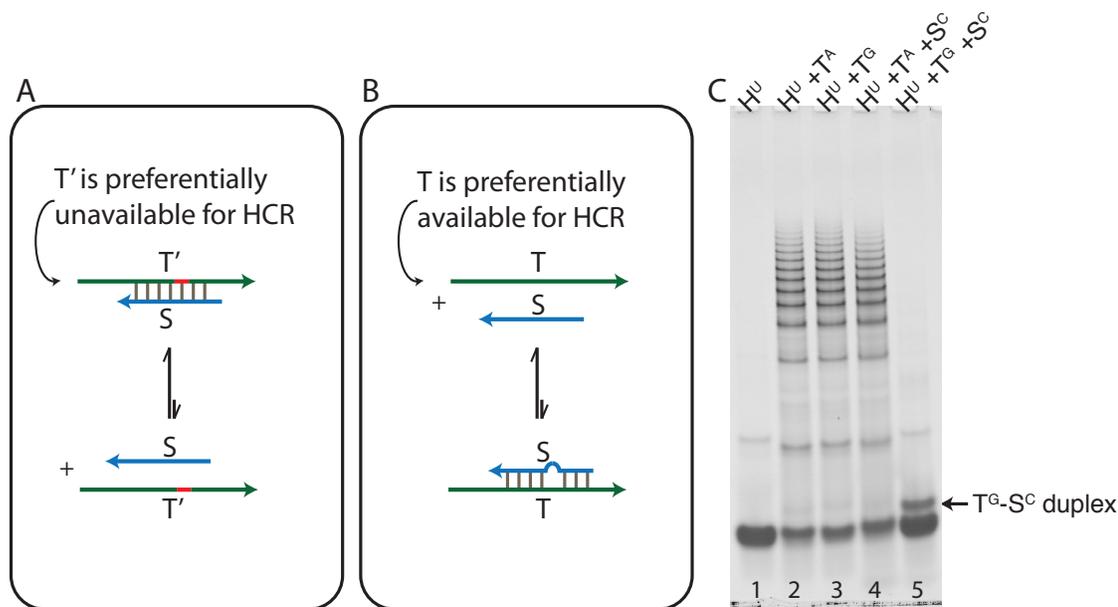


Figure 2.5: Scavenger- and HCR-mediated SNP detection. (A, B) Scavenger is a single-stranded oligonucleotide that complements T' and has an SNP with T. The energetics of Scavenger are chosen such that Scavenger equilibrium tends toward duplex formation in the presence of T' , but Scavenger remains mostly unhybridized in the presence of T. (C) The hardest SNP to detect in RNA-RNA hybridization is G→A, since the detector-target ($H^U \cdot T^A$) energetics are almost isoenergetic to the energetics of the detector-off-target ($H^U \cdot T^G$), thereby leading to diminished discrimination ratios in the absence of Scavenger (lanes 2 and 3, respectively). The presence of Scavenger dramatically improves discrimination (lanes 4 and 5).

to detect the BRAF T→A, JAK2 G→T and PTEN C→G cancer markers were tested in the presence of their respective targets, i.e., the mutated states of the genes, as well as in the presence of all three other nucleotide variants in the naturally occurring SNP site. Incubating each of six different HCR systems with four different target sequences (lanes 2–5 in Figure 2.6) yielded five additional cases (denoted by red numbers in Figure 2.6) in which we sought to improve discrimination. Scavenger was successfully employed in all of these cases (green numbers in Figure 2.6).

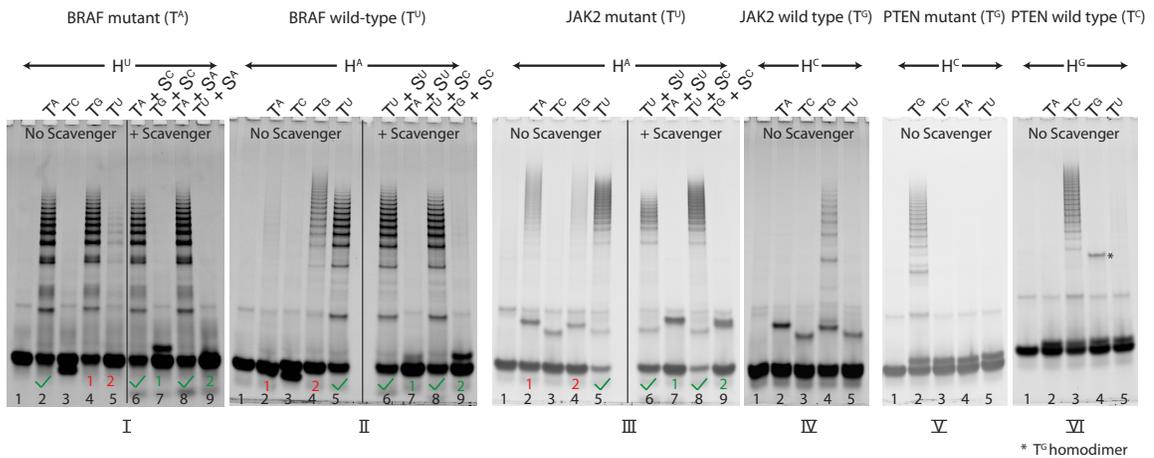


Figure 2.6: Generality of Scavenger- and HCR-mediated selectivity. HCR systems designed to detect the mutated forms of BRAF T→A, JAK2 G→T and PTEN C→G were incubated in the presence of their cognate targets, as well as in the presence of all other three SNP variants at the cancer mutation position. In all cases (marked in red numbers) where a selectivity improvement in the performance of HCR was sought, an appropriate Scavenger was designed and successfully implemented (lanes 6–9 in panels I–III). Discrimination improvement in the presence of Scavenger requires that undesired polymerization (numbered, in red) is reduced to a minimum, and desired polymerization (denoted in green checkmarks) is maintained. Amongst the six cases in which discrimination was improved in this figure, five are novel, since panel I lane 4 was demonstrated in Figure 2.5.

2.8 SNP Profiling via HCR Multiplexing

In Section 2.6 we demonstrated that multiplexing with two HCR systems in a single test tube is feasible. Since the ability to multiplex more than two genes at once is highly useful in a variety of applications including pathogenic bacteria genotyping [9, 16] and pathogenic virus genotyping [54], we sought to expand 2-color multiplexing to 4-color multiplexing. Accordingly, we proceed to demonstrate highly selective SNP profiling of four target sequences that vary from each other by an SNP (Figure 2.7). To this end, four HCR systems, each labeled with a different fluorophore, were mixed in a single test tube. In each of the four target sequences that were tested (Figure 2.7), only one HCR system formed the majority of the polymers observed in the test tube. Thus, an arbitrary nucleotide identity was selectively determined using four HCR systems in a single test tube. False negatives were avoided by generating a different signal for each of the four possible nucleotide identities.

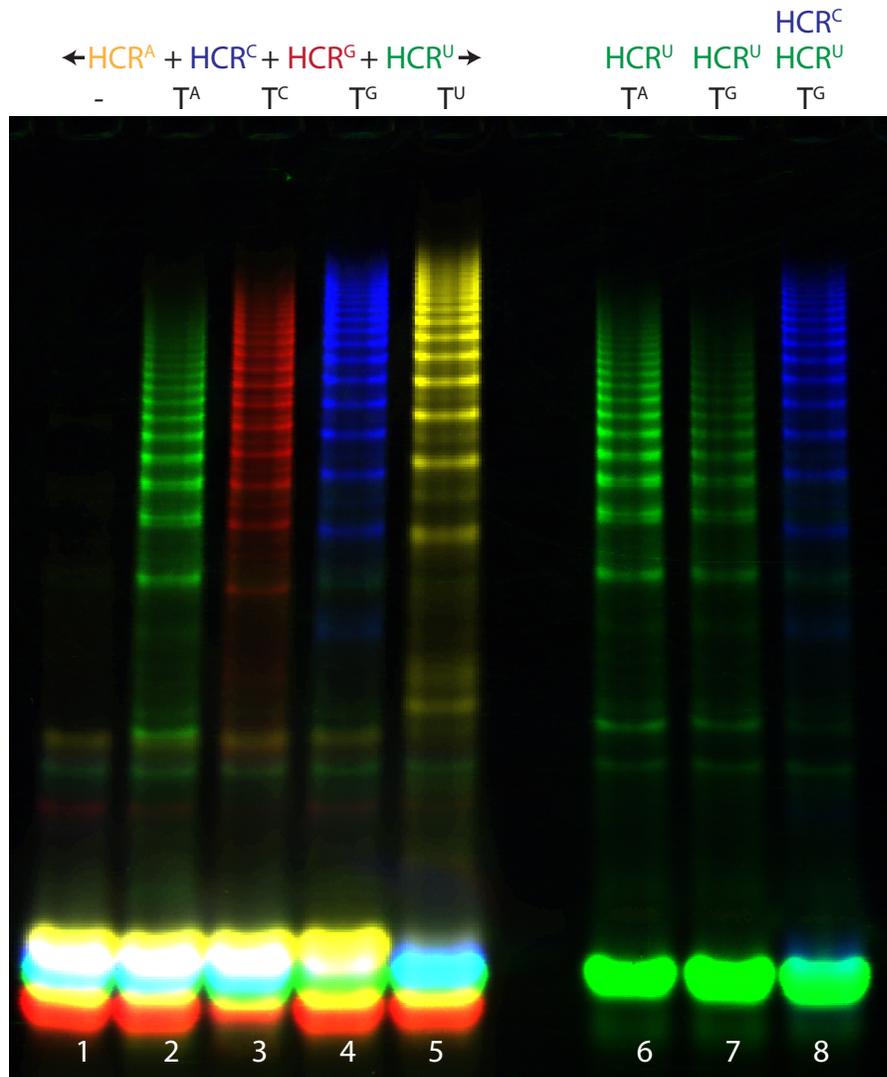
In order to obtain optimal results from SNP profiling, i.e., reduce incorrect signal as much as possible, the kinetic behavior of the various HCR systems ought to be controlled by tuning either ΔG (Section 2.3) or the concentration of the various HCR systems. Since a concentration change is both easier to implement and more economical than a redesign of a fluorescently labeled HCR system, we adopt that approach preferentially when tuning multiple HCR systems to work at the same time point. Using the four HCR systems that are presented in Figure 2.7, we found it useful to increase the concentration of H^C and H^G $2\times$ with respect to H^A and H^U . We increased the concentration of H^C so as to inhibit H^U from forming polymers with T^G via G·U wobble pair formation. We increased the concentration of H^G as it was relatively slow to polymerize compared to the other HCR systems. A further requirement of HCR multiplexing is that the HCR systems not share too much sequence space with each other, since this leads to mixed, multi-fluorophore-containing, polymers. Figure 2.7 depicts the sequence windows that were detected by each of the four HCR systems.

Lastly, we note that the use of multiple HCR systems in a single test tube results in a situation in which the HCR system that hybridizes with the target variant most readily (due to energetics or concentration advantage) scavenges the target and, hence, prevents

other HCR systems from hybridizing with the target. As a result, test tubes containing multiple HCR systems can exhibit higher discrimination ratios than test tubes containing a single HCR system (lanes 4 and 8 vs. 7, Figure 2.7). Multiplexing, therefore, provides a form of selective target scavenging, which results in increased selectivity.

2.9 Detection of SNPs in Long RNA

Highly selective HCR was experimentally demonstrated with a variety of short RNA targets throughout this chapter. While short RNA provides a proof-of-concept that HCR selectively discriminates between SNPs, it is desirable that similar selectivity be demonstrated with long RNA targets which may be found in complex biological samples such as blood and saliva. To demonstrate SNP detection in the presence of long RNA, an 872-nt transcript, consisting primarily of the d2EGFP sequence, as well as a C→G SNP-variant of this transcript, were *in vitro* transcribed. Following purification of the transcription reactions, each of the two transcripts was incubated with a Cy3-modified HCR system that is designed to detect the original transcript, as well as with a Cy5-modified HCR system that is designed to detect the mutated variant of the transcript. As is shown in Figure 2.8, SNP detection was successfully demonstrated with these two long RNA strands.



T^A : CUGUAAAGCUGGAAAGGGAAAGAACUGGUGUAAUGAU
 T^C : CUGUAAAGCUGGAAAGGGACGAAACUGGUGUAAUGAU
 T^G : CUGUAAAGCUGGAAAGGGAGGAACUGGUGUAAUGAU
 T^U : CUGUAAAGCUGGAAAGGGGAUGAACUGGUGUAAUGAU

Figure 2.7: SNP profiling via HCR multiplexing. Each of four distinct targets triggers the polymer formation of its complementary HCR system but not of other HCR systems. The undesired polymerization that occurs in an $H^U + T^G$ test tube (lane 7) is reduced in an $H^U + H^C + T^G$ test tube, because the target T^G is scavenged by the HCR system with which T^G is designed to interact (H^C , lane 8). The sequence variance between the four targets that were used in this study is marked in orange. The sequence window with which each of the four HCR systems is designed to hybridize is marked with a solid line (toehold of h1) and a dashed line (stem of h1).



Figure 2.8: Detection of a C→G SNP in long RNA. A transcript of 872-nt consisting primarily of the sequence of d2EGFP gene, as well as a C→G SNP variant of it were incubated with a Cy3-modified HCR system that is designed to detect the unmodified transcript, and a Cy5-modified HCR system that is designed to detect the SNP variant of the transcript. High selectivity was obtained for both target variants.

2.10 Conclusion

We demonstrated that HCR can selectively detect single nucleotide polymorphisms. Further, we demonstrated that for the most challenging SNPs, HCR yields increased discrimination ratios with the use of either Scavenger strands or HCR multiplexing. Throughout the studies we conducted, all SNP targets we tested were selectively discriminated from their non-SNP-containing analogues. We therefore expect that HCR can be used as a selective SNP detector for all SNPs. We have also shown that SNPs can be detected quickly; a time frame of 6 minutes was demonstrated in Figure 2.2, and we presume that detection in the time frame of seconds is achievable as well. While we have not been able to establish SNP detection in target concentrations that fall below $\sim 50\text{nM}$, we expect that supplementary technologies and instrumentation will improve this sensitivity limit. Overall, HCR provides highly-selective, isothermal, enzyme-free, rapid SNP detection—characteristics that establish HCR as a promising candidate for molecular diagnostic and genotyping applications.

References

- [1] Cantwell-Dorris, E. R., O’Leary, J. J. & Sheils, O. M. BRAFV600E: implications for carcinogenesis and molecular therapy. *Molecular Cancer Therapeutics* **10**, 385–94 (2011).
- [2] Jelinek, J. *et al.* JAK2 mutation 1849G>T is rare in acute leukemias but can be found in CMML, Philadelphia chromosome-negative CML, and megakaryocytic leukemia. *Blood* **106**, 3370–3 (2005).
- [3] Kato, H. *et al.* Functional evaluation of p53 and PTEN gene mutations in gliomas. *Clinical Cancer Research* **6**, 3937–3943 (2000).
- [4] Ghossaini, M. *et al.* Multiple loci with different cancer specificities within the 8q24 gene desert. *Journal Of The National Cancer Institute* **100**, 962–966 (2008).
- [5] Greenman, C. *et al.* Patterns of somatic mutation in human cancer genomes. *Nature* **446**, 153–8 (2007).
- [6] Katz, L. S. *et al.* Using single-nucleotide polymorphisms to discriminate disease-associated from carried genomes of *Neisseria meningitidis*. *Journal of Bacteriology* **193**, 3633–41 (2011).
- [7] da Silva, T. A. *et al.* SNPs in DNA repair genes associated to meningitis and host immune response. *Mutation Research* **713**, 39–47 (2011).
- [8] Lin, C.-Y. *et al.* IL28B SNP rs12979860 is a critical predictor for on-treatment and sustained virologic response in patients with hepatitis C virus genotype-1 infection. *PloS one* **6**, e18322 (2011).
- [9] Blaschitz, M. *et al.* Real-time PCR for single-nucleotide polymorphism detection in the 16S rRNA gene as an indicator for extensive drug resistance in *Mycobacterium tuberculosis*. *The Journal of Antimicrobial Chemotherapy* 1–4 (2011).

- [10] Song, J. *et al.* Simultaneous pathogen detection and antibiotic resistance characterization using SNP-based multiplexed oligonucleotide ligation-PCR (MOL-PCR). *Advances in Experimental Medicine and Biology* **680**, 455–64 (2010).
- [11] Chapman, P. B. *et al.* Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *The New England Journal of Medicine* **364**, 2507–2516 (2011).
- [12] Health, C. f. D. & Radiological. Recently-Approved Devices - cobas 4800 BRAF V600 Mutation Test - P110020.
- [13] Laing, R. E., Hess, P., Shen, Y., Wang, J. & Hu, S. X. The role and impact of SNPs in pharmacogenomics and personalized medicine. *Current Drug Metabolism* **12**, 460–86 (2011).
- [14] van Ooij, C. Pharmacogenetics: A SNP for hepatitis C treatment failure. *Nature Reviews Genetics* **10**, 738–739 (2009).
- [15] Planz, J. V. *et al.* Enhancing resolution and statistical power by utilizing mass spectrometry for detection of SNPs within the short tandem repeats. *Forensic Science International: Genetics Supplement Series* **2**, 529–531 (2009).
- [16] Marras, S. A., Russell Kramer, F. & Tyagi, S. Multiplex detection of single-nucleotide variations using molecular beacons. *Genetic Analysis: Biomolecular Engineering* **14**, 151–156 (1999).
- [17] Mhlanga, M. M. & Malmberg, L. Using molecular beacons to detect single-nucleotide polymorphisms with real-time PCR. *Methods* **25**, 463–71 (2001).
- [18] Ye, S., Dhillon, S., Ke, X., Collins, A. R. & Day, I. N. M. An efficient procedure for genotyping single nucleotide polymorphisms. *Nucleic Acids Research* **29**, e88 (2001).
- [19] Wolford, J. K., Blunt, D., Ballecer, C. & Prochazka, M. High-throughput SNP detection by using DNA pooling and denaturing high performance liquid chromatography (DHPLC). *Human Genetics* **107**, 483–487 (2000).

- [20] Brockman, W. *et al.* Quality scores and SNP detection in sequencing-by-synthesis systems. *Genome Research* **18**, 763–70 (2008).
- [21] Brizova, H., Kalinova, M., Krskova, L., Mrhalova, M. & Kodet, R. Quantitative measurement of cyclin D1 mRNA, a potent diagnostic tool to separate mantle cell lymphoma from other B-cell lymphoproliferative disorders. *Diagnostic Molecular Pathology* **17**, 39–50 (2008).
- [22] Tiveljung-Lindell, A. *et al.* Development and implementation of a molecular diagnostic platform for daily rapid detection of 15 respiratory viruses. *Journal of Medical Virology* **81**, 167–175 (2009).
- [23] Ragoussis, J. & Elvidge, G. Affymetrix GeneChip system: moving from research to the clinic. *Expert Review of Molecular Diagnostics* **6**, 145–152 (2006).
- [24] Li, R. *et al.* SNP detection for massively parallel whole-genome resequencing. *Genome Research* **19**, 1124–32 (2009).
- [25] Zhang, J., Finney, R. P., Clifford, R. J., Derr, L. K. & Buetow, K. H. Detecting false expression signals in high-density oligonucleotide arrays by an in silico approach. *Genomics* **85**, 297–308 (2005).
- [26] Draghici, S., Khatri, P., Eklund, A. C. & Szallasi, Z. Reliability and reproducibility issues in DNA microarray measurements. *Trends in Genetics* **22**, 101–9 (2006).
- [27] Bustin, S. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology* **25**, 169–193 (2000).
- [28] Corless, C. E. *et al.* Contamination and sensitivity issues with a real-time universal 16S rRNA PCR. *Journal of Clinical Microbiology* **38**, 1747–52 (2000).
- [29] Nolan, T., Hands, R. E., Ogunkolade, W. & Bustin, S. A. SPUD: a quantitative PCR assay for the detection of inhibitors in nucleic acid preparations. *Analytical Biochemistry* **351**, 308–10 (2006).

- [30] Bustin, S. A. Molecular medicine, gene-expression profiling and molecular diagnostics: putting the cart before the horse. *Biomarkers in Medicine* **2**, 201–7 (2008).
- [31] Tyagi, S. & Kramer, F. R. Molecular beacons: probes that fluoresce upon hybridization. *Nature Biotechnology* **14**, 303–308 (1996).
- [32] Li, J. J., Chu, Y., Lee, B. Y.-H. & Xie, X. S. Enzymatic signal amplification of molecular beacons for sensitive DNA detection. *Nucleic Acids Research* **36**, e36 (2008).
- [33] Dirks, R. M. & Pierce, N. A. Triggered amplification by hybridization chain reaction. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 15275–8 (2004).
- [34] Beck, V. A. Personal communication (2010).
- [35] Demidov, V. V. & Frank-Kamenetskii, M. D. Two sides of the coin: affinity and specificity of nucleic acid interactions. *Trends in Biochemical Sciences* **29**, 62–71 (2004).
- [36] Zhang, D. Y. & Winfree, E. Control of DNA strand displacement kinetics using toehold exchange. *Journal of the American Chemical Society* **131**, 17303–17314 (2009).
- [37] Zadeh, J. N., Wolfe, B. R. & Pierce, N. A. Nucleic acid sequence design via efficient ensemble defect optimization. *Journal of Computational Chemistry* **32**, 439–452 (2011).
- [38] Dirks, R. M. & Pierce, N. A. An algorithm for computing nucleic acid base-pairing probabilities including pseudoknots. *Journal of Computational Chemistry* **25**, 1295–1304 (2004).
- [39] Dirks, R. M., Bois, J. S., Schaeffer, J. M., Winfree, E. & Pierce, N. A. Thermodynamic Analysis of Interacting Nucleic Acid Strands. *SIAM Review* **49**, 65 (2007).
- [40] Zadeh, J. N. *et al.* NUPACK: Analysis and design of nucleic acid systems. *Journal of Computational Chemistry* **32**, 170–173 (2011).
- [41] Serra, M. J. & Turner, D. H. Predicting thermodynamic properties of RNA. *Methods in Enzymology* **259**, 242–61 (1995).

- [42] Davies, H. *et al.* Mutations of the BRAF gene in human cancer. *Nature* **417**, 949–54 (2002).
- [43] Maldonado, J. L. *et al.* Determinants of BRAF mutations in primary melanomas. *Journal of the National Cancer Institute* **95**, 1878–90 (2003).
- [44] Uribe, P., Wistuba, I. I. & González, S. BRAF mutation: a frequent event in benign, atypical, and malignant melanocytic lesions of the skin. *The American Journal of Dermatopathology* **25**, 365–370 (2003).
- [45] N Poynter, J. *et al.* BRAF and NRAS mutations in melanoma and melanocytic nevi. *Melanoma Research* **16**, 267–273 (2006).
- [46] Rubinstein, J. C. *et al.* Incidence of the V600K mutation among melanoma patients with BRAF mutations, and potential therapeutic response to the specific BRAF inhibitor PLX4032. *Journal of translational medicine* **8**, 67 (2010).
- [47] Lovly, C. M. *et al.* Routine multiplex mutational profiling of melanomas enables enrollment in genotype-driven therapeutic trials. *PloS One* **7**, e35309 (2012).
- [48] Ascierto, P. A. *et al.* The role of BRAF V600 mutation in melanoma. *Journal of Translational Medicine* **10**, 85 (2012).
- [49] Steensma, D. P. JAK2 V617F in myeloid disorders: molecular diagnostic techniques and their clinical utility. *The Journal of Molecular Diagnostics* **8**, 397–411 (2006).
- [50] Murugesan, G. *et al.* Identification of the JAK2 V617F mutation in chronic myeloproliferative disorders using FRET probes and melting curve analysis. *American Journal of Clinical Pathology* **125**, 625–633 (2006).
- [51] Kanaya, T. *et al.* Association of mismatch repair deficiency with PTEN frameshift mutations in endometrial cancers and the precursors in a Japanese population. *American Journal of Clinical Pathology* **124**, 89–96 (2005).
- [52] Kurman, R. J. & Shih, I.-M. Molecular pathogenesis and extraovarian origin of epithelial ovarian cancer—shifting the paradigm. *Human Pathology* **42**, 918–31 (2011).

- [53] Mathews, D. H., Sabina, J., Zuker, M. & Turner, D. H. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *Journal of Molecular Biology* **288**, 911–40 (1999).
- [54] Vet, J. A. *et al.* Multiplex detection of four pathogenic retroviruses using molecular beacons. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 6394–9 (1999).

Chapter 3

Transducing Sequence to Light with Quenched HCR

3.1 Introduction

Molecular beacons [1] are nucleic acid sequences that fold into a toehold-free hairpin structure. Upon hybridization with a target of choice via their loop domain, molecular beacons undergo a conformational change that disrupts the integrity of their stem. As the 5'-end of molecular beacons is conjugated to a fluorophore and their 3'-end is conjugated to a quencher, hybridization to target disrupts their fluorophore-quencher FRET pair. This, in turn, leads to an increase in emitted fluorescence. Because molecular beacons directly hybridize with nucleic acid targets and offer high sequence selectivity, they have found broad use in biological applications. Some examples include SNP detection in homogenous solution [2], *in situ* RNA imaging [3], and dynamic monitoring of RNA in live cells [4–6]. While the isothermal, direct mode of target detection provided by molecular beacons has facilitated their broad use, the fact that a maximum of one molecular beacon can hybridize with a target site has limited their sensitivity to the nanomolar scale [7].

One way to overcome this sensitivity limit is to enrich the number of targets via the use of PCR. An increasing number of target sites is generated through the amplification cycles of PCR, and, as a result, an increase in fluorescence is recorded in real time, resulting in one variant of real-time PCR [8]. Real-time PCR can also be used to detect RNA targets (following a reverse transcription step) in a reaction known as “real-time reverse

transcription PCR (qRT-PCR).” Since qRT-PCR is both sensitive and semi-quantitative¹, it has recently found some diagnostic use. At the same time, the sensitivity improvement seen with qRT-PCR requires the use of a thermocycler. Large, expensive, and in need of continuous electricity supply, thermocyclers limit the use of qRT-PCR as a point-of-care diagnostic on the one hand, and render qRT-PCR unsuitable for both *in situ* and live cell applications on the other hand.

In an attempt to capture the isothermal mode of detection offered by molecular beacons while simultaneously achieving target amplification, Li et al. developed the Nicking Enzyme Signal Amplification (NESA) method and obtained ~ 6.2 pM sensitivity [9]. At the same time, to create the recognition sequence² required by the nicking enzyme with a target of choice, the authors incorporated heat inactivation steps into their protocol, thus re-introducing the dependence on a thermocycler. As NESA requires both a thermocycler and enzymes (DNA ligase and DNA polymerase), NESA’s limitations in a point-of-care setup are similar to those faced by qRT-PCR.

In this chapter, we explore Quenched HCR as a new method for nucleic acid detection in bulk. We recall from Section 2.2 that HCR entails the self-assembly of hairpins into polymers in the presence of a target sequence with which the hairpins are designed to hybridize [10]. To date, all applications in which direct readout of HCR polymers was sought required either gel analysis [10], immobilization to surfaces [11], or tethering to immobilized RNA *in situ* followed by a wash [12]. Here, we discuss Quenched HCR, as the first wash-free method for HCR readout in bulk. Quenched HCR combines some of the positive attributes of molecular beacons with those of qRT-PCR. Like molecular beacons, Quenched HCR offers an isothermal, non-enzymatic mode of detection. Like qRT-PCR, Quenched HCR provides signal amplification, though not to the same extent. For these reasons, we propose that Quenched HCR may be integrated into *in vitro* applications, e.g., diagnostics, and expression profiling in a lysate, *in situ* applications, e.g., HCR-ISH [12],

¹A comparison of the fluorescent signal generated by a transcript of interest (via the transcript’s cDNA proxy) and analogous fluorescent signals generated by multiple housekeeping genes, provides quantitative ratios.

²Li et al. proposed Rolling Circle Amplification (RCA) as a method of transducing between a target sequence of choice and the recognition sequence required by a nicking enzyme of choice. By introducing RCA to their method, the authors were able to reach fM sensitivity.

and possibly, live cell applications, e.g., cell sorting.

3.2 Design of 2-Hairpin and 4-Hairpin Quenched HCR

We sought to design Quenched HCR systems that polymerize in the time scale of dozens of minutes or more so as to allow Quenched HCR optical measurement from $\sim t_0$ and onwards. Accordingly, DNA hairpins composed of toehold/loop sizes of 6-nt, and stem sizes of 18–19-nt were designed. The domain sizes and sequence compositions of hairpins were chosen such that the free energy [13, 14] of the hairpins’ stems is ≥ 25 kcal/mol, and the free energy of their toehold and loop domains is 7–10.5 kcal/mol³.

The design of Quenched HCR systems necessitates consideration of the distance between the fluorophore and quencher. This distance, which is mostly controlled by the ΔG^4 between the fluorophore and quencher, and the length and flexibility of the linkers used to conjugate these two moieties, ought to be within the range of quenching in the hairpin conformation and mostly out of the range of quenching in the polymer conformation. Hence, we required that the quencher be placed in the 5'-end or 3'-end of h1 (in accordance with IDT’s catalog orders), and that the fluorophore be placed across from the quencher (internally, between the toehold and the stem domains, as is illustrated in Figure 3.1, panels A, B).

The combination of these design criteria, for the relatively small-toehold systems we focused on, suggests that the use of 4-hairpin-periodicity Quenched HCR (Figure 3.1, panel A) is preferable to the use of 2-hairpin-periodicity Quenched HCR (Figure 3.1, panel B). In 4-hairpin-periodicity Quenched HCR, the distance between the fluorophore and the quencher in the polymer conformation is increased via the insertion of “insulating,” unlabeled hairpins between every dually labeled hairpin and its nearest neighbor on the same side of the double-helix HCR polymers. In 2-hairpin Quenched HCR, the distance between the

³NUPACK calculations were based on parameters that were obtained by SantaLucia et al. [14]. Calculations did not account for the ΔG contribution of fluorophore-quencher modifications. Toehold and loop free energy calculations utilized the sequences of these domains and their reverse DNA complements, without considering the nearest neighbors of these two domains. Lastly, the free energy range of 7–10.5 kcal/mol characterizes elongation steps of the HCR reaction, but does not characterize the free energy of T·h1 formation from T and h1, since this free energy value is derived from the addition of an DNA h1 to an RNA target.

⁴Differences in $\Delta\Delta G$ between hairpins that are modified with fluorophore-quencher pairs and their unmodified analogues are expected, and may require different stem dimensions for the dually modified hairpins than for the unmodified hairpin analogues.

fluorophore and quencher within the polymer can be too small (Figure 3.1 panel B). With internally labeled 2-hairpin Quenched HCR (Figure 3.1 panel C), however, the distance between the quencher and the fluorophore within the polymer is predicted to be large enough to facilitate the use of small DNA hairpins of toehold / loop sizes of 6-nt, and stem sizes of 18–19-nt. Internally-labeled 2-hairpin Quenched HCR was not studied here, because internally labeled hairpins are difficult to synthesize. We note that quenching distance drops as $\frac{1}{R^6}$, where R is the distance between the fluorophore and quencher [15], and for our choice of quenchers, the distance in which the quencher is at $\sim 50\%$ efficiency is reported to be between 3–7nm by the manufacturer (IDT) [16].

Lastly, two fluorophore-quencher pairs with minimal cross-talk were chosen so as to allow for multiplexing. As Integrated DNA Technologies (IDT) offer a limited set of five different internal fluorophore modifications (6-FAMK, Fluorescein, Cy3, TAMRA and Cy5), the maximum number of spectrally distinct Quenched HCR systems (known to be sufficiently resistant to photobleaching) that we could test was two. Accordingly, Cy3 and Cy5 were chosen as the fluorophores with which to label the Quenched HCR systems.

3.3 Analysis of 2-Hairpin and 4-Hairpin Quenched HCR

To demonstrate the feasibility of Quenched HCR without placing any requirements on the sequences of the Quenched HCR systems, the mutated states of three of the most-commonly occurring SNPs in cancer were chosen as targets: BRAF T→A, JAK2 G→T, and PTEN C→G, denoted by T¹, T², and T³, respectively.

4-hairpin-periodicity and 2-hairpin-periodicity DNA-based Quenched HCR systems (Figure 3.1, panel A and B, respectively) were designed according to the specifications outlined in Section 3.2. Both the 4-hairpin-periodicity and the 2-hairpin-periodicity Quenched HCR systems utilize the same dually labeled h1 that hybridizes with target. They differ from each other, however, in the loop of h2, which contains a reverse-complement sequence for h1 in the 2-hairpin-periodicity system, but does not contain a reverse-complement sequence to h1’s toehold in 4-hairpin-periodicity system. The 4-hairpin-periodicity system also utilizes two additional hairpins (h3 and h4).

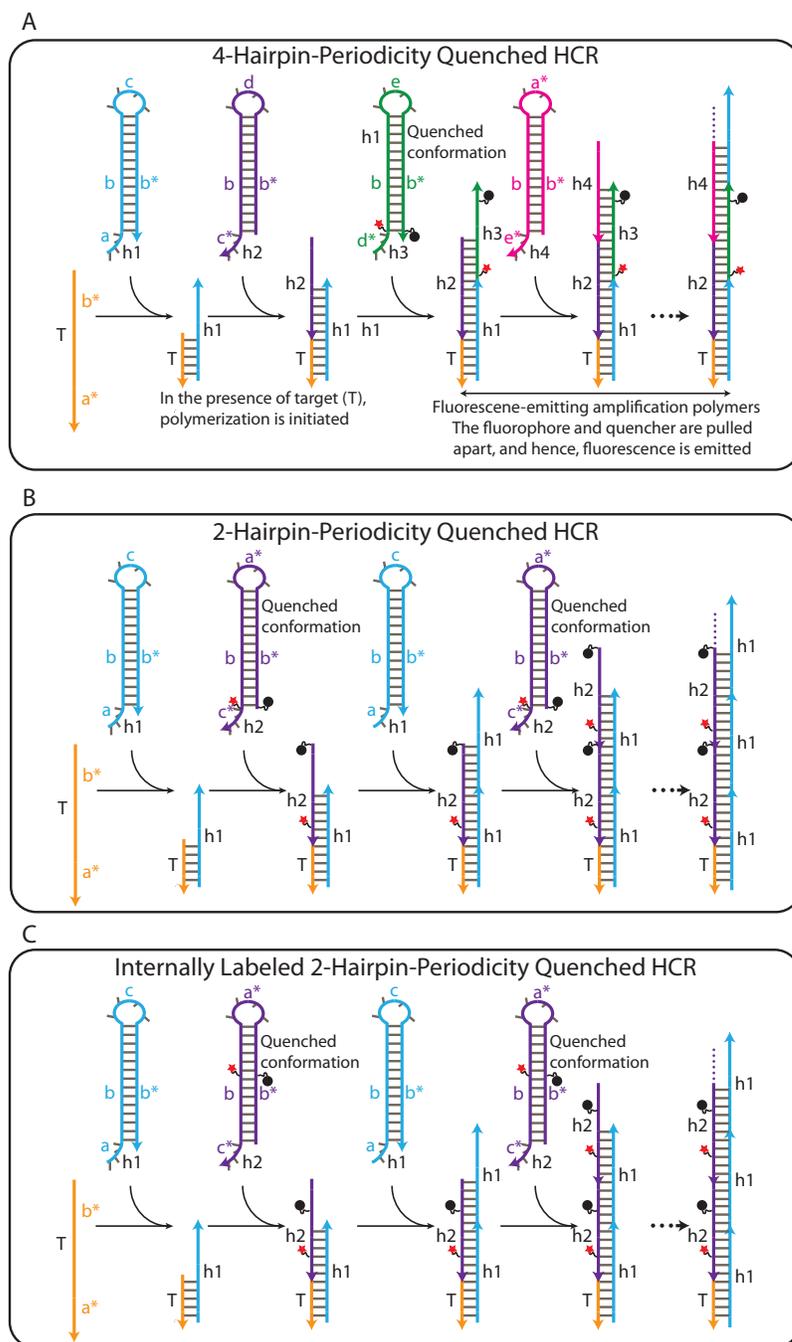


Figure 3.1: (A) 4-hairpin-periodicity Quenched-HCR. Hairpins labeled with a fluorophore-quencher pair are integrated into an HCR polymer that has a period of 4 hairpins, i.e., one fluorophore-quencher-labeled hairpin is integrated into HCR per every 4 hairpins of which the HCR polymer consists. Upon integration into polymers, the fluorophore-quencher-labeled hairpin undergoes a conformational change that pulls the fluorophore-quencher pair apart, thus resulting in an increase in the emitted light. (B) 2-hairpin-periodicity HCR is predicted to yield a short fluorophore-quencher distance in HCR polymers thus maintaining the quencher at above 50% efficiency when hairpins of toehold / loop sizes of 6-nt, and stem sizes of 18–19-nt are used. (C) Internally-labeled 2-hairpin-periodicity Quenched HCR is predicted to yield long fluorophore-quencher distances even when small hairpins are used.

To analyze Quenched HCR, we designed 2-hairpin-periodicity and 4-hairpin-periodicity Quenched HCR systems that detect cancer markers BRAF T→A, JAK2 G→T, and PTEN C→G (Section 2.6). We denote these targets as T^1 , T^2 , and T^3 , respectively, the 2-hairpin-periodicity Quenched HCR systems that detect them as P^1 , P^2 , and P^3 , respectively, and the 4-hairpin-periodicity Quenched HCR that detect the targets as Q^1 , Q^2 , and Q^3 , respectively. As demonstrated in Figure 3.2, 4-hairpin-periodicity Quenched HCR systems were successfully designed for all three targets. Among the 2-hairpin-periodicity Quenched HCR systems that were tested, P^3 and P^2 were successfully implemented, with the former performing better. As expected (Section 3.2), 4-hairpin-periodicity Quenched HCR outperformed 2-hairpin-periodicity Quenched HCR.

In addition to engineering 4-hairpin Quenched HCR in a long time scale, focusing on monitoring the kinetic performance of Quenched HCR starting at $\sim t_0$, we also designed a fast Quenched HCR system (Figure B.4). Guided by the ΔG tuning principle (Section 2.3), fast Quenched HCR utilized hairpins of toehold / loop sizes of 10-nt, and stem sizes of 26-nt. Since 10-nt toeholds were used in this instance, as opposed to the 6-nt toeholds utilized in Figure 3.2, fast Quenched HCR utilized 2-hairpin periodicity (Figure 3.1, panel B). Lastly, we note that an RNA-based 4-hairpin periodicity Quenched HCR was also designed, and successfully implemented (Appendix B.3).

3.4 Multiplexing Quenched HCR

The utilization of fluorophores that emit at different wavelengths allows for multiplexed gene detection in a single test tube. To demonstrate this, two test tube compositions ($T^2 + Q^2 + Q^3$, and $T^3 + Q^2 + Q^3$) are analyzed in this section; additional conditions are presented in Figure B.1. The total fluorescence traces generated from this study demonstrate that multiplex Quenched HCR performs well: both test tubes produce monotonically increasing optical trace for the Quenched HCR system that they are expected to turn ‘on,’ but not for the Quenched HCR system that should be kept ‘off’ (Figure 3.3).

We note that, in addition to the shapes of the curves, the intersections with the y-axis

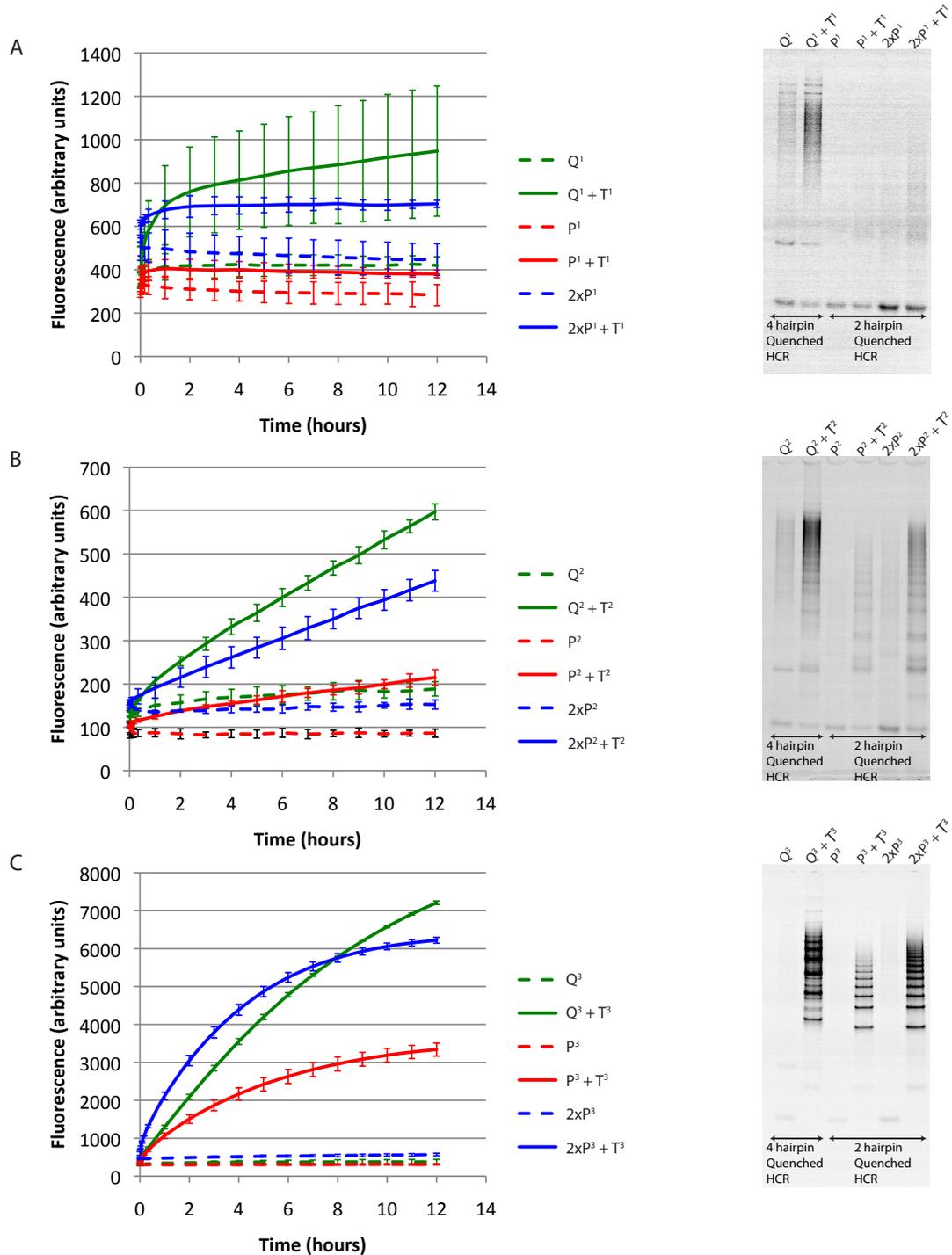


Figure 3.2: Each of three RNA targets (T^x) was incubated with a DNA end-labeled 2-hairpin-periodicity Quenched HCR system (P^x), as well as with a 4-hairpin-periodicity Quenched HCR (Q^x). Total fluorescence as a function of time was recorded on a real-time PCR machine. Plotted data represent mean and standard deviation for two experiments. Samples from the first experiment were analyzed on a 10% TBE native gel. (A) Target T^1 is the mutated state of the BRAF $T \rightarrow A$ cancer marker; Quenched HCR systems P^1 and Q^1 are Cy5-labeled. (B) Target T^2 is the mutated state of the JAK2 $G \rightarrow T$ cancer marker; Quenched HCR systems P^2 and Q^2 are Cy3-labeled. (C) Target T^3 is the mutated state of the PTEN $C \rightarrow G$ cancer marker; Quenched HCR systems P^3 and Q^3 are Cy5-labeled.

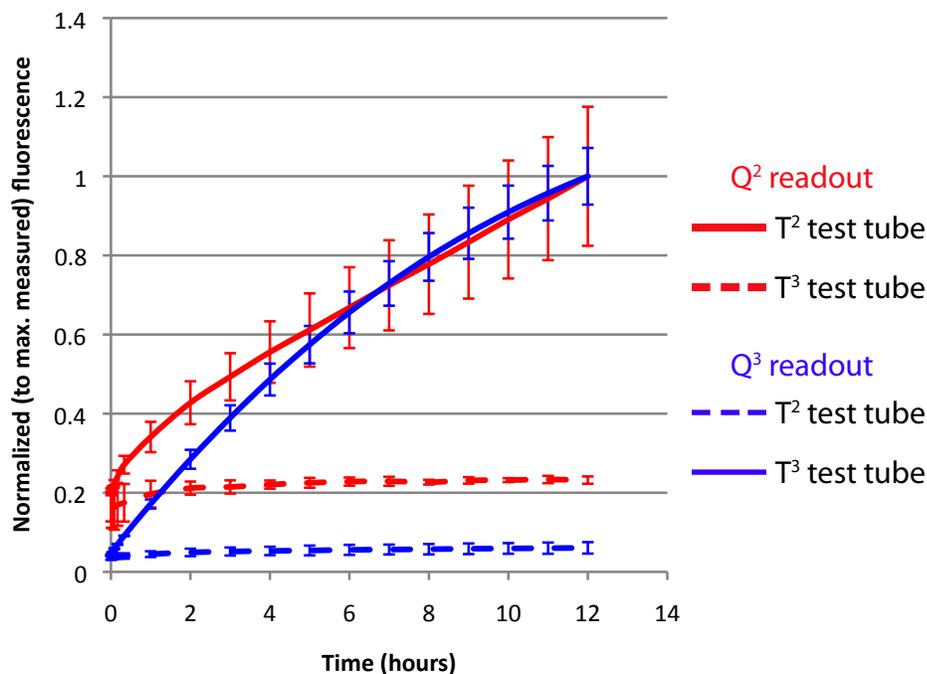


Figure 3.3: Multiplex detection of cancer gene markers. Two test tubes containing either the JAK2 cancer marker T^2 or the PTEN cancer marker T^3 were incubated in the presence of two 4-hairpin-periodicity Quenched HCR systems (Q^2 and Q^3 designed to detect T^2 and T^3 , respectively). Each of the test tubes reveals an increase in signal with time in a manner that agrees with its target gene content. Quenched HCR system 2 is labeled with Cy3, and is monitored using 515–530nm excitation filter and a 560–580nm emission filter. Quenched HCR system 3 is labeled with Cy5, and is monitored using 620–650nm excitation filter and 675–690nm emission filter. Plotted data represent mean and standard deviation of two experiments, normalized by the maximum mean value in each channel.

provide some indication⁵ of the extent to which the dually labeled hairpin is quenched, where the closer the intersection is to zero, the more quenched the dually labeled hairpin. With respect to this, we note that the Cy5-labeled system that detects T^3 appears more quenched than the Cy3-labeled system that detects T^2 . This agrees with the data collected in the single-channel experiments (Figure 3.2).

⁵The exact interpretation of the intersect with the y-axis is the ratio between total signal that is observed in a test tube in the first measurement of the time-course and the highest fluorescence measurement throughout the experiment.

3.5 Conclusion

Quenched HCR using DNA or RNA hairpins (Figure B.3) was demonstrated to selectively detect RNA targets. Quenched HCR can be parallelized when used *in vitro*; in our experiments we analyzed as many as 36 test tubes at the same time, but the upper limit, which depends on the optical device used, is significantly higher.

In its basic *in vitro* implementation, HCR requires the use of an additional assay to determine the extent to which polymers are present in solution. Traditionally, this assay takes the form of gel electrophoresis.

Quenched HCR has solved, or ameliorated, four problems that pertain to gel analysis of HCR. First, single-time-point monitoring of HCR polymers, as provided by gel analysis, can be replaced with continuous monitoring of the polymerization signal. Second, the time (normally ~ 30 minutes or more) that elapses between an HCR reaction and the completion of both the gel run and the gel imaging has been effectively reduced to zero. Third, the materials, chemical waste, and costs associated with running a gel have also been dramatically reduced. Fourth, the challenge of quantifying often blurred bands from the background of a gel, which is not transparent through a large portion of the visible light spectrum, has been eliminated, and instead, an instrument that monitors fluorescence is required [17].

In addition to *in vitro* applications e.g., diagnostics and wash-free expression profiling in lysate, Quenched HCR may prove useful *in situ* and in live cells, as it provides a wash-free approach to mapping genetic expression.

References

- [1] Tyagi, S. & Kramer, F. R. Molecular beacons: probes that fluoresce upon hybridization. *Nature Biotechnology* **14**, 303–308 (1996).
- [2] Tyagi, S., Bratu, D. P. & Kramer, F. R. Multicolor molecular beacons for allele discrimination. *Nature Biotechnology* **16**, 49–53 (1998).
- [3] Lenaerts, J., Lappin-Scott, H. M. & Porter, J. Improved fluorescent in situ hybridization method for detection of bacteria from activated sludge and river water by using DNA molecular beacons and flow cytometry. *Applied and Environmental Microbiology* **73**, 2020–3 (2007).
- [4] Bratu, D. P., Cha, B.-J., Mhlanga, M. M., Kramer, F. R. & Tyagi, S. Visualizing the distribution and transport of mRNAs in living cells. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 13308–13313 (2003).
- [5] Mhlanga, M. M., Vargas, D. Y., Fung, C. W., Kramer, F. R. & Tyagi, S. tRNA-linked molecular beacons for imaging mRNAs in the cytoplasm of living cells. *Nucleic Acids Research* **33**, 1902–1912 (2005).
- [6] Bratu, D. P., Catrina, I. E. & Marras, S. A. E. Tiny molecular beacons for in vivo mRNA detection. *Methods in Molecular Biology (Clifton, N.J.)* **714**, 141–57 (2011).
- [7] Zhang, P., Beck, T. & Tan, W. Design of a Molecular Beacon DNA Probe with Two Fluorophores. *Angewandte Chemie (International ed. in English)* **40**, 402–405 (2001).
- [8] Manganelli, R., Tyagi, S. & Smith, I. Real Time PCR Using Molecular Beacons : A New Tool to Identify Point Mutations and to Analyze Gene Expression in *Mycobacterium tuberculosis*. *Methods in Molecular Medicine* **54**, 295–310 (2001).
- [9] Li, J. J., Chu, Y., Lee, B. Y.-H. & Xie, X. S. Enzymatic signal amplification of molecular beacons for sensitive DNA detection. *Nucleic Acids Research* **36**, e36 (2008).
- [10] Dirks, R. M. & Pierce, N. A. Triggered amplification by hybridization chain reaction.

Proceedings of the National Academy of Sciences of the United States of America **101**, 15275–8 (2004).

- [11] Yang, L., Liu, C., Ren, W. & Li, Z. Graphene surface-anchored fluorescence sensor for sensitive detection of microRNA coupled with enzyme-free signal amplification of hybridization chain reaction. *ACS Applied Materials & Interfaces* **4**, 6450–3 (2012).
- [12] Choi, H. M. T. *et al.* Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nature Biotechnology* **28**, 1208–1212 (2010).
- [13] Zadeh, J. N., Wolfe, B. R. & Pierce, N. A. Nucleic acid sequence design via efficient ensemble defect optimization. *Journal of Computational Chemistry* **32**, 439–452 (2011).
- [14] SantaLucia, J. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proceedings of the National Academy of Sciences* **95**, 1460–1465 (1998).
- [15] Lackowicz Joseph R. *Principles of Fluorescence Spectroscopy (Google eBook)* (Springer, 2009).
- [16] Integrated DNA Technologies - Probes. URL <http://www.idtdna.com/catalog/DNaProbes/Page1.aspx>.
- [17] Chang, B.-Y. Smartphone-based Chemistry Instrumentation: Digitization of Colorimetric Measurements. *Bulletin of the Korean Chemical Society* **33**, 549–552 (2012).

Chapter 4

Sequence Transduction with Conditional Probes

4.1 Introduction

As was discussed in Chapter 2, HCR is a fast, enzyme-free, isothermal, selective nucleic acid detector. While these properties suggest that HCR could be used as a diagnostic, the dependence of HCR on a gel-based [1], or immobilization-based assay [2] for purposes of differentiating between HCR polymers and hairpin-monomers diminishes the utility of HCR in some diagnostic settings.

In Chapter 3 we discussed Quenched HCR, which was developed to eliminate HCR's dependence on gel electrophoresis. Quenched HCR is a technique that measures fluorescent signal as a proxy for polymer formation, thus eliminating the loss of time and reagents associated with running gel electrophoresis. While Quenched HCR maintains many of the positive attributes of HCR, Quenched HCR also poses non-trivial synthesis challenges. Specifically, the dually labeled molecules that Quenched HCR requires are costly and time-consuming to generate.

There is, therefore, a need for a molecule that hybridizes with both a target sequence of choice and Quenched HCR. A transducer molecule of this nature would allow for the detection of a target of choice using Quenched HCR without the need to design, synthesize, and run quality control experiments on Quenched HCR systems every time the detection of a new gene is sought.

For wash-free applications of the kind discussed in this chapter, as well as in Chapter 3,

the transducer must be conditional, that is, the transducer should only expose an initiator sequence for the cognate Quenched HCR, or other triggered hybridization system of choice, if target is present. An unconditional probe, by comparison, would trigger the formation of Quenched HCR independent of the absence/presence of target, thus removing the diagnostic capability of Quenched HCR. Since the probe design discussed in this chapter is conditional, we refer to it as “Conditional Probe,” or “CP.”

Conditional Probe executes molecular logic similar to that of molecular beacons. Molecular beacons are nucleic acids that assume a stem-loop structure in the absence of target, but lose the integrity of their stems in the presence of target (Figure 4.1 panel A) [3]. Molecular beacons can therefore be viewed as conditional probes that undergo a triggered reaction wherein their stem-sequestered sequences turn into single-stranded sequences only in the presence of target. Hence, these molecules carry out the following logic operation: in the presence of sequence A, expose sequence space B. Although molecular beacons are inherently conditional sequence transducers, they have found very limited use as such, and instead have been used primarily as fluorescent reporters¹. The design of CP is inspired by molecular beacons. Unlike molecular beacons that transduce sequence A-to-light, however, we propose to engineer CPs to transduce sequence A-to-sequence B.

Upon hybridization with a target, via the CP loop, the stem of CP is disrupted, exposing single stranded tails that can serve as initiators for other reactions. For our purposes, the motif we wish to turn ‘on’ in the presence of target and CP is Quenched HCR.

Other groups have utilized sequence transducers. Seelig et al. [4], for example, have implemented a sequence transducer based on strand displacement reactions using duplexes. Special care is required when duplex molecular motifs are introduced to biological samples, since an excess of one strand over the other can lead to false positives. With Seelig’s sequence transducer, the output sequence is released upon detection of the input sequence. By comparison, the unimolecular CP studied here remains bound to target, and is thus suitable for *in situ* applications and for substrate-based assays.

Wilner’s and Ellington’s groups have used unimolecular sequence transducers that as-

¹Since molecular beacons are dually labeled with a fluorophore and a quencher on the 5’-end and 3’-end of their stems, respectively, the target-mediated stem disruption that molecular beacons undergo results in a measurable increase of fluorescent signal.

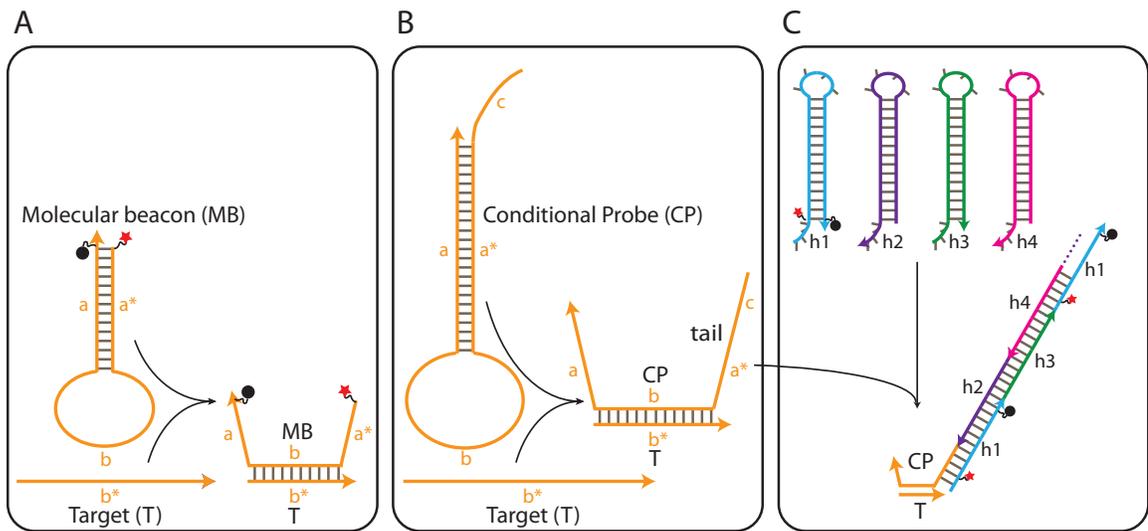


Figure 4.1: (A) Molecular beacons are toehold-free hairpins that are composed of ~ 18 – 30 -nt loop domains and ~ 5 – 7 -nt stem domains. Molecular beacons are covalently linked to a fluorophore on their 5'-end and to a quencher on their 3'-end. In the presence of target, quenching is disrupted, and a sequence that is orthogonal to that of the target is exposed [3]. (B) Adaptation of the sequence-transduction capability of molecular beacons to Conditional Probe involved an enlargement of the stem domain, in addition to omission of the fluorophore and quencher modifications. Sequence domain a^* is exposed in the presence of target sequence b^* . (C) Only in the presence of both target and CP, Quenched HCR is triggered.

sume a hairpin structure that is similar to that of molecular beacons. Wilner’s group has used sequence transducers to form DNAzyme nanowires [5, 6], and Ellington’s group has used sequence transducers for DNA catalytic circuits² [7].

4.2 Design of Conditional Probes

As discussed in Section 4.1, CP is a molecular motif designed to carry out the following operation: if sequence A is present, expose sequence B. To test this molecular logic, we sought to utilize CP as a sequence transducer between Quenched HCR and targets orthogonal to the sequence of Quenched HCR. Specifically, our objective was to trigger Quenched HCR systems that were already in our possession, thus refraining from introducing any size or sequence restrictions to either CP or Quenched HCR.

The Quenched HCR systems that were chosen for our studies (Section 4.3) are Q^2 and Q^3 , which were discussed in Chapter 3. Because these Quenched HCR systems consist of 6-nt toehold domains and 18–19-bp stem domains, triggering with a full-sized reverse complement sequence requires a CP with a 24–25-nt tail (Figure 4.1, panels B and C).

To ensure that CP stems are sufficiently open in the presence of targets so as to expose a sequence of 24–25-nt, our CPs feature a combined toehold+stem length of 24–25-nt. We devised three test cases for Q^2 and three test cases for Q^3 (Section 4.3). To trigger Q^2 , CPs composed of stem-toehold domain sizes of 25-0-nt, 20-5-nt and 15-10-nt were utilized. To trigger Q^3 CPs composed of stem-toehold domain sizes of 24-0-nt, 19-5-nt and 14-10-nt were used.

The main difference between these CPs is the extent to which their stems are expected to be open during hybridization with target and, hence, the speed with which Quenched HCR is expected to proceed. The longer the toehold, the more favorable the free energy of the reaction $T \cdot CP + h1 \rightarrow T \cdot CP \cdot h1$ is, and hence, the faster Quenched HCR polymerization is expected to initiate. At the same time, the longer the toehold, the more the undesired reaction of Quenched HCR polymerization is expected to occur in the presence

²It may be that the sequence transducer used by Ellington’s group is leaky, i.e., not conditional, since no data is provided for test tubes that contain the sequence transducer as well as the DNA circuit, but are absent of target.

of only Conditional Probe, i.e., without target. Hence, Conditional Probe stems cannot be arbitrarily short.

Our molecular designs focused on DNA-based conditional probes, due to their resistance to degradation and low cost, and RNA-based targets, due to their biological interest.

4.3 Experimental Verification of Conditional Probe Function

To verify the utility of the CP designs discussed in Section 4.2, Quenched HCR systems were incubated in the presence of either CP (to test for undesired triggering of Quenched HCR), or CP and target, at 25°C. The Quenched HCR systems utilized here are Q^2 and Q^3 , which are the better performing DNA-based Quenched HCR systems from Section 3.3. For each of these Quenched HCR systems, we sought to engineer a CP that can trigger Quenched HCR when in the presence of two orthogonal targets. For this purpose, T^1 and T^3 served as orthogonal targets for Q^2 , and T^1 and T^2 served as orthogonal targets for Q^3 . $CP^{1,2}$ and $CP^{3,2}$ were utilized to trigger Q^2 in the presence of T^1 and T^3 , respectively. Similarly, $CP^{1,3}$ and $CP^{2,3}$ were utilized to trigger Q^3 in the presence of T^1 and T^2 respectively.

As mentioned in Section 4.2, CPs of three varieties were tested: long stem and no toehold, middle stem and middle toehold, and lastly, short stem, and long toehold. Each of $CP^{1,2}$, $CP^{3,2}$, $CP^{1,3}$, and $CP^{2,3}$ was tested with all three of these varieties. Hence, a total of 12 CPs were analyzed. To distinguish between the three varieties of CPs, we affix the letters L, M, or S, which stand for long, medium, or short stem length, to each of the conditional probes that appear in Figure 4.2.

Conditional probes function in a desirable manner only when they contain a toehold (Figure 4.2). When toehold-containing conditional probes are utilized, Quenched HCR systems remain mostly dark when in the presence of only CP, but they undergo dark-to-light transformation when in the presence of both CP and T, as desired. This observation suggests that if the CP stem is too long, the CP does not sufficiently expose the HCR initiator upon target binding.

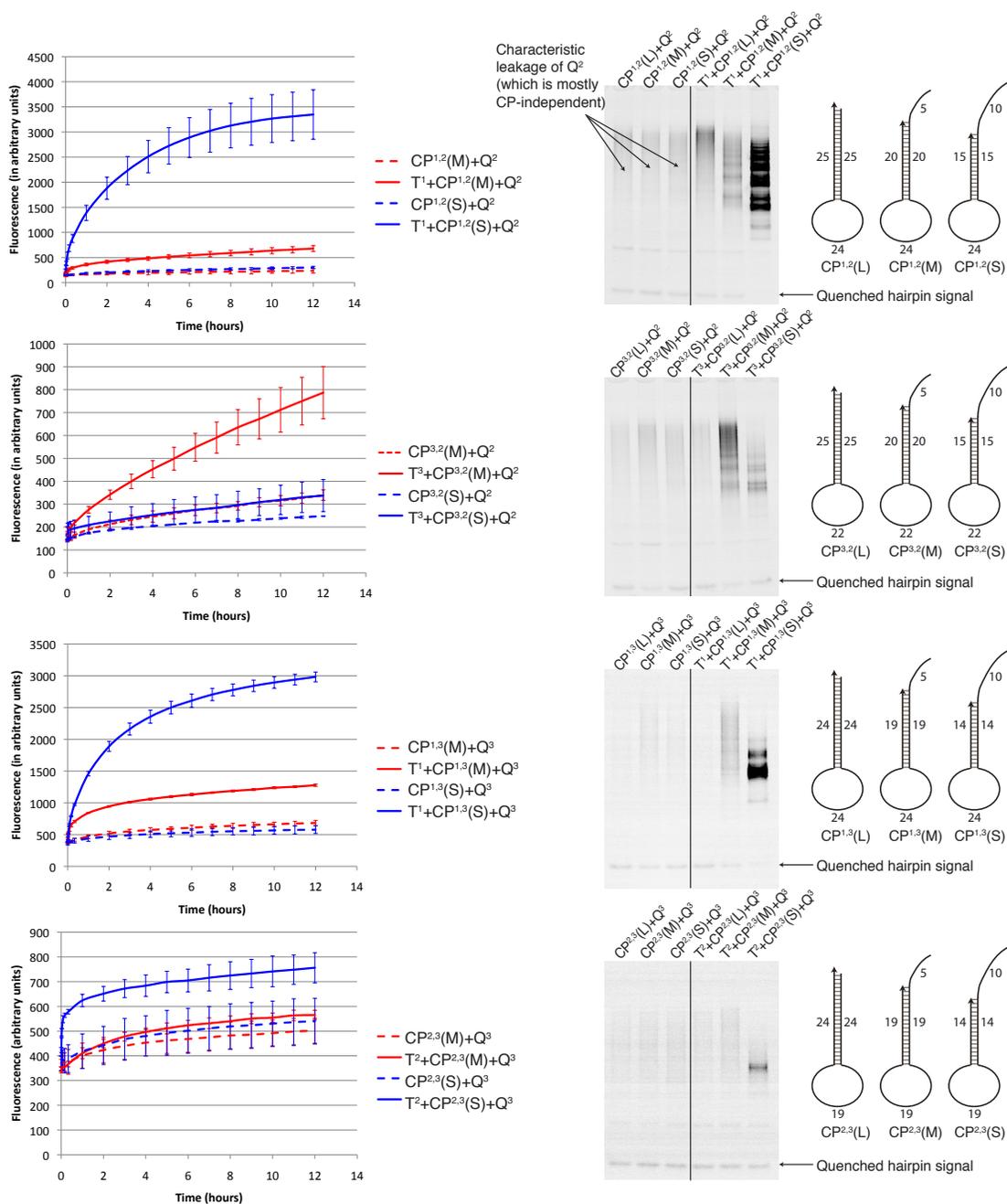


Figure 4.2: Sequence transduction with conditional probes. Cancer markers BRAF T→A, JAK2 G→T, and PTEN C→G, denoted by T¹, T², and T³, respectively, are orthogonal to each other. In Chapter 3, it was demonstrated that these target sequences selectively trigger Quenched HCR systems Q¹, Q², and Q³. Here, Quenched HCR systems Q² and Q³ are triggered by target sequences that are orthogonal to their own in the presence of relevant toehold-containing conditional probes. Conditional probes (CPs) are marked with superscript indices that denote the target entity and the Quenched HCR system between which the CPs mediate. L, M, and S denote long, medium, and short CP stem lengths. Plotted data represent mean and standard deviation of two experiments. Traces for CP(L) are not included in the time course data, since CP(L) data very closely matched background data in all experiments. Intended CP secondary structure is illustrated at the right side of each panel.

4.4 Multiplexing via Conditional Probe and Quenched HCR

In Section 4.3, it was demonstrated that CP is an effective sequence transducer between a target of choice and a Quenched HCR system of choice. Next, we wished to demonstrate that CP allows for target detection in multiplex. To demonstrate multiplexing, Q² and Q³ were used as the Quenched HCR systems to be triggered by the orthogonal target sequences T¹ and T⁴. Q², Q³, and T¹ are the same sequences utilized throughout Chapter 3 and 4. T⁴ is a short RNA sequence of the destabilized, enhanced, green fluorescent gene (d2EGFP).

The fluorescent traces shown in Figure 4.3, panels A and B indicate that CP and Quenched HCR can be multiplexed. The gel analysis in Figure 4.3, done upon completion of the real-time fluorescence measurements (of the first of two experiments), provides further indication that CP and Quenched HCR can be used in multiplex. Specifically, Q³ provides the majority of signal in lane 1, Q² provides the majority of signal in lane 2, and both Q² and Q³ provide signal in lane 3, as desired.

4.5 Detection of Long RNA

In Section 4.3 it was demonstrated that CP allows for selective triggering of Quenched HCR in the presence of short RNA targets that are orthogonal to Quenched HCR. As a follow-up to this study, we sought to initiate Quenched HCR using the same conditional probes, but for long RNA targets rather than short RNA targets.

To test the performance of CP with long RNA, a transcript of 1213-nt consisting primarily of a (C→G) variant of the PTEN gene was transcribed *in vitro* using T7 polymerase. Subsequently, this transcript was incubated at 25°C with Q², and with either CP^{3,2}(M) or CP^{3,2}(S) (Section 4.3). As is demonstrated in Figure 4.4, CP successfully detects long RNA targets.

4.6 Conclusion

In Chapter 3 we demonstrated that Quenched HCR undergoes a dark-to-light transition in the presence of a target of choice. Since Quenched HCR hybridizes to targets, its sequence

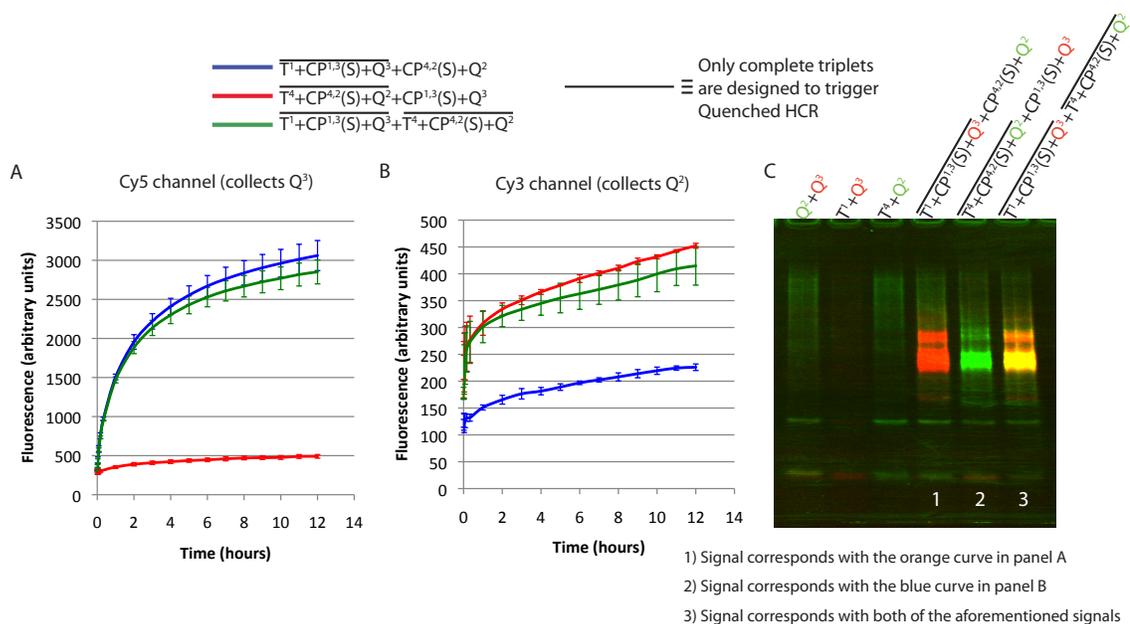


Figure 4.3: Conditional Probe Multiplexing. (A, B) Target sequences T^1 and T^4 are orthogonal to 4-hairpin Quenched HCR systems Q^2 and Q^3 . Conditional Probe $CP^{1,3}(S)$ triggers Q^3 only in the presence of T^1 , as can be shown by the blue and green curves in panel A. Similarly, $CP^{4,2}(S)$ triggers Q^2 in the presence of T^4 . Plotted data represent mean and standard deviation of two experiments. We recall from Chapter 3 that Q^2 self-assembles into polymers to some extent even in the absence of a complementary target. (C) Native gel analysis preceded real-time PCR fluorescence measurements of the first experiment, and is in agreement with the total fluorescence traces.

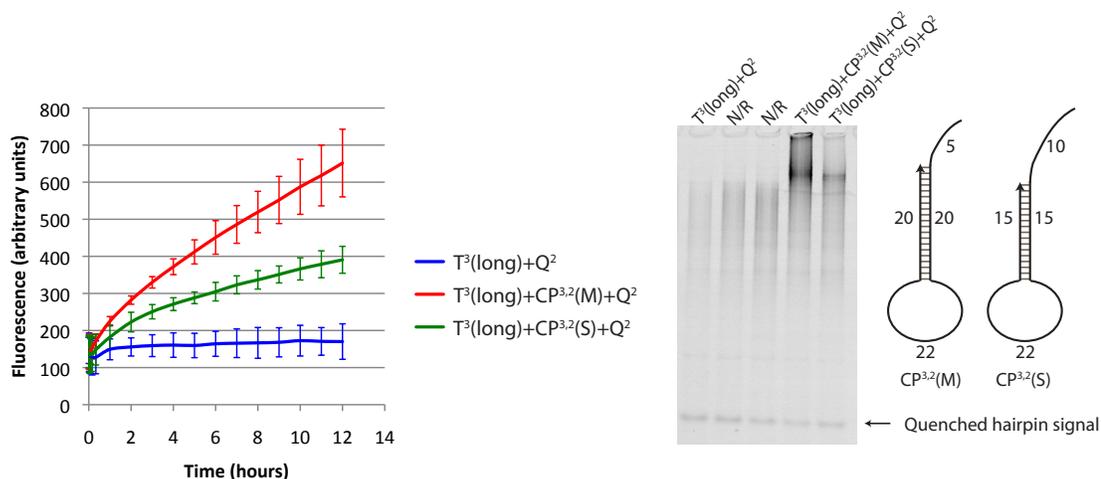


Figure 4.4: Conditional Probe mediated sequence transduction triggered by long RNA ($T^3(\text{long})$; Appendix C.2). Q^2 undergoes dark-to-light transition in the presence of $CP^{3,2}(M)$ and, to a lesser extent, in the presence of $CP^{3,2}(S)$. Plotted data represent mean and standard deviation of two experiments.

largely depends on the sequence of the target. Hence, any new target one wishes to detect requires the design, synthesis, and characterization of a new Quenched HCR system. Due to synthesis and purification challenges, the preparation of dually labeled DNA is much slower and significantly more costly than the preparation of unmodified DNA. Moreover, *in vitro* characterization of Quenched HCR is time-consuming, and cell characterization of Quenched HCR is highly time-consuming. This fact should be taken into consideration since we predict that Quenched HCR will be used in cells.

To lessen the time and cost limitations associated with the iterative design and characterization of Quenched HCR, we engineered conditional probes: hairpins whose stems open when their loop hybridizes with a target of choice. Since a CP contains a stem sequence that is not dependent on CP's loop sequence, Quenched HCR systems that are mixed with conditional probes can detect target sequences that are orthogonal to their own sequences. The requirement of having to design, synthesize, and characterize a new Quenched HCR system (that consists of 2-hairpins at a minimum, of which at least one is dually labeled) for every new target sequence is, therefore, replaced with the requirement of having to design, synthesize, and characterize one unmodified CP. We expect that the latter will often be a more convenient approach for users of Quenched HCR. The ability to design and verify conditional probes, rather than new Quenched HCR systems, could be of high value in genotyping applications in which multiple sequence windows along a single gene are detected.

Conditional probes could also potentially be used to initiate HCR that is fluorescently labeled, but not quenched, thus reducing the amount of time required for HCR *in situ* hybridization protocols by eliminating the need to wash out unused probes [2]. Additionally, CPs could trigger a variety of DNA circuits [5–7]. Lastly, we expect that the CP design process will benefit from the fact that a CP's stem design can be achieved, almost entirely, without having to take into consideration the target sequence. When Quenched HCR hybridizes directly with target, i.e., without the use of conditional probes, the length of Quenched HCR's hairpin stems is dependent on the G/C content of the target. When Quenched HCR hybridizes with Conditional Probe, however, the stem lengths of Quenched

HCR hairpins cease to depend on the sequence of the target.

References

- [1] Dirks, R. M. & Pierce, N. A. Triggered amplification by hybridization chain reaction. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 15275–8 (2004).
- [2] Choi, H. M. T. *et al.* Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nature Biotechnology* **28**, 1208–1212 (2010).
- [3] Tyagi, S. & Kramer, F. R. Molecular beacons: probes that fluoresce upon hybridization. *Nature Biotechnology* **14**, 303–308 (1996).
- [4] Seelig, G., Soloveichik, D., Zhang, D. Y. & Winfree, E. Enzyme-free nucleic acid logic circuits. *Science* **314**, 1585–8 (2006).
- [5] Wang, F., Elbaz, J., Orbach, R., Magen, N. & Willner, I. Amplified analysis of DNA by the autonomous assembly of polymers consisting of DNAzyme wires. *Journal of the American Chemical Society* **133**, 17149–51 (2011).
- [6] Shimron, S., Wang, F., Orbach, R. & Willner, I. Amplified detection of DNA through the enzyme-free autonomous assembly of hemin/G-quadruplex DNAzyme nanowires. *Analytical Chemistry* **84**, 1042–8 (2012).
- [7] Li, B., Ellington, A. D. & Chen, X. Rational, modular adaptation of enzyme-free DNA circuits to multiple detection methods. *Nucleic Acids Research* **39**, e110 (2011).

Appendices

Appendix A

Appendix to Chapter 2

A.1 Methods

A.1.1 Preparation of Oligonucleotides

All oligonucleotides used in Chapter 2 were synthesized by Integrated DNA Technologies (Coralville, Iowa) and were either HPLC-purified by IDT, or purified using a denaturing PAGE gel in our lab. Some of the hairpins underwent fluorophore labeling in our lab according to Life Technologies' "Amine-Reactive-Probes" protocol. These hairpins are denoted with parentheses around the fluorophore moiety in Tables A.1–A.9.

A.1.2 Test Tube Preparation

All oligonucleotides were stored at -20°C and brought to room temperature prior to the initiation of the experiment. Extinction-coefficient-based quantification of nucleic acid species was carried out on a spectrophotometer prior to each experiment. Following quantification, all of the nucleic acid strands, apart from long RNA of d2EGFP in Figure 2.8, underwent snap-cooling, which involves a 95°C incubation for 90 seconds, immediately followed by incubation on ice for 30 seconds, followed by an additional 15 minutes, or more, of equilibration at room temperature.

In reactions that contained either scavengers or multiple HCR systems, non-target reagents were expelled onto the sides of the test tube to ensure that all strands can experience 'first contact' with targets at the same time. Reaction tubes contained strands at $1\mu\text{M}$ concentration, but for three exceptions: 1) Scavengers were utilized in $2\mu\text{M}$; 2) the

experiment presented in Section 2.8 contained $2\mu\text{M}$ of each hairpin of HCR systems H^{C} and H^{G} ; and 3) the experiment presented in Section 2.9 utilized $0.715\mu\text{M}$ targets. Test tubes contained $1\times$ PKR buffer (20mM HEPES pH 7.5, 4mM MgCl_2 , 100mM KCl). HCR reactions were carried out at 37°C for 1 hour, other than the reaction that involved detection of long RNA (Section 2.9), which was carried out for 18 hours and 23 minutes. Following incubation at 37°C , $5\times$ loading buffer (50% glycerol, 1xPKR) was added to each reaction mixture, or a fraction of it, and the resulting mixture was run (150V) on a 10 % native TBE gel (Bio-Rad, Hercules, CA).

A.1.3 Imaging

Imaging of all hairpins, other than those which were conjugated to Alexa750, was carried out on an FLA-5100 laser scanner (Fujifilm, Stamford, CT). Unmodified hairpins were stained with SYBR Gold according to manufacturer directions and were imaged using a 473 nm laser and a $530\text{nm} \pm 10$ band-pass filter. Fluorescently-labeled hairpins were excited with either 473, 532, or 635 nm lasers, and their emission was collected using an appropriate choice of 530 ± 10 nm or 570 ± 10 nm band-pass filters or a 665 nm long-pass filter. Imaging of Alexa750-modified hairpins was carried out on Li-Cor's (Lincoln, NE) Odyssey machine, which utilizes a 785nm laser and collects emission above 810nm. When Odyssey-rendered images were superimposed with FLA-5100 images (Figure 2.7 and left panel of Figure 2.4), the TurboReg plugin of Image J was utilized.

A.2 Strand Sequences

Expected RNA strand from T7 *in vitro* transcription of d2EGFP C→G; SNP is marked in red

GGGAGACCCAAGCUGGCUAGCAUGGUGAGCAAGGGCGAGGAGCUGUUCACCG
GGGUGGUGCCCAUCCUGGUCGAGCUGGACGGCGACGUAAACGGCCACAAGUU
CAGCGUGUCCGGCGAGGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCGUG
AAGUUCAUCUGCACCACCGGCAAGCUGCCCGUGCCCUUGGCCACCCUCGUGAC
CACCCUGACCUACGGCGUGCAGUGCUUCAGCCGCUACCCCGACCACAUGAAGC
AGCACGACUUCUUAAGUCCGCCAUGCCCGAAGGCUACGUCCAGGAGCGCAC
CAUCUUCUUAAGGACGACGGCAACUACAAGACCCGCGCCGAGGUGAAGUUC
GAGGGCGACACCCUGGUGAACCGCAUCGAGCUGAAGGGCAUCGACUUCAAGG
AGGACGGCAACAUCCUGGGGCACAAGCUGGAGUACAACUACAACAGCCACAA
CGUCUAUAUCAUGGCCGACAAGCAGAAGAAUGGCAUCAAGGUGAACUUCAAG
AUCCGCCACAACAUCGAGGACGGCAGCGUGCAGCUCGCCGACCACUACCAGC
AGAACACCCCCAUCGGCGACGGCCCCGUGCUGCUGCCCGACAACCACUACCUG
AGCACCCAGUCCGCCUGAGCAAAGACCCCAACGAGAAGCGCGAUCACAUGG
UCCUGCUGGAGUUCGUGACCGCCGCCGGGAUCACUCUCGGCAUGGACGAGCU
GUACAAGAAGCUUAGCCAUGGCUUCCCGCCGGAGGUGGAGGAGCAGGAUGAU
GGCACGCUGCCCAUGUCUUGUGCCAGGAGAGCGGGAUGGACCGUCACCCUG
CAGCCUGUGCUUCUGCUAGGAUCA AUGUGUAGCUUAA

Table A.1: Strands Utilized in Figure 2.1

Material	Name	# (nt)	Sequence (5' to 3')	Stock #
RNA	h1	36	CAGUCCUCCUUUCCAGGAAACUGGAAAGGGAGGA	285
RNA	h2	36	CUGGAAAGGGAGGAACUGUCCUCCUUUCCAGUUUC	286
RNA	T	26	CUGUAAAGCUGGAAAGGGAGGAACUG	193
RNA	T'	26	CUGUAAAGCUGGAAAGGGACGAACUG	194

Table A.2: Strands Utilized in Figure 2.2

Material	Name	# (nt)	Sequence (5' to 3')	Stock #
RNA	h1 slow	36	GAUUUCUCUGUAGCUAGAGACUUCUAGCUACAGAGA	281
RNA	h2 slow	36	UCUAGCUACAGAGAAAUCUCUCUGUAGCUAGAAGUC	282
RNA	h1 medium	36	UCUCUGUAGCUAGACCAACAAAUUGGUCUAGCUACA	500
RNA	h2 medium	36	UUGGUCUAGCUACAGAGAUGUAGCUAGACCAAUUUG	501
RNA	h1 fast	36	CAGAGAAAUCUCGACAGAUCGGAUUUCUCUGUAGC	518
RNA	h2 fast	36	UCUGUCGAGAUUUCUCUGGCUACAGAGAAAUCUCGA	519
RNA	T	36	GAUUUUGGUCUAGCUACAGAGAAAUCUCGAUGGAGU	277
RNA	T'	36	GAUUUUGGUCUAGCUACAGUAAAUCUCGAUGGAGU	278
RNA	T''	36	GAUUUUGGUCUAGCAACAGUGAAAUCUCGAUGGAGU	504

Table A.3: Strands Utilized in Figure 2.3

Material	Name	Panel	# (nt)	Sequence (5' to 3')	Stock #
RNA	h1	A, B	36	CACCACACCACCGGAGACGUGUGGUGUGUGUG	314
RNA	h2	A, B	36	UCUCCGUGUGUGUGUGGUGCACACACCACACCACG	315
RNA	T	A, B	18	CACACACCACACCACAG	316
RNA	T'1	A	18	CAGACACCACACCACAG	317
RNA	T'2	A	18	CACAGACCACACCACAG	318
RNA	T'3	A	18	CACACACCAGACCACAG	319
RNA	T'4	A	18	CACACACCACACCAGAG	320
RNA	T'1	B	18	CATACACCACACCACAG	328
RNA	T'2	B	18	CACAUACCACACCACAG	329
RNA	T'3	B	18	CACACACCAUACCACAG	330
RNA	T'4	B	18	CACACACCACACCAUACG	331
RNA	h1	C, D	36	CUGACUUGACUGACUACUGUCAGUCAAGUCAGGUCA	364
RNA	h2	C, D	36	AGUAGUCAGUCAAGUCAGUGACCUGACUUGACUGAC	365
RNA	T	C, D	18	UGACCUGACUUGACUGAC	363
RNA	T'1	C	18	UUACCUGACUUGACUGAC	366
RNA	T'2	C	18	UGACCUUACUUGACUGAC	367
RNA	T'3	C	18	UGACCUGACUUUACUGAC	368
RNA	T'4	C	18	UGACCUGACUUGACUUAC	369
RNA	T'1	D	18	UGACCUGACUUGACUAAAC	410
RNA	T'2	D	18	UGACCUGACUUAAACUGAC	411
RNA	T'3	D	18	UGACCUAAACUUGACUGAC	412
RNA	T'4	D	18	UAACCUGACUUGACUGAC	413

Table A.4: Strands Utilized in Figure 2.4

Material	Name	Panel	# (nt)	Sequence (5' to 3')	Stock #
RNA	h1 (H ^U)	left	36	CAGAGAAAUCUCGACAAUUCGAGAUUUCUCUGUAGC	449
RNA	h2 (H ^U)	left	36	AUUGUCGAGAUUUCUCUGGCUCACAGAGAAAUCUCGA/iSp9//3AmM0/(Alexa750)	450*
RNA	h1 (H ^A)	left	36	UCACUGUAGCUAGACCAACAAAUUGGUCUAGCUACA/iSp9//3AlexF546N/	435
RNA	h2 (H ^A)	left	36	UUGGUCUAGCUACAGUGAUGUAGCUAGACCAAUUUG	225
RNA	T ^A	left	36	GAUUUUGGUCUAGCUACAGAGAAAUCUCGAUGGAGU	277
RNA	T ^U	left	36	GAUUUUGGUCUAGCUACAGUGAAAUCUCGAUGGAGU	278
RNA	h1 (H ^A)	middle	36	UCCACAGAAACAUACUCCUCUUGGAGUAUGUUUCUG	423
RNA	h2 (H ^A)	middle	36	/5A1ex647N//iSp9/GGAGUAUGUUUCUGUGGACAGAAACAUACUCCAAGA	424
RNA	h1 (H ^C)	middle	38	GUAUGUGUCUGUGGAAGAUAUCCACAGACACAUCUCCA	447
RNA	h2 (H ^C)	middle	38	UUCUCCACAGACACAUCUGGAGUAUGUGUCUGUGGA/iSp9//3AmM0/(Alexa532)	461*
RNA	T ^U	middle	36	UUGUUUUAAAUAUGGAGUAUGUUUCUGUGGAGAC	269
RNA	T ^G	middle	36	UUGUUUUAAAUAUGGAGUAUGUGUCUGUGGAGAC	270
RNA	h1 (H ^C)	right	36	GAGGAACUGGUGUAAACAACACACAGUCCUCCUU	489
RNA	h2 (H ^C)	right	36	UGUUACACCAGUCCUCAAGGGAGGAACUGGUGUA/iSp9//3AmM0/(Cy5)	490*
RNA	h1 (H ^G)	right	36	CAGUUCGUCCUUUCCAGGAAACUGGAAAGGGACGA/iSp9//3Cy3Sp/	295
RNA	h2 (H ^G)	right	36	CUGGAAAGGGACGAACUGUCGUCCUUUCCAGUUUC	264
RNA	T ^C	right	36	CUGUAAAGCUGGAAAGGGACGAACUGGUGUAAUGAU	487
RNA	T ^G	right	36	CUGUAAAGCUGGAAAGGGAGGAACUGGUGUAAUGAU	488

Table A.5: Strands Utilized in Figure 2.5.

Material	Name	# (nt)	Sequence (5' to 3')	Stock #
RNA	h1	36	UCUCUGUAGCUAGACCAACAAAUUGGUCUAGCUACA	500
RNA	h2	36	UUGGUCUAGCUACAGAGAUGUAGCUAGACCAAUUUG	501
RNA	T ^A	26	GAUUUUGGUCUAGCUACAGAGAAAUC	189
RNA	T ^G	26	GAUUUUGGUCUAGCUACAGGAAAUC	371
RNA	S ^C	11	UUUCCUGUAG	389

Table A.6: Strands Utilized in Figure 2.6.

Material	Name	Panel	# (nt)	Sequence (5' to 3')	Stock #
RNA	h1	1	36	UCUCUGUAGCUAGACCAACAAUUGGUCUAGCUACA	191
RNA	h2	1	36	UUGGUCUAGCUACAGAGAUGUAGCUAGACCAAUUUG	192
RNA	h1	2	36	UCACUGUAGCUAGACCAACAAUUGGUCUAGCUACA	502
RNA	h2	2	36	UUGGUCUAGCUACAGUGAUGUAGCUAGACCAAUUUG	503
RNA	T ^A	1, 2	26	GAUUUUGGUCUAGCUACAGAGAAAUC	189
RNA	T ^C	1, 2	26	GAUUUUGGUCUAGCUACAGCGAAAUC	370
RNA	T ^G	1, 2	26	GAUUUUGGUCUAGCUACAGGGAAAUC	371
RNA	T ^U	1, 2	26	GAUUUUGGUCUAGCUACAGAGAAAUC	190
RNA	S ^A	1	12	AUUUCACUGUAG	496
RNA	S ^C	1, 2	11	UUUCCUGUAG	389
RNA	S ^U	2	12	AUUUCUCUGUAG	499
RNA	h1	3	36	UCCACAGAAACAUAUCCUCUUGGAGUAUGUUUCUG	423
RNA	h2	3	36	GGAGUAUGUUUCUGUGGACAGAAACAUAUCCCAAGA	509
RNA	h1	4	38	GUAUGUGUCUGUGGAAGAAUCCACAGACACAUACUCUA	521
RNA	h2	5	38	UUCUCCACAGACACAUACUGGAGUAUGUGUCUGUGGA	448
RNA	T ^A	3, 4	36	UUGGUUUUAAAUAUGGAGUAUGUAUCUGUGGAGAC	497
RNA	T ^C	3, 4	36	UUGGUUUUAAAUAUGGAGUAUGUCUCUGUGGAGAC	498
RNA	T ^G	3, 4	36	UUGGUUUUAAAUAUGGAGUAUGUCUCUGUGGAGAC	270
RNA	T ^U	3, 4	36	UUGGUUUUAAAUAUGGAGUAUGUUUCUGUGGAGAC	269
RNA	S ^U	3	12	ACAGAUACAUAUAC	512
RNA	S ^C	3	11	CAGACACAUAUAC	513
RNA	h1	5	36	CAGUCCUCCUUUCCAGGAAACUGGAAAGGGAGGA	285
RNA	h2	5	36	CUGGAAAGGGAGGAACUGUCCUCCUUUCCAGUUUC	286
RNA	h1	6	36	CAGUUCGUCCUUUCCAGGAAACUGGAAAGGGACGA	287
RNA	h2	6	36	CUGGAAAGGGACGAACUGUCGUCCUUUCCAGUUUC	288
RNA	T ^A	5, 6	26	CUGUAAAAGCUGGAAAGGGGAAGAACUG	374
RNA	T ^C	5, 6	26	CUGUAAAAGCUGGAAAGGGGACGAACUG	194
RNA	T ^G	5, 6	26	CUGUAAAAGCUGGAAAGGGGAGAACUG	193
RNA	T ^U	5, 6	26	CUGUAAAAGCUGGAAAGGGGAUGAACUG	375

Table A.7: Strands Utilized in Figure 2.7

Material	Name	# (nt)	Sequence (5' to 3')	Stock #
RNA	h1 (H ^A)	36	/5A1ex750N//iSp9/AAGGGAUGAACUGGUCUACCGAUUCAUCCCUUCCA	492
RNA	h2 (H ^A)	36	AUGACCAGUUCAUCCCUUUGGAAAGGGAUGAACUGG	493
RNA	h1 (H ^C)	36	GAGGAACUGGUGUAAACAUAACCCAGUUCUCCUU	489
RNA	h2 (H ^C)	36	UGUUUACACCAGUUCUCAAGGGAGGAACUGGUGUA/iSp9//3AmM0/(Cy5)	517*
RNA	h1 (H ^G)	36	CAGUUCGUCCUUUCCAGGAAACUGGAAAGGGACGA/iSp9//3Cy3Sp/	295
RNA	h2 (H ^G)	36	CUGGAAAGGGACGAACUGUCGUCCUUUCCAGUUUC	264
RNA	h1 (H ^U)	38	CACUAGUUCUCCCUUCCUCAAGGAAAGGGGAAGAACU/iSp9//3AmM0/(Alexa488)	516*
RNA	h2 (H ^U)	38	GGAAAGGGAAGAACUAGUGAGUUCUCCCUUCCUUGA	486
RNA	T ^A	36	CUGUAAAAGCUGGAAAGGGGAAGAACUGGUGUAAUGAU	484
RNA	T ^C	36	CUGUAAAAGCUGGAAAGGGGACGAACUGGUGUAAUGAU	487
RNA	T ^G	36	CUGUAAAAGCUGGAAAGGGGAGAACUGGUGUAAUGAU	488
RNA	T ^U	36	CUGUAAAAGCUGGAAAGGGGAUGAACUGGUGUAAUGAU	491

Table A.8: Strands Utilized in Figure 2.8.

Material	Name	# (nt)	Sequence (5' to 3')	Stock #
RNA	h1	36	AGCUGACCCUGAAGGACAUUACGGGUCAGCUUGCC	441
RNA	h2	36	UGUCCU/iCy3/UCAGGGUCAGCUGGCAAGCUGACCCUGAAG	442
RNA	h1	36	ACCGUGAAGUUCUAACCAUGAACUUCACGGUCAGC	432
RNA	h2	36	GGUUUAUGAACUUCACGGGUCUGACCGUGAAGUUCAU/iSp9//3Cy5Sp/	433

Appendix B

Appendix to Chapter 3

B.1 Methods

B.1.1 Preparation of Oligonucleotides

With the exception of one strand (#359* from Table B.2)¹, all oligonucleotides used in Chapter 3 were synthesized and HPLC-purified by Integrated DNA Technologies (Coralville, Iowa).

B.1.2 Experimental Procedure

Oligonucleotides were stored at -20°C and brought to room-temperature prior to the initiation of the experiment. Extinction-coefficient-based quantification of nucleic acid species was carried out on a spectrophotometer prior to each experiment. Following quantification, all of the nucleic acid strands underwent snap-cooling, a process that involves 95°C incubation for 90 seconds, immediately followed by 30 seconds of incubation on ice and an additional 15 minutes, or more, equilibration at room temperature. The hairpins were snap-cooled in $1\times$ PKR buffer (20mM HEPES pH 7.5, 4mM MgCl_2 , 100mM KCl).

In the experiments presented in Chapter 3, all strands were in $1\mu\text{M}$ final concentration, with the exception of the RNA-based Quenched HCR studies, which utilized $0.8\mu\text{M}$ final concentration of each strand. The final volume in the sensitivity-studies was $5\mu\text{L}$; the final volume in the multiplexing studies as well as in the RNA-based Quenched HCR studies was

¹Strand 359* was ligated using NEB's RNA ligase I ligation protocol; upon completion of the ligation reaction, this strand was ethanol-precipitated.

10 μ L; the final volume in all other studies in Chapter 3 was 8 μ L.

All non-target nucleic acid strands were expelled onto the sides of the test tube to ensure that all strands experience ‘first contact’ with target at the same time. Reagent mixing entailed rapid test tube flicking, followed by brief centrifugation. Once in a homogenous mixture, the test tubes were placed in Bio-Rad’s CFX-96 Real-Time PCR machine, and incubated at 25°C. Quenched HCR reactions were run for 12 hours, with the exception of the RNA-based Quenched HCR study described in Section B.4, which was run for 2 hours. During HCR reactions, the lid covering the test tubes was maintained at 105°C.

Following the optical recording of Quenched HCR, test tubes were removed from the real-time PCR machine and analyzed on a gel. To this end, 5 \times loading buffer (50% glycerol, 1 \times PKR) was added to each test tube. The resulting mixtures were run (150V) on a 10% native TBE gel (Bio-Rad, Hercules, CA).

All figures presented in Chapter 3, other than Figure 3.3², provide background-subtracted data. Background is defined as the average of three, or four, measurements that took place prior to the addition of test tubes to the real-time PCR machine.

²Multiplexing Figure 3.3 contains background-subtracted data, normalized to the maximum recorded in each of the two channels.

B.2 Background-subtracted Data of Quenched HCR Multiplexing

In this section we provide additional data (Figure B.1) for Quenched HCR multiplexing. In Figure 3.3 we provided normalized fluorescence data for two test tubes: the test tube containing $T^2+Q^2+Q^3$ and the test tube containing $T^3+Q^2+Q^3$. Here, we provide background-subtracted data for the same two test tubes, in addition to data for the test tube containing $T^1+T^2+Q^2+Q^3$. The term “background” is defined in Section B.1. Plotted data represent mean and standard deviation of two experiments.

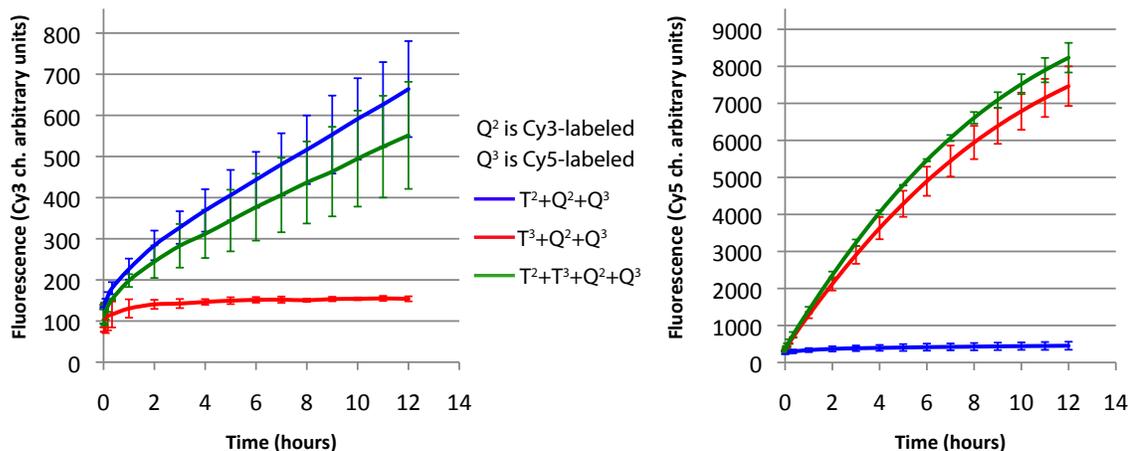


Figure B.1: Background-subtracted data of multiplexing with Quenched HCR. Cy3-labeled Quenched HCR system designed to detect JAK2 G→T cancer marker, and Cy5-labeled Quenched HCR system designed to detect PTEN C→G cancer marker were mixed with a JAK2 G→T-containing test tube, a PTEN C→G-containing test tube, and a test tube containing both of these cancer markers. Plotted data represent mean and standard deviation of two experiments.

B.3 Sensitivity of Quenched HCR

To demonstrate the sensitivity of Quenched HCR, $1\mu\text{M}$, 100nM , and 10nM T^3 were each incubated with Q^3 (B.2). Sensitivity of up to and including 100nM was repeatedly obtained. At 10nM T^3 , some experiments provided fluorescent signal which was within a standard deviation of the background traces.

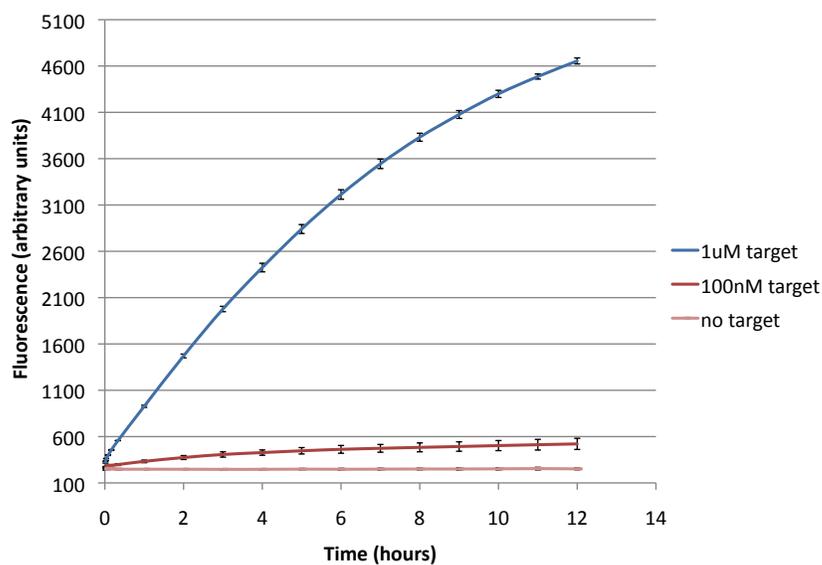


Figure B.2: Sensitivity of detection of PTEN C \rightarrow G gene marker (T^3) via 4-hairpin-periodicity Quenched HCR (Q^3). Plotted data represent mean and standard deviation of three experiments. 100nM sensitivity was obtained in all three experiments; 10nM target, however, was within the error of background in some experimental runs.

B.4 RNA-based Quenched HCR

In addition to testing a variety of DNA-based Quenched HCR systems, we also wanted to demonstrate that Quenched HCR can be obtained with an RNA-based, 4-hairpin-periodicity Quenched HCR system. This system consists of domain sizes that are equal to most of the HCR systems that were used in Chapter 2, specifically, 4-nt toehold, 14-nt stem, and 4-nt loop. The fluorescence curves that were produced with this system are shown in figure B.3.

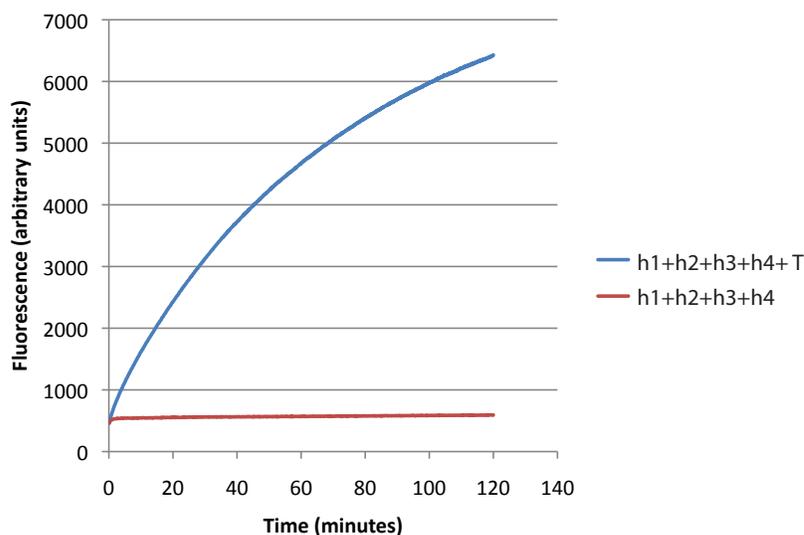


Figure B.3: Detection of an RNA target with an RNA-based Quenched HCR. 4-hairpin-periodicity RNA-based Quenched HCR was incubated with an RNA target (blue curve), as well as without a target (red curve). Target-free fluorescent trace remains almost flat throughout the experiment. Target-containing fluorescent trace is monotonically increasing throughout the experiment.

B.5 Fast Target Detection with Quenched HCR

Fluorescent traces that demonstrate successful implementation of Quenched HCR are described in Section 3.3, Figure 3.2. Since we sought to monitor the polymerization of Quenched HCR from $\sim t_0$, we engineered the systems described in Section 3.3 to have long polymerization time scales. All three Quenched HCR systems included in Figure 3.2 exhibit monotonically increasing fluorescent signal throughout the 12 hours in which Quenched HCR was monitored.

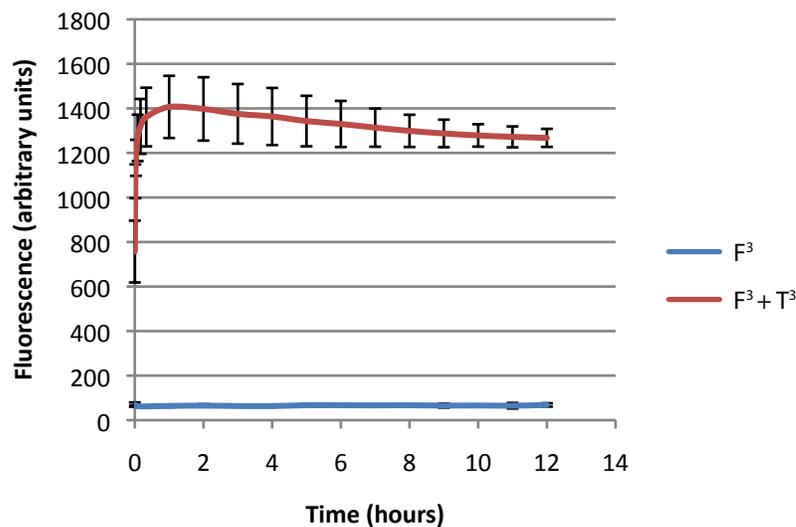


Figure B.4: 2-hairpin-periodicity Quenched HCR exhibits fast polymerization when in the presence of T^3 , as desired. Plotted data represent mean and standard deviation of two experiments. The engineering of fast-to-polymerize Quenched HCR systems employed the ΔG tuning principle, which predicts that the more negative the value of ΔG of addition of hairpin toehold domains to their reverse-complements, as can be calculated using thermodynamic principles, the faster the formation of HCR polymers. Details about this principle are presented in Section 2.3.

Here, we demonstrate that Quenched HCR systems can also undergo fast polymerization. To design a fast-to-polymerize Quenched HCR system, we implemented ΔG tuning, which is described in Section 2.3. As demonstrated in Figure B.4, we were able to engineer a Quenched HCR system (F^3) that reaches a maximum fluorescence signal in ~ 1 hour. The steepness of the curve around the first time point demonstrates that the majority of polymerization occurs in the first few minutes of the reaction. F^3 and $F^3 + T^3$ do not intersect with the y-axis at the same value due to a short time interval, ~ 2 minutes, between mixing the reaction tubes and the first measurement on the real-time PCR machine. This fast-to-polymerize Quenched HCR system entailed long-toehold hairpins and, hence, was realized in 2-hairpin-periodicity (Figure 3.1).

B.6 Strand Sequences

Table B.1: Strands Utilized in Chapter 3

Material	Name	# (nt)	Sequence (5' to 3')	Stock #
DNA	h1 (P ¹ , Q ¹)	50	/5IAbRQ/GCTACAGAGAAATCTCGATATGAGGATCGAGATTCTCTGTAGC/iCy5/TAGACC	582
DNA	h2 (Q ¹)	50	CCTCAT ATCGAGATTTCTCTGTAGCCAGATGTCTACAGAGAAATCTCGAT	532
DNA	h3 (Q ¹)	50	GCTACAGAGAAATCTCGATTACTCCATCGAGATTTCTCTGTAGC ATCTGG	533
DNA	h4 (Q ¹)	50	GGAGTAATCGAGATTTCTCTGTAGCGGTCTAGCTACAGAGAAATCTCGAT	534
DNA	"h2" (P ¹)	50	CCTCATATCGAGATTTCTCTGTAGCGGTCTAGCTACAGAGAAATCTCGAT	535
DNA	h1 (P ² , Q ²)	50	/5IAbkFQ/AGTATGTTTCTGTGGAGACTGACGTGTCTCCACAGAAACATACT/iCy3/CCATAA	536
DNA	h2 (Q ²)	50	ACGTGAGTCTCCACAGAAACATACTAATACCAGTATGTTTCTGTGGAGAC	537
DNA	h3 (Q ²)	50	AGTATGTTTCTGTGGAGACACAGCAGTCTCCACAGAAACATACTGTGATT	538
DNA	h4 (Q ²)	50	TGCTGTGTCTCCACAGAAACATACTTTATGGAGTATGTTTCTGTGGAGAC	539
DNA	"h2" (P ²)	50	ACGTGAGTCTCCACAGAAACATACTTTATGGAGTATGTTTCTGTGGAGAC	540
DNA	h1 (P ³ , Q ³)	48	/5IAbRQ/GAAAGGGAGGAAGTGGTGTGACAGTCCAGTTCTCCCTTTC/iCy5/CAGCTT	541
DNA	h2 (Q ³)	48	ACTGCACACCAAGTTCCCTTTCCTCGAAAGGGAGGAAGTGGT	542
DNA	h3 (Q ³)	48	GAAAGGGAGGAAGTGGTGAACGACACACCAAGTTCCCTTTCGAGGAA	543
DNA	h4 (Q ³)	48	TGTCGTCCAGTTCCTCCCTTTCAGCTGAAAGGGAGGAAGTGGT	544
DNA	"h2" (P ³)	48	ACTGCACACCAAGTTCCCTTTCAGCTGAAAGGGAGGAAGTGGT	545
DNA	h1 fast (Q ³)	72	/5IAbRQ/GAAAGGGAGGAAGTGGTGAATGATGACTAGTACAATCATTACACCAAGTTCCCTTTC/i5-TAMK/AGCTTTACAG	556
DNA	h2 fast (Q ³)	72	TGTAAGTGTGATCATTACACCAAGTTCCCTTTCCTGTAAGCTGAAAGGGAGGAAGTGGTGAATGAT	557
RNA	T1	36	GAUUUUGGUCUAGUCACAGAGAAAUCUGGAUGGAGU	277
RNA	T2	36	UUGUUUUUUUUUUGGAGUAUGUUUCUGUGGAGAC	269
RNA	T3	36	CUGUAAAGCUGGAAAGGGAGGAACUGGUGUAAUGAU	488

Table B.2: Additional Strands Utilized in Appendix B

Material	Name	# (nt)	Sequence (5' to 3')	Stock #
DNA	h1 fast (F ³)	72	/5IAbRQ/GAAAGGGAGGAAGTGGTGAATGATGACTAGTACAATCATTACACCAAGTTCCCTTTC/i5-TAMK/AGCTTTACAG	556
DNA	h2 fast (F ³)	72	TGTAAGTGTGATCATTACACCAAGTTCCCTTTCCTGTAAGCTGAAAGGGAGGAAGTGGTGAATGAT	557
RNA	h1	36	GAACCCUUCUUUUGUCUGCAUUAAGAAGGUUCUGCU	B3
RNA	h2	36	CAGACAUUAAGAAGGUUCUCUGUAACCCUUCUUUUG	305
RNA	h3	36	/5IAbRQ/GAACCCUUCUUUUGAGACCAUUAAGAAGGUU/iCy5/CACGA	359*
RNA	h4	36	GUCUCAUUAAGAAGGUUCAGCAGAACCCUUCUUUUG	307

Appendix C

Appendix to Chapter 4

C.1 Methods

Test tubes in Chapter 4 contained a final volume of $8\mu\text{M}$, consisting of $1\mu\text{M}$ of each strand in $1 \times \text{PKR}$, with the exception of the long RNA target PTEN C→G, which was utilized in Section 4.5, and was at a final concentration of $0.9375\mu\text{M}$. Long RNA of PTEN C→G was not snap-cooled. Other than the presence of conditional probes in Chapter 4, the methods followed in this chapter are identical to the methods that are described in Section B.1.

C.2 Strand Sequences

Expected RNA strand from T7 *in vitro* transcription of PTEN C→G; SNP is marked in red

```

AUGACAGCCAUCAUCAAGAGAUCGUUAGCAGAAACAAAAGGAGAUUAUCAAG
AGGAUGGAUUCGACUUAGACUUGACCUAUUUUAUCCAAACAUUAUUGCUAU
GGGAUUUCCUGCAGAAAGACUUGAAGGCGUAUACAGGAACAAUUAUGAUGAU
GUAGUAAGGUUUUUGGAUUCAAAGCAUAAAAACCAUACAAGAUUAACAAUC
UUUGUGCUGAAAGACAUUAUGACACCGCCAAAUUUAAUUGCAGAGUUGCACA
AUAUCCUUUUGAAGACCAUAACCCACCACAGCUAGAACUUAUCAAAACCCUUU
GUGAAGAUCUUGACCAAUGGCUAAGUGAAGAUGACAAUCAUGUUGCAGCAAU
UCACUGUAAAGCUGGAAAGGGAGGGAACUGGUGUAAUGAUUAUGUGCAUUAUU
AUUACAUCGGGGCAAUUUUUAAAGGCACAAGAGGCCCUAGAUUUCUAUGGG
GAAGUAAGGACCAGAGACAAAAGGGAGUAACUAUCCCAGUCAGAGGCGCU
AUGUGUAUUUUUAUAGCUACCUGUUAAGAAUCAUCUGGAUUUAUAGACCAGU
GGCACUGUUGUUUCACAAGAUGAUGUUUGAAACUAUUCCAAUGUUCAGUGGC
GGAACUUGCAAUCCUCAGUUUGUGGUCUGCCAGCUAAAGGUGAAGAUUAUU
CCUCCAAUUCAGGACCCACACGACGGGAAGACAAGUUCAUGUACUUUGAGUU
CCCUCAGCCGUUACCUGUGUGUGGUGAUUAUCAAGUAGAGUUCUCCACAAA
CAGAACAAGAUGCUAAAAAAGGACAAAUGUUUCACUUUUGGGUAAAUACAU

```

UCUUCAUACCAGGACCAGAGGAAACCUCAGAAAAAGUAGAAAAUGGAAGUCU
AUGUGAUCAAGAAAUCGAUAGCAUUUGCAGUAUAGAGCGUGCAGAUAAUGAC
AAGGAAUAUCUAGUACUUACUUUACAAAAAUGAUCUUGACAAAGCAAUA
AAGACAAAGCCAACCGAUACUUUCUCCAAAUUUUAAAGGUGAAGCUGUACUU
CACAAAAACAGUAGAGGAGCCGUCAAAUCCAGAGGCUAGCAGUUCAACUUCU
GUAACACCAGAUGUUAGUGACAAUGAACCUGAUCAUUUAUAGAUUUCUGACA
CCACUGACUCUGAUCCAGAGAAUGAACCUUUUGAUGAAGAUACAGCAUACACA
AAUUACAAAAGUCUGA

Table C.1: Strands Utilized in Chapter 4

Material	Name	# (nt)	Sequence (5' to 3')	Stock #
DNA	h1 (Q ¹)	50	/5IABkFQ/GCTACAGAGAAATCTCGATATGAGGATCGAGATTTCTCTGTAGC/i6-FAMK/TAGACC	531
DNA	h2 (Q ¹)	50	CCTCAT ATCGAGATTTCTGTAGCCCAGATGCTACAGAGAAATCTCGAT	532
DNA	h3 (Q ¹)	50	GCTACAGAGAAATCTCGATTACTCCATCGAGATTTCTCTGTAGC ATCTGG	533
DNA	h4 (Q ¹)	50	GGAGTAATCGAGATTTCTGTAGCGGTCTAGCTACAGAGAAATCTCGAT	534
DNA	h1 (Q ²)	50	/5IABkFQ/AGTATGTTTCTGTGGAGACTGACGTGTCCACAGAAACATACT/iCy3/ CCATAA	536
DNA	h2 (Q ²)	50	ACGTCAGTCTCCACAGAAACATACTAATACCAGTATGTTTCTGTGGAGAC	537
DNA	h3 (Q ²)	50	AGTATGTTTCTGTGGAGACACAGCAGTCTCCACAGAAACATACTGGTATT	538
DNA	h4 (Q ²)	50	TGCTGTGTCTCCACAGAAACATACTTTATGGAGTATGTTTCTGTGGAGAC	539
DNA	h1 (Q ³)	48	/5IAbRQ/GAAAGGGAGGAACTGGTGTGACAGTACCAGTTCCTCCCTTTC/iCy5/CAGCTT	541
DNA	h2 (Q ³)	48	ACTGCACACCAGTTCCTCCCTTTCCTCCGAAAGGGAGGAACTGGTG	542
DNA	h3 (Q ³)	48	GAAAGGGAGGAACTGGTGTGACGACACACCAGTTCCTCCCTTTCGAGGAA	543
DNA	h4 (Q ³)	48	TGTCGTCACCAGTTCCTCCCTTTCAGCTGGAAAGGGAGGAACTGGTG	544
RNA	T ¹	36	GAUUUUGUCUAGCUACAGAGAAUUCUGAUGGAGU	277
RNA	T ²	36	UUGUUUUAAAUAUGGAGUAUGUUUCUGGAGAC	269
RNA	T ³	36	CUGUAAAAGCUGGAAAGGGAGGAAACUGGUGUAUUGAU	488
RNA	T ⁴	31	CAAGCUGACCCUGAAGUUAUCUGCACCACC	173
DNA	CP ^{1,2} (L)	74	GTCTCCACAGAAACATACTCCATAATTTCTCTGTAGCTAGACCAAAATCTTATGGAGTATGTTTCTGTGGAGAC	564
DNA	CP ^{1,2} (M)	69	CACAGAAACATACTCCATAATTTCTGTAGCTAGACCAAAATCTTATGGAGTATGTTTCTGTGGAGAC	565
DNA	CP ^{1,2} (S)	64	AAACATACTCCATAATTTCTGTAGCTAGACCAAAATCTTATGGAGTATGTTTCTGTGGAGAC	566
DNA	CP ^{3,2} (L)	72	GTCTCCACAGAAACATACTCCATAATTACACCAGTTCCTCCCTTTCCTTATGGAGTATGTTTCTGTGGAGAC	573
DNA	CP ^{3,2} (M)	67	CACAGAAACATACTCCATAATTACACCAGTTCCTCCCTTTCCTTATGGAGTATGTTTCTGTGGAGAC	574
DNA	CP ^{3,2} (S)	62	AAACATACTCCATAATTACACCAGTTCCTCCCTTTCCTTATGGAGTATGTTTCTGTGGAGAC	575
DNA	CP ^{1,3} (L)	72	CACCAGTTCCTCCCTTTCAGCTTTTCTCTGTAGCTAGACCAAAATCAAGCTGGAAAGGGAGGAACTGGTG	546
DNA	CP ^{1,3} (M)	67	GTTCTCCCTTTCAGCTTTTCTCTGTAGCTAGACCAAAATCAAGCTGGAAAGGGAGGAACTGGTG	547
DNA	CP ^{1,3} (S)	62	TCCCTTTCAGCTTTTCTCTGTAGCTAGACCAAAATCAAGCTGGAAAGGGAGGAACTGGTG	548
DNA	CP ^{2,3} (L)	67	CACCAGTTCCTCCCTTTCAGCTTCTCCACAGAAACATACTCCAAGCTGGAAAGGGAGGAACTGGTG	549
DNA	CP ^{2,3} (M)	62	GTTCTCCCTTTCAGCTTCTCCACAGAAACATACTCCAAGCTGGAAAGGGAGGAACTGGTG	550
DNA	CP ^{2,3} (S)	57	TCCCTTTCAGCTTCTCCACAGAAACATACTCCAAGCTGGAAAGGGAGGAACTGGTG	551
DNA	CP ^{4,2} (S)	71	AAACATACTCCATAAAGTGGTGCAGATGAATTCAGGGTCAGCTTGTATGGAGTATGTTTCTGTGGAGAC	581