

A TRIPARTILE BIOSENSOR FOR REAL-TIME SNSs
DETECTION IN DNA HAIRPIN MOTIF

by

CAMHA NGUYEN

A thesis submitted in partial fulfillment of the requirements
for the Honors in the Major Program in Molecular Biology
in the College of Medicine
and in The Burnett Honors College
at the University of Central Florida
Orlando, Florida

Spring Term 2011

Thesis Chair: Dr. Dmitry Kolpashchikov

ABSTRACT

The hybridization between two complementary strands of nucleic acid is the basis for a number of applications in DNA and RNA analysis, including in vivo RNA monitoring, microarrays, SNPs detection, and so on. The short oligonucleotide probes form Watson/Crick base pairs (A -T and G – C) with the analyzed nucleic acid. Molecular beacon (MB) probe is one of the most advantageous tools for nucleic acid analysis in real-time. A traditional MB probe consists of a DNA strand folded in hairpin motif with a fluorophore attached to the 5' end and a quencher attached to the 3' end. The loop segment is complementary to the analytes. Upon hybridization to a complementary single-stranded nucleic acid, MB probe switches to the elongated conformation, which separates the fluorophore from the quencher, resulting in high fluorescence signal. However, DNA or RNA folded in hairpin motifs are difficult to analyze by a conventional MB probes. Inefficient formation of the duplex between the secondary analyte and the MB probe results in low or undetectable fluorescent signal.

In this project, we developed a tripartite probe consisting of one MB probe and two adaptor strands to genotype single nucleotide polymorphism (SNPs) in DNA hairpin motifs in real-time fluorescent assays. Each adaptor strand contains a fragment complementary to the analyte and a fragment complementary to an MB probe. One adaptor strand hybridizes to the analyte and unwinds its secondary structure, and the other strand forms stable complex only with the fully complementary analyte sequence. The tri-component probe promises to simplify nucleic acid analysis at ambient temperatures in such application as in vivo RNA monitoring and isothermal detection of specific DNA/RNA targets.

DEDICATIONS

For my grandparents and my parents, thank you for your encouragement and love.

For my mentor, Dr. Kolpashchikov thank you for your guidance and encouraging me to achieve my highest goals.

And especially, for my mother Doantrinh Quach, my best friend. You created me and gave me strength to make my dream come true.

ACKNOWLEDGMENTS

I express sincere thanks and gratitude to my committee members, who have supported me to finish my project with their guidance, wisdom, and experience. Special thanks go to my thesis chair, Dr. Dmitry Kolpashchikov, for his guidance and his great patience. Thanks to Dr. Yulia Gerasimova for her instruction and constructive feedback on this projects. Thanks also to my friends, Jeffrey Grimes and Evan Cornett thank you for your instructions and keeping my spirits high throughout this process.

This research was funded by UCF – Office of Undergraduate Research, OUR grant Spring 2011.

TABLE OF CONTENTS

Introduction.....	1
Hypothesis and Specific Aims	4
Experimental Designs	7
Results.....	14
The tricomponent probe can analyze DNA hairpins disregarding to the stem length.	14
Selectivity of the tricomponent probe in detecting 3 SNPs in DNA hairpin	19
Limits of detection of the tri-component probe in analysis of all 7 analytes ..	Error! Bookmark not defined.
Discussions	21
Conclusion	24
Reference	25

LIST OF FIGURES

Figure 1. MB probe fluoresces when hybridizing to a complementary nucleic acid sequence.	2
Figure 2. Molecular beacon probe for the analysis of DNA analyte	5
Figure 3. Structure of T13 analyte and its analogs used in this study	8
Figure 4. The hybridization reaction between the tricomponent probe and the target T1	9
Figure 5. Comparison the detection efficiency of the MB probe approach with the tricomponent probe approach.....	14
Figure 6. Detection limits of tri-component probe.	17
Figure 7 Selectivity of the tricomponent probe approach in detecting 3 SNPs.....	19

LIST OF TABLES

Table 1. Fluorescent assay conditions.....	10
Table 2. Oligonucleotides used in the study.....	12-13

Introduction

The human genome consists of more than 3 billion base pairs. Single nucleotide polymorphism (SNP) is the most common genetic variation between two individuals. Two individuals have one base pair difference for every 1000-2000 base pairs (1). According to “A Haplotype Map of Human Genome”, there are more than 3 million SNPs validated; however, the actual number of SNPs in the human genome may range up to 10 million (2). Some SNPs are believed to cause diseased conditions such as cystic fibrosis (3), muscular dystrophy (4), insulin dependent diabetes mellitus (5), breast cancer (6) and others. Therefore, the International Hap Map Project started in 2002 is to genotype the SNPs and create a SNPs database for human genome. The SNPs data is important not only to give us a better understanding about SNPs role in population epidemiology but also to create an extensive source for other fields such as SNPs “finger print” in forensic science (7), pharmacogenomics drug development (7) as well as molecular analysis of infectious diseases (8-10).

The hybridization between two complementary strands of nucleic acid is the basis for many applications in DNA and RNA analysis. The probes form Watson-Crick base pairs (A-T and G-C) with the DNA or RNA analytes, which is very simple for the unconstrained single-stranded DNA (ssDNA) or RNA. Real-time detection techniques allow instant read out of fluorescent signal at ambient temperature after the probe is hybridized with the analyte, thus avoiding the need to separate the probe-analyte hybrid from the unbound probe (11-14).

Molecular beacon (MB) probe is a powerful tool for real-time nucleic acid analysis (15, 16). A traditional MB probe consists of a strand of oligonucleotide folded in hairpin motif, a fluorophore at the 5' end, and a quencher at the 3' end (Fig. 1, left). Upon hybridization to a complementary single-stranded DNA or RNA, MB probe switches to the elongated

conformation, when the fluorophore is separated from the quencher, thus resulting in high fluorescence signal (Fig. 1, right). The simplicity and the elegance of MB design have made this probe a popular tool for nucleic acid analysis (17). MB probes have been extensively used in real-time PCR assays (18). In addition MB probes were suggested for RNA monitoring in living cells (19, 20). Importantly, not only does the stem-loop structure bring the fluorophore close to the quencher, but it also improves the probe specificity by increasing the energy required to separate the fluorophore and quencher which will only occur as the MB probe-analyte duplex is matched. MB probes discriminate between two nucleic acid sequences that differ by a single nucleotide in a wider temperature range than linear hybridization probes do (21). These properties are very important for SNP genotyping.

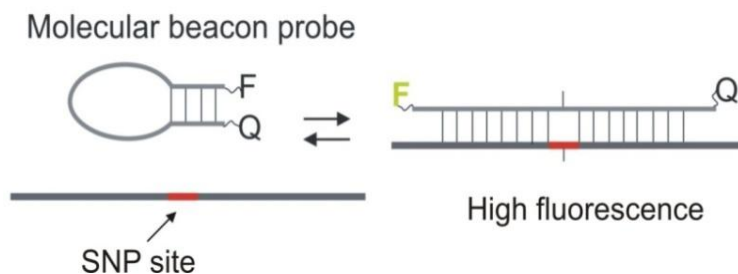


Figure 1. MB probe fluoresces when hybridizing to a complementary nucleic acid sequence.

However, the ssDNA and naturally occurring RNA often fold into secondary and tertiary structures due to the self-complementarities in their sequences. The analysis of such folded analyte is often complicated because the sequence of interest is shielded in the secondary structure and the thermodynamic cost to unwind the complex is high and sometimes is unfavorable for the hybridization between the probe and the nucleic acid analyte (22). As a result, the probe sensitivity is greatly reduced, which sometimes lead to the inability of the probe to detect the SNPs

There are several approaches to analyze the SNPs in secondary-structure forming DNA, particularly DNA folded in hairpin loop, which is very thermodynamically stable (23). According to Kushon, peptide nucleic acid (PNA) probes can be used to analyze the DNA hairpin motif. PNA molecule resembles a ssDNA, but it has an *N*-(2-aminoethyl)glycine backbone instead of a phosphodiester backbone. PNA molecule hybridizes to complementary strand based on the Watson-Crick rule (24). Kushon's group demonstrated the ability of PNA probe to unwind the DNA hairpins and to form a stable complex between the DNA analyte and the PNA probes. PNA probes lack the electrostatic repulsion due to their neutral polyamide backbones, which allow favorable free energy upon their hybridization with the DNA analyte (25). Owing to PNA's neutral polyamide backbone, the PNA strand, however, is more hydrophobic than a regular DNA strand; as a result, it is difficult to synthesize and purify long chain of PNA. Furthermore, the long PNA strands tend to aggregate due to the hydrophobic nature of the polyamide backbone, which prevents the strands from dissolving to interact with the DNA analytes. Therefore, the complication in the design and purification of the PNA probes using High Pressure Liquid Chromatography (HPLC) increase the cost of multiplex analysis. In addition, PNA probes can't be analyzed the PNA using polyacrylamide gel electrophoresis (PAGE), while this method is a powerful tool to purify and characterize DNA or RNA molecules.

Hypothesis and Specific Aims

In this project we hypothesized that traditional MB probe was unable to directly analyze the DNA targets that were folded in a stable secondary structure. We suggested a novel approach that allows MB-probe to be used for the analysis of such structurally constrained analytes. Figure 2 demonstrates MB probe hybridization to a stem-loop folded nucleic acid analyte. The equilibrium of this reaction may be shifted toward dissociated MB-probe and the analyte, since this state is stabilized by intramolecular hybridization of the two oligonucleotides. Increase in concentrations of the two oligonucleotides would shift the reaction toward hybrid formation. However, in many analytical applications analysis of low concentrations (1-100 nM) of nucleic acids is required. Hence, an analytical tool that can detect 1-100 nM nucleic acids is required.

In order to achieve SNP genotyping of hairpin-folded analytes, we suggested the following approach. The analyte was incubated with a universal MB (**UMB**) probe and a set of two target-specific adaptor oligonucleotides, **strand a** and **strand b** (Figure 2b). **Strands a** and **b** contain fragments complementary to the analyte (analyte-binding arms) and fragments complementary to the MB probe (MB-binding arms). In the absence of specific nucleic acid analyte, the MB-binding arms of **strands a** and **b** only weakly interact with MB probe because each arm hybridizes to a relatively short (9 nucleotides) fragment of **UMB**. In the presence of analyte, the long analyte-binding arm of **strand b** hybridizes to the fragment adjacent to the SNP site and thus unwinds the analyte secondary structure. The resultant complex contains the SNP site in a single-stranded region accessible for analysis. Then, hybridization of **strand a** and MB probe results in the formation of a quadripartite complex, which contains UMB probe in highly fluorescent opened conformation. Importantly, **strand a** contains a short analyte-binding arm which forms a stable hybrid only with the fully complementary analyte. According to

Kolpashchikov, the splitting of the DNA probe into two parts increases the selectivity of the probe because a mismatch base pair in short sequence results in a significant energetic penalty (26). Such two-component probes can recognize long DNA sequences (>15 nucleotides) and be highly sensitive to a single nucleotide substitution in the analyte. For the probe shown in Figure 2b, the presence of a single base mismatch in the case of mutation will destabilize the quadripartite complex resulting in its decomposition into MB probe, **strand a** and **strand b**-analyte duplex. In this study we show that this approach allows SNP analysis in model DNA analytes containing secondary structures in the form of stem-loop (hairpin).

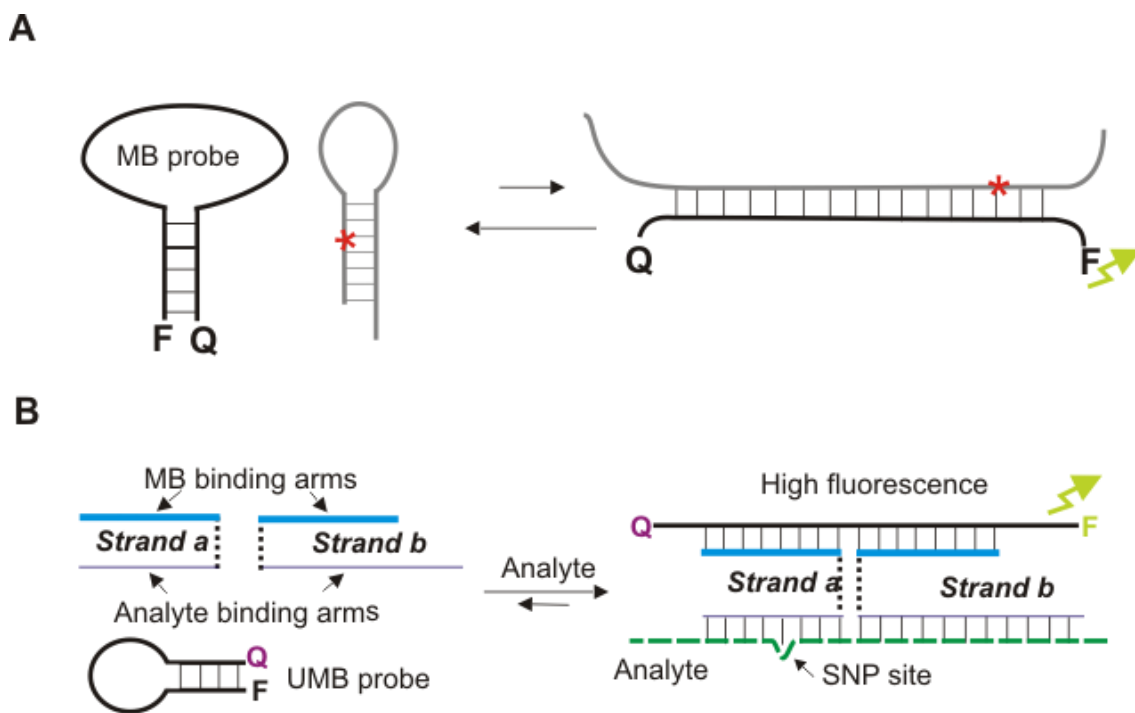


Figure 2. Molecular beacon probe for the analysis of specific nucleic acids. A: General scheme of hybridization of MB probe to a hairpin-folded analyte. SNP site is indicated by an asterisk. The equilibrium is shifted toward the dissociated state. B: The tripartite MB-based probe for SNP analysis in secondary structure-forming analytes. Probe designs. **Strands a** and **b** contain analyte-binding arms and MB-binding arms. SNP site in the analyte is indicated.

The goal of this project was to demonstrate the ability of MB-based tripartite probe to fluorescently report the presence of stem-loop folded DNA sequences in solution. Moreover, the accurate analysis of single nucleotide substitutions in stem-loop folded analytes would be achieved. We would pursue the following two specific aims.

Aim 1: Real time fluorescent detection of stem-loop folded nucleic acid analytes.

MB-based tripartite probe would be designed to fluorescently report DNA analytes that contain 5-9, 11, and 13 nucleotide stems. The efficiency of the detection of analytes containing different stem-loop lengths would be assessed by comparison of signal-to-background ratios generated by the probe for each analyte. High signal to background ratio regardless the stem lengths of DNA analytes would be achieved. Achieving this aim would demonstrate the efficiency of the approach in analysis of DNA containing as long as 13 nucleotide stems, which are very stable at room temperature.

Aim 2: Single nucleotide substitution (SNS) analysis in secondary structure-folded nucleic acid analytes at room temperature. Tripartite probe would be used for recognition of single base substitutions at different positions in a DNA analyte: the first SNS was located in the loop, the second SNS was located at the edge of the loop, and the third SNS was located in the middle of the stem (the stem is 13-base pair long). High signal to background ratio was expected in the presence of the true target while the mutants would have significantly lower signal to background ratio. Achieving this aims would demonstrate that the suggested approach was applicable for the analysis of mutations disregarding of their location in a DNA hairpin.

Experimental Designs

Hairpin motif does occur in vivo. For instance in the regions of genomic DNA that regulate gene expression and replication. For example, such region is located in the boundary of exon 10 coding for Tau protein, which has role in the stabilization for the microtubule in the nervous tissue filament. This sequence also contains the hairpin motif for the pre-mRNA, and the SNPs in this sequence destabilize the stem structure, and eventually lead to the alternate splicing of exon 10 in the tau mRNA. As a result, the protein product from this mRNA is hyperphosphorylated and destabilizes the microtubule in the nervous tissue filament. The malformation of the filament in the neuron is the main reason for many neurodegenerative tauopathies such as Alzheimer's disease (27).

The model DNA we proposed for our experiment was an oligonucleotide based on the gene encoding for Tau protein. However, the sequence of the DNA analyte was modified to vary the lengths of the stem from 5 to 13 bases. It is important to systematically study the effect of stem length on the ability of both conventional MB probe and tripartite probe to analyze such stem-loop folded nucleic acids. In addition to the DNA true targets, we analyzed the 3 SNSs located at different positions in the DNA hairpin, which were the T13-8A, T13-12A and T13-16C substitutions (Fig. 3).

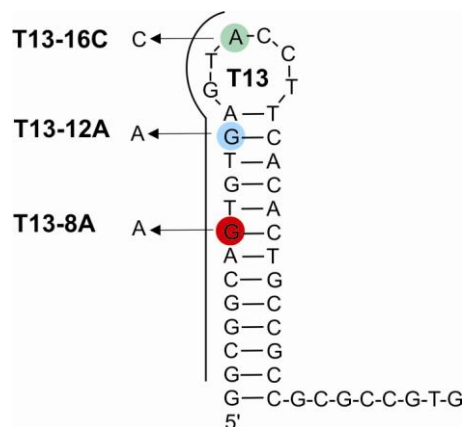


Figure 3. Structure of T13 analyte and its analogs used in this study. Three T13 mutants T13-8A, T13-12A and T13-16C contained base substitutions at 8, 12 and 16 positions respectively. The line indicates nucleotides complementary to MB1 probe. (30)

In order to analyze the three SNSs, we designed two pairs of adaptor strands. **Strand a-1** and **b-1** were used to analyze the T13-12A and T13-16C mutants, while **strand a-2** and **strand b-2** were used for the T13-8A mutant. The schematic hybridization of analyte T13 with the two pairs of adaptor strands and UMB are shown Figure 4. The fluorescent assay condition for this analysis is shown in Table 1C (30).

Additionally, we systematically analyzed 7 DNA targets folded in hairpin structure (T5-T9, T11, and T13) with different stem lengths using the multi-component biosensor. The two adaptor strands were sequence-dependent and were designed to hybridize to targeted parts of the DNA analytes. Only one **strand a-1** was used to analyze all 7 targets. Strand b's were varied to maximize the fluorescent signals: **strand b-T5** → T5, T6 and T7, **strand b-T8** → T8 and T9, **strand b-T11** → T11, and **strand b-1** → T13. Fluorescent assays were conducted for both the tripartite probes and the MB probes to compare the efficiency of the two types of probes in detecting secondary-structured DNA analytes (Table 1). The signal to background ratio (S/B) is the ratio of the fluorescent signals of the reaction mixture in the presence of the DNA analyte to that of the reaction mixture without DNA analyte.

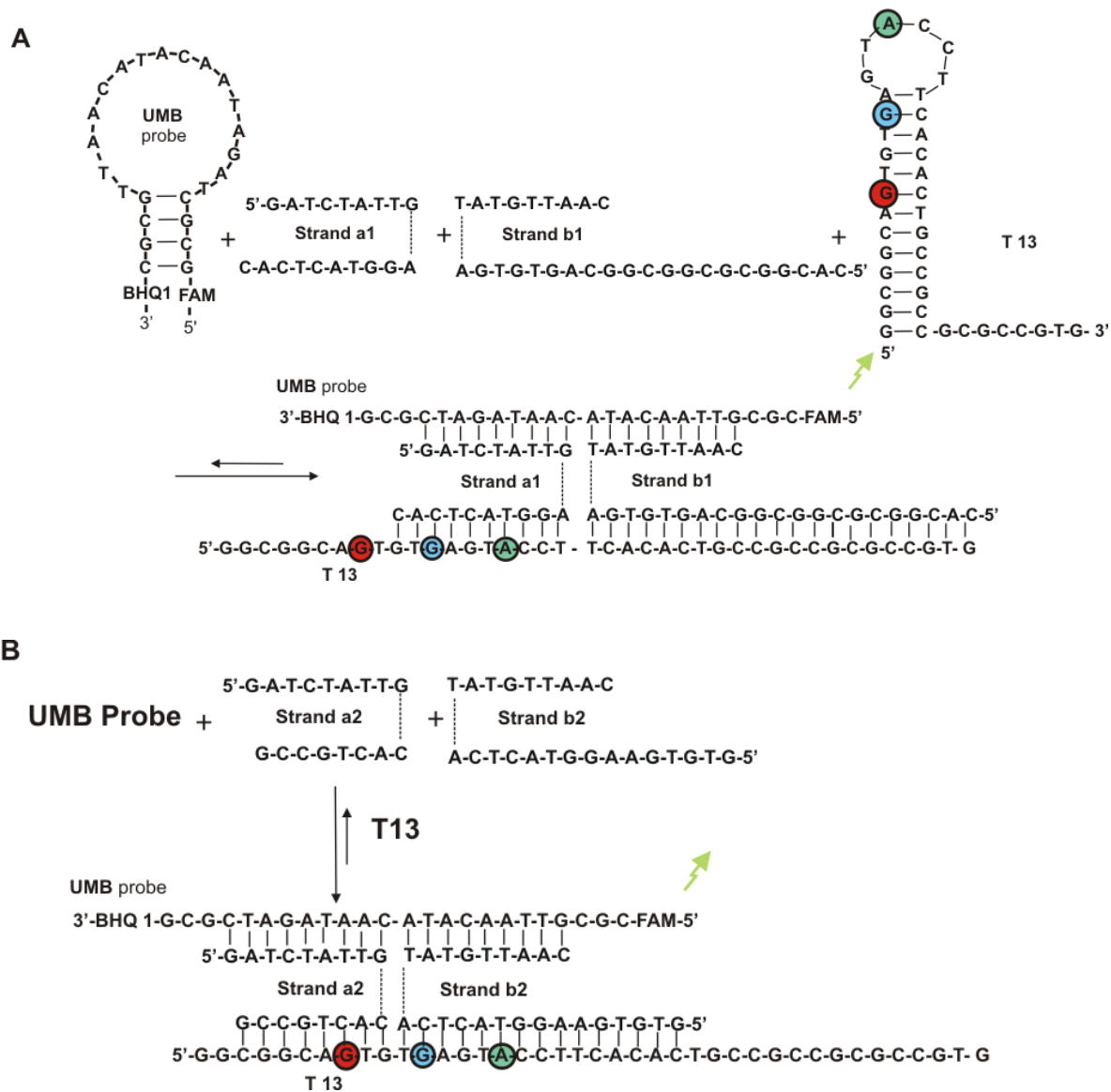


Figure 4. The hybridization reaction between the tri-component probe the target T13: The dotted lines in the two adaptor strands indicate the TEG linker between the DNA binding arm and the UMB binding arm. A: The DNA binding arm of **strand a-1** can analyze T13-12A and T13-16C substitutions. B: The DNA binding arm of **strand a-2** can analyze the T13-8A substitution. The 4-way junction is shifted 9 bases toward the 5' end of the T13.

Table 3. Fluorescent assay conditions.

A. Tricomponent probe approach:

Buffer condition: 20 mM MgCl ₂ , 50 mM Tris-HCl, pH 7.4, 15 minutes incubation time at 22°C.	MB probe	Strand a	Strand b	Targets
	UMB (50 nM)	Strand a-1 (100 nM)	Strand b-T5 (100 nM)	Analytes T5/T6/ T7 (100 nM)
	UMB (50 nM)	Strand a-1 (100 nM)	Strand b-T8 (100 nM)	Analytes T8/T9 (100 nM)
	UMB (50 nM)	Strand a-1 (100 nM)	Strand b-T11 (100 nM)	Analytes T11 (100 nM)
	UMB (50 nM)	Strand a-1 (100 nM)	Strand b-1 (100 nM)	Analytes T13 (100 nM)

B. MB probe approach:

Buffer condition: 2mM MgCl ₂ , 50 mM Tris-HCl, pH 7.4, 15 minutes incubation time at 22°C.	MB probe	Targets
	MB 1 (100 nM)	Analytes T5/T6/T7/T8/T9/T11/T13 (100 nM)

C. Fluorescent assay condition in the analysis of the three mutants:

Buffer condition: 20 mM MgCl ₂ , 50 mM Tris-HCl, pH 7.4, 15 minute incubation time at 22°C.	MB probe	Strand a	Strand b	Targets
	UMB (50 nM)	Strand a-1 (100 nM)	Strand b-1 (100 nM)	Analytes T13/T13-12A/ T13-16C (100 nM)
	UMB (50 nM)	Strand a-2 (100 nM)	Strand b-2 (1000 nM)	Analytes T13/T13-8C (1000 nM)

The limits of detection (LOD) of tri-components probes were determined in this study. The tri-component probe was used to analyze the 7 DNA analytes at low concentrations (1, 2, 3, 4, 5, 10 nM). Fluorescent signals from three measurements were averaged and plotted. The limit of detection (LOD) is equal to the average fluorescent signal of the blank plus three standard deviation of the blank. The concentration corresponding to this fluorescent value, extrapolated from the graph is the LOD.

All oligonucleotides were custom synthesized by Integrated DNA Technology, Inc (Coralville, IA). The fluorescent signals were measured on a Perkin-Elmer (San Jose, CA) LS-55 Luminescence Spectrometer with a Hamamatsu xenon lamp (excitation at 485 nm; emission 517 nm) after 15 min of incubation at room temperature (22 °C). Data for each experiment were triplicate, and standard deviations would be shown as error bars on the graph

Table 4. Oligonucleotides used in the study.

Name	Sequence	Purification method
UMB	5'-FAM- <u>CGCG</u> TTAA CATA CAAT AGAT <u>CGCG</u> -BQ1-3'	HPLC
MB1	5'-FAM- <u>CCGC</u> TACT CACA CTGC CGC <u>GCGG</u> -BQ1-3'	HPLC
C15	5'- GCGGCAGTGTGAGTA-3'	SD
T5	5'-GGCG GCAG TGTG AGTA CCTT CACA <u>ACTT</u> <u>TTTT GTTC</u> C-3'	SD
T6	5'-GGCG GCAG TGTG AGTA CCTT CACA <u>CCTT</u> <u>TTTT GTTCC</u> -3'	SD
T7	5'-GGCG GCAG TGTG AGTA CCTT CACA <u>CTTT</u> <u>TTTT GTTCC</u> -3'	SD
T8	5'-GGCG GCAG TGTG AGTA CCTT CACA <u>CTGT</u> <u>TTTT GTTCC</u> -3'	SD
T9	5'-GGCG GCAG TGTG AGTA CCTT CACA CTGC <u>TTTT GTTC</u> C-3'	SD
T11	5'-GGCG GCAG TGTG AGTA CCTT CACA CTGC <u>CGTT GTTCC</u> -3'	SD
T13	5'-GGCG GCAG TGTG AGTA CCTT CACA CTGC CGCC GCGC CGTG-3'	SD
T13-16C	5'-GGCG GCAG TGTG AGT <u>C</u> CCTT CACA CTGC CGCC GCGC CGTG-3'	SD
T13-8A	5'-GGCG GCA <u>A</u> TGTG AGTA CCTT CACA CTGC CGCC GCGC <u>CGTG</u> -3'	SD
T13-12A	5'-GGCG GCAG TG <u>T</u> A AGTA CCTT CACA CTGC CGCC GCGC CGTG-3'	SD
Strand a-1	5'-GAT CTA TTG /TEG/ AGG TAC TCAC-3'	SD
Strand b-1	5'-CACG GCGC GGCG GCAG TGTG A/TEG/TATGTTAAC-3'	SD
Strand a-2	5'-GAT CTA TTG /TEG/ ACT GCC GCC-3'	SD
Strand b-2	5'-CACG GCGC GTCA GTAA TGTG AAGG TACT CAC /TEG/ TAT GTT AAC-3'	SD
Strand b T5	5'-GGAA CAAA AAAG TTGT GA/TEG/TAT GTT AAC-3'	SD

Strand b T8	5'-GGAA CAAA AACA GTGT GA/TEG/TAT GTT AAC-3'	SD
Strand b T11	5'-GGA ACA ACG GCA GTG TGA/TEG/TATG TTAAC-3'	SD

“/TEG/” indicates triethylene glycol linker inserts. Stem-forming nucleotides of molecular beacon probes and wild-type analytes are shown in italic. MB-binding arm of the strands are underlined. Mutated nucleotide positions in the analytes are depicted in bold and underlined. “SD” is standard desalting. “HPLC” is high pressure liquid chromatography. The SNSs are in bold and underlined.

Results

The tri-component probe can analyze DNA hairpins regardless of the stem length.

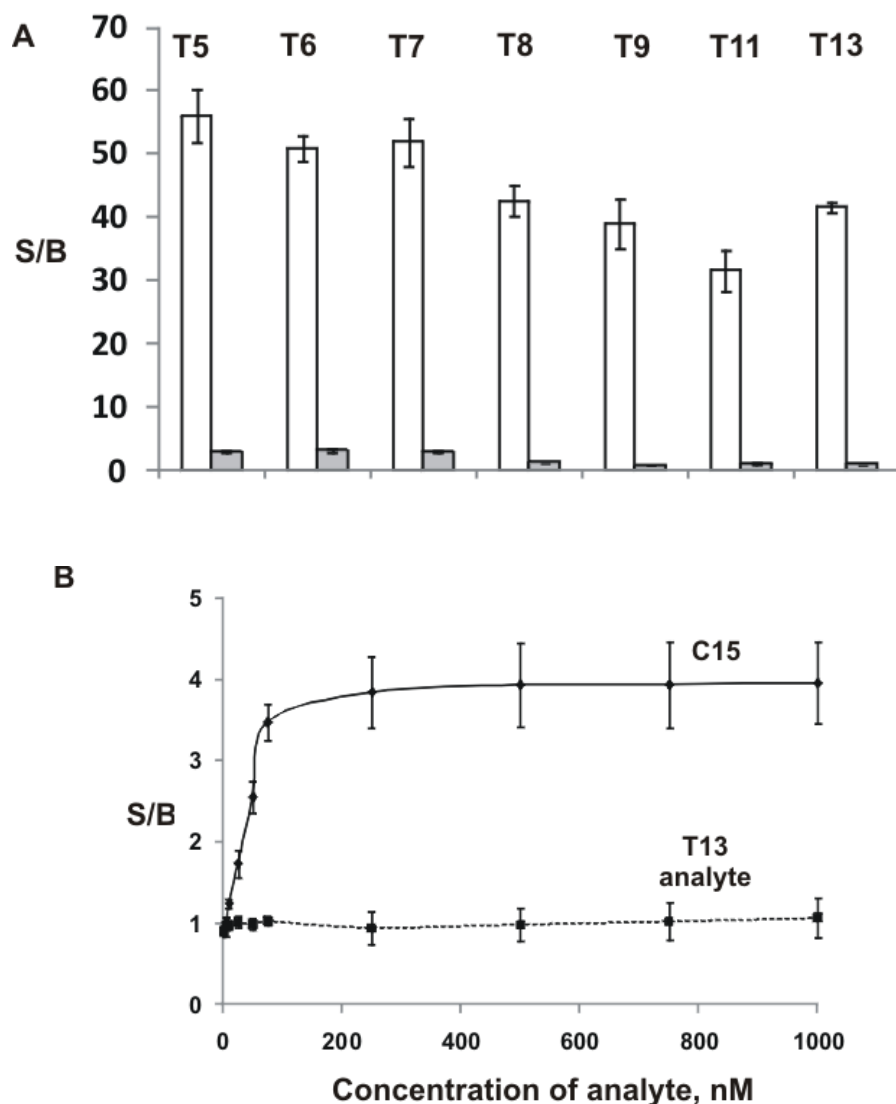


Figure 5. Comparison between the MB probe approach and the tri-component probe approach. A. Comparison on signal to background ratio between the MB probe approach and the tripartite probe approach in detecting the presence of different targets. The gray bars show the data for the MB probe approach, and the white bars show the data for the tripartite probe approach. Fluorescent assay condition is indicated in table 1A and 1B. B. The signal to background ratio for the MB probe with different concentration of its fully complementary strand C15 (5'-GCGGCAGTGTGAGTA-3') and analyte T13. The DNA target is folded in stem loop and is fully complementary to the loop portion of the MB probe. Fluorescent assays condition: MB probe 100 nM; DNA targets from 1-100 nM; buffer MgCl₂ 2 mM, Tris-HCl: 50 mM; and incubation time 15 minutes.

The seven DNA analytes were analyzed with the conventional MB probe and the tripartite probes. The fluorescent signals were collected and the signals to background ratios were calculated. The conventional MB probe approach failed to analyze DNA hairpins with different stem lengths (Figure 5a). For Figure 5a, the maximum S/B ratio was only about 3 for the targets T5, T6 and T7, which have relatively short stem lengths. For other targets, the fluorescent data were only about background level and thus MB probe in its traditional format could not report the presence of these targets. Annealing experiment or longer incubation time did not improve the S/B ratio. The MB1 probe was unable to form the duplex with the DNA hairpins, since the hybridization reaction between the DNA analytes and the MB probe was not a favored outcome (30).

The S/B ratio of MB probe was calculated at different concentrations of its complementary strands C15 (solid line) and the hairpin-shaped target T13 (dotted line) (Figure 5b). MB probe in the presence of its complementary strand significantly increased the fluorescent signal in the concentration of C15 as low as 5 nM, but MB probe failed to form the duplex with T13 even at very high concentration of T13 (1000 nM for T13). Therefore, the MB1 probe was unable to directly analyze the DNA hairpins, but only able to hybridize with linear target C15 (30).

The tripartite probe approach, on the other hand, showed that S/B ratio was significantly higher than those of MB probe approach. Fluorescent assays were conducted at room temperature and the samples were not heated. In addition, the background signals were as low as 8-12 a.u. (fluorescent arbitrary units) after 15 minutes of incubation (data for background noise are not shown). The addition of the true targets triggered a significant increasing of fluorescent. Regardless of the stem lengths, the S/B ratio were high and consistent among the targets.

Although the S/B ratio for the targets varied from 30 to 50, this approach was capable of detecting the presence of DNA hairpin in real time. The background ratio of targets T11 and T13 although lower but were still high enough to efficiently report the presence of the targets (30).

Limits of detection of the tri-component probe in analysis of all 7 analytes

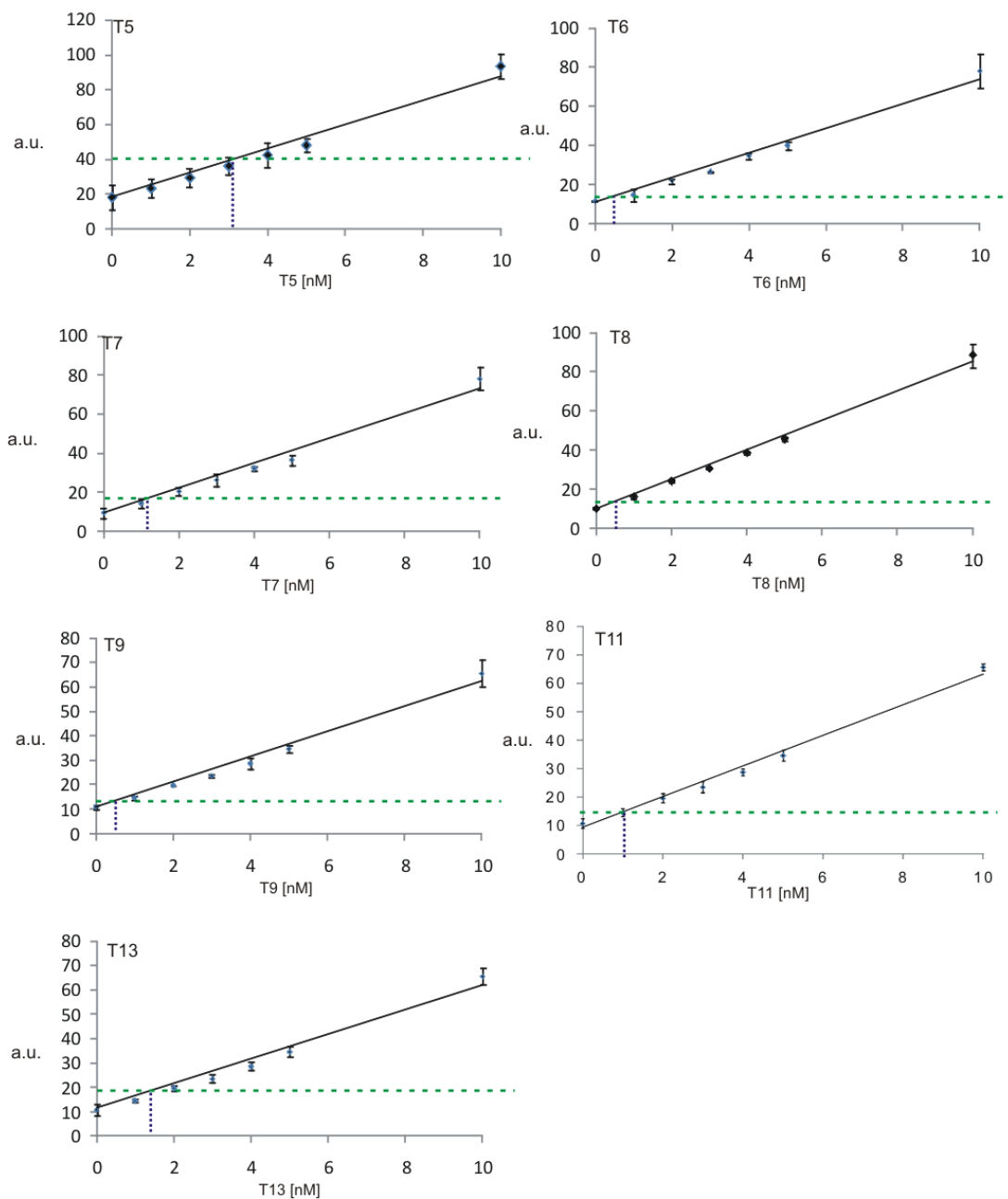


Figure 6. Detection limits of tri-component probe. Tripartite probe was used to analyze low concentrations of the targets. Fluorescent signals from three measurements were averaged and plotted; Error bars were included in the graph. The limit of detection (LOD) equals to the average fluorescent signal of the blank plus three standard deviation of the blank. The concentration corresponding to this fluorescent value, extrapolated from the graph is the LOD. Fluorescent assays conditions: Low concentrations of DNA analytes (1, 2, 3, 4, 5, and 10 nM), and other components of the probe were the same as table 1A.

In order to find limits of detection (LOD) of the tripartite probe, low concentration of the 7 DNA analytes (1, 2, 3, 4, 5, 10 nM) were incubated in the presence of all three components (**strand a**, **strand b**, and UMB probe - the concentration of each component was the same as those mentioned in Table 1A). From three independent measurements, the fluorescent signals were averaged and plotted against the analyte concentration. The statistical trend lines were graphed based on these sets of data point. Standard deviations were included as error bars in the graph. The average fluorescent value of the blank was added to three standard deviations of the blank, and the concentration corresponding to this value, extrapolated from the graph is the LOD (28).

The average fluorescence of the blank for analyte T6 was 11.28 plus three standard deviation 0.27. Therefore, the fluorescent value of the LOD would equal: $11.28 + 0.27 \times 3 = 12.09$. The concentration corresponding to this fluorescence is 0.5 nM. In general, the LODs for all 7 analytes are ranged from about 0.5-3 nM, which are relatively low and close to the detection limit of the conventional MB probe.

Selectivity of the tricomponent probe in detecting 3 SNSs in DNA hairpin

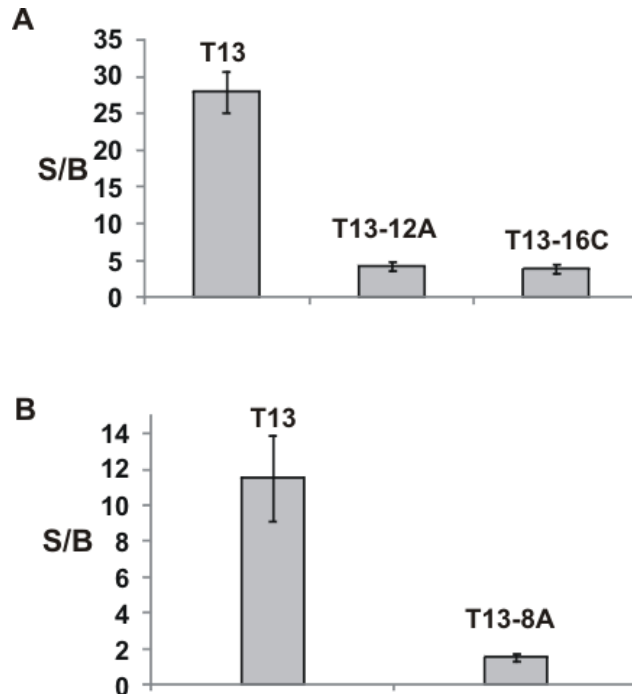


Figure 7 Selectivity of the tricomponent probe approach in detecting 3 single nucleotide substitutions (SNSs) as indicated in Figure 3. A. S/B ratio in the presence of the true target T13, and 2 mismatched T13-12A and T13-16C respectively. B. S/B ratio in the presence of the true target T13 and the mismatched T13-8A. Fluorescent assays condition is specified in table 1C.

Two tripartite probes were used to analyze the three SNSs located at different sites in T13 analyte (Figure 3). The two probes used the same UMB probe but two pairs of adaptor strands. **Strand b** has long analyte binding arm in order to unwind the long stem of T13 analyte while **strand a** has shorter analyte binding arm for which improves hybridization selectivity. In other words, the four-way junction is only formed in optimized condition if **strand a** is fully complementary to the true target T13 and this resulted in high fluorescent signal, while the mismatched has low fluorescent signal. For the 12-A and the 16-C substitution, **strand b-1** and **strand a-1** were used. The S/B ratio for the true target T13 was as high as 28 while those for the two analytes 12-A and 16-C were only about 5 (Figure 7a) (30).

The 8-C SNS was analyzed with **strand b-2** and **strand a-2**. For this target, the SNS is located in the middle of the stem, which caused much complication due to self-complementarity between the two adaptor strands. In order to solve this issue, **strand b-2** has a stretch of mismatched nucleotides to disrupt the self-complementarities between the two adaptor strands while still allowing the formation of the 4-way junction. In addition, high concentration for both the analytes (1000 nM) and **strand b-2** (1000 nM) were used to enable efficient hybridization for both of them and thus the formation of the four-way junction. The S/B ratio was about 12 for the true analyte T13 and was about 2 for the mismatched 8-A, which was sufficient to distinguish between the true target T13 and the mismatched T13 8-A (30).

Additionally, we also calculated the discrimination factor (DF) ————— to show the ability of the tri-component probe in detecting change in the DNA sequence ($F_{\text{mismatched}}$ is the fluorescent signal for the substituted target, F_{matched} is the fluorescent signal for the true targets, and B is the background). The DF is the average from three independent measurements. DF should be close to 1 for the mutants because the mismatched generate no significant increase of fluorescent signal relative to the background. DF's were 0.879, 0.893 and 0.950 for the analytes T13 16-C, T13 12-A and T13 8-A respectively. DF's for the two analytes 16-C and 12-A were close to 0.9. Although these two mutants did generate fluorescent signal above the background (Figure 7A), the probe was able to report the change in the DNA sequence. The DF of the T13 8-A mutant was as high as 0.95; therefore, we could distinguish between the two analytes T13 and T13 8-A.

Discussions

The MB probe has been a popular tool for nucleic analysis owing to its high specificity and simple design. However, the design of MB probe is not as simple as the stem arms can't be arbitrarily attached to any loop sequence complementary to the targets. One drawback is the stem invasion, in which the stem arm is complementary to a portion of the loop and thus results in the partial opening conformation of the MB probe. The partial quenching of the fluorophore leads to high background noise in the assay. Therefore, optimization for the probe is an important step in the design of each new probe for each new analyte and increases the cost for the probe in multiplex assays (29).

Another disadvantage of MB probe is its inefficiency in the analysis of DNA folded in secondary structures. The stem of T13 analyte has the melting temperature as high as 88.3°C, which reflects the high stability of the secondary structure at room temperature. This stable secondary structure prevents application of conventional MB probe to analyze this target. The duplex between the MB1 probe and the hairpin DNA analytes were not formed or formed at low efficiency to detect the fluorescence. The hybridization reaction was unfavorable since the opening of the long stem of the DNA analytes and the probe requires more energy input. The fluorescent ratios of the MB1 probe in this study were much lower than those of the multi-component probe and thus proved this drawback of the MB probe. This limits the use of the powerful MB probe to analyze secondary structured DNA targets.

However, it is practically important to develop a tool that allows analysis of such DNA analytes since many *in vivo* DNA sequences involve secondary structures. The MB-based tripartite probe in this study can circumvent the two biggest drawbacks mentioned above. This method utilizes the elegance of MB probe and applies it in the analysis of very stable DNA

hairpins. The analyte T13 has the melting temperature as high as 88.3°C, which we could not use the conventional MB probe to analyze this target.

Our approach has the following advantages. First, the design of the UMB probe is independent to the DNA targets, so only one pre-optimized molecular beacon is used to analyze different analytes. The stem invasion issue is, therefore, avoided, by using the tripartite probe. The introduction of the two adaptor **strands a** and **b** has solved this major problem in the optimization for MB probe design. In this study, only one UMB was used to analyze all ten DNA targets while this is impossible as using the MB probe approach. Second, the design for each adaptor strand is straightforward as each strand has one portion complementary to the UMB probe and one portion complementary to the analyte. **Strand b** has long analyte binding arm to unwind the target while **strand a** has a short analyte binding arm to detect the substitutions in the target sequence. Therefore, this type of probe allows detecting in the presence of DNA hairpin as well as SNSs in real-time condition. Moreover, the detection limit of this probe for all the 7 hairpins is only 0.5-3.0 nM. The LOD of the conventional MB probes is also close to these values. Therefore, the multi-component probe is highly sensitive in detect the presence of these DNA targets.

Furthermore, the amount of each adaptor strand can be varied to obtain optimum fluorescence in each assay. The fluorescent intensities for the tri-component probes are significantly higher than the background in the analysis of all 7 DNA hairpins. The probe has high selectivity for SNSs in the sequence as we can distinguish the true target T13 from its three analogs containing mismatched simply based on the S/B ratio. The DFs of the probes were relatively close to 1; especially that of T13 8-A mutant was 0.95, which also display the probes ability to report small changes in a DNA sequence. Although we had to use high concentration of

strand b and the targets in the assays for T13 and T13-8A, the S/B ratio is high enough to select for the true target from the mutants. One explanation for this complication is because the target has relatively long stem (13 bp) causing high background noise due to undesired interaction of the adaptor strands with the UMB and low kinetic in the unwinding of this target. The introduction of several mismatched into **strand b-2** had temporarily solved this issue, but further optimization step remains for future studies (30).

Conclusion

The tripartite biosensor is an elegant tool for SNSs detection for DNA secondary structure. The two adaptor strands allows flexibility in the assay since the concentration of **strand a** and **b** can be optimized for the best S/B results. The experimental setup is simple and the instant result can be obtained in real-time fluorescent assays. Another advantage of this approach is the possibility to use one UMB for the analysis of many targets and SNSs. The only change in the probe is the DNA binding arms of the two adaptor strands, which are inexpensive, compared to the designing new a MB probe for each DNA target. Only one type of UMB is necessary to analyze 3 SNSs at different positions in the DNA hairpin while the same result is impossible to achieve using solely the MB probe. This contributes to the cost efficiency of the suggested approach in case of multiplex analysis. The multi-component probe is highly sensitive to detect the presence of DNA targets folded in secondary structures at relatively low concentration. Hence, it was be demonstrated that the tripartite probe is excellent for discrimination single nucleotide substitution in DNA constrained by secondary structures. The aforementioned characteristics of tricomponent probe promise the possibility to monitor RNA and detect DNA in vivo with real time fluorescent assays. Indeed, the tripartite biosensor is highly accurate for real time SNSs genotyping, which may be applied in clinical SNPs multiplex assay at low price.

Reference

1. R. Sachidanandam, D. Weissman, S. C. Schmidt, J. M. Kakol, L. D. Stein, G. Marth, S. Sherry, J. C. Mullikin, B. J. Mortimore, D. L. Willey, S. E. Hunt, C. G. Cole, P. C. Coggill, C. M. Rice, Z. Ning, J. Rogers, D. R. Bentley, P. Y. Kwok, E. R. Mardis, R. T. Yeh, B. Schultz, L. Cook, R. Davenport, M. Dante, L. Fulton, L. Hillier, R. H. Waterson, J. D. McPherson, B. Gilman, S. Schaffner, W. J. Van Etten, D. Reich, J. Higgins, M. J. Daly, B. Blumenstiel, J. Baldwin, N. Stange-Thomann, M. C. Zody, L. Linton, E. S. Lander, D. Altshuler. "A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms" *Nature* 2001, 409, 928-933
2. The International HapMap Consortium. "A Haplotype Map of the Human Genome." *Nature* 437.7063 (2005): 1299-320.
3. Blaisdell, Carol J., Timothy D. Howard, Augustus Stern, Penelope Bamford, Eugene R. Bleecker, and O Colin Stine. "CLC-2 Single Nucleotide Polymorphisms (SNPs) as Potential Modifiers of Cystic Fibrosis Disease Severity." *BMC Medical Genetics* 5:26 (2004).
4. Vainzof, Mariz, Maria Rita Passosbueno, Reinaldo I. Takata, Rita De Cassia M. Pavanello, and Mayana Zatz. "Intrafamilial Variability in Dystrophin Abundance Correlated with Difference in the Severity of the Phenotype." *Journal of the Neurological Sciences* 119.1 (1993): 38-42.
5. Bell, G. I., Horita, S. & Karam, J. H. A polymorphic locus near the human insulin gene is associated with insulin-dependent diabetes mellitus. *Diabetes* 33, 176–183 (1984).
6. Burwinkel, Barbara, Kalai S. Shanmugam, Kari Hemminki, Alfons Meindl, Rita K. Schmutzler, Christian Sutter, Barbara Wappenschmidt, Marion Kiechle, Claus R.

- Bartram, and Bernd Frank. "Transcription Factor 7-like 2 (TCF7L2) Variant Is Associated with Familial Breast Cancer Risk: a Case-control Study." *BMC Cancer* 6 (2006): 268.
7. Budowle, Bruce, and Angela Van Daal. "Forensically Relevant SNP Classes." *BioTechniques* 44 Supplement.4 (2008): 603-10.
 8. Guzey, Cuneyt, and Olav Spigset. "Genotyping as a Tool to Predict Adverse Drug Reactions." *Current Topics in Medicinal Chemistry* 4.13 (2004): 1409-419.
 9. Beaudet, Arthur L., and John W. Belmont. "Array-Based DNA Diagnostics: Let the Revolution Begin." *Annual Review of Medicine* 59.1 (2008): 113-29.
 10. Jawaid, Ansar, Rachael Moore, Ruth March, and Bryan J. Barratt. "Gene Mapping Strategies for Complex Disease and Drug Response." *Drug Discovery Today: Technologies* 3.2 (2006): 131-36.
 11. Morrison, Larry E., Thomas C. Halder, and Lucy Stols M. Stols. "Solution-phase Detection of Polynucleotides Using Interacting Fluorescent Labels and Competitive Hybridization." *Analytical Biochemistry* 183.2 (1989): 231-44.
 12. Cardullo, Richard A., Sudhir Agrawal, Carlos Flores, Paul C. Zamecnik, and David E. Wolf. "Detection of Nucleic Acid Hybridization by Nonradiative Fluorescence Resonance Energy Transfer." *Proceedings of the National Academy of Sciences* 85.23 (1988): 8790-794.
 13. Santangelo, Philip, Nitin Nitin, and Gang Bao. "Nanostructured Probes for RNA Detection in Living Cells." *Annals of Biomedical Engineering* 34.1 (2006): 39-50.

14. K., Ebata, Masuko M., Ohtani H., and Kashiwasake-Jibu M. "Nucleic Acid Hybridization Accompanied with Excimer Formation from Two Pyrene-labeled Probes." *Photobiochemistry and Photobiology* 62 (1995): 836-39.
15. Tyagi, Sanjay, and Fred Russell Kramer. "Molecular Beacons: Probes That Fluoresce upon Hybridization." *Nature Biotechnology* 14.3 (1996): 303-08.
16. Marras, Salvatore A.E., Sanjay Tyagi, and Fred Russell Kramer. "Real-time Assays with Molecular Beacons and Other Fluorescent Nucleic Acid Hybridization Probes." *Clinica Chimica Acta* 363.1-2 (2006): 48-60.
17. Wang, Kemin, Zhiwen Tang, Chaoyong James Yang, Youngmi Kim, Xiaohong Fang, Wei Li, Yanrong Wu, Colin D. Medley, Zehui Cao, Jun Li, Patrick Colon, Hui Lin, and Weihong Tan. "Molecular Engineering of DNA: Molecular Beacons." *Angewandte Chemie International Edition* 48.5 (2009): 856-70.
18. Whitman, Douglas F., and Sherry A. Dunbar. "Real-Time Polymerase Chain Reaction Detection Methods." *Recent Patents on DNA & Gene Sequences* 2.1 (2008): 20-26.
19. Bao, Gang, Won Jong Rhee, and Andrew Tsourkas. "Fluorescent Probes for Live-Cell RNA Detection." *Annual Review of Biomedical Engineering* 11.1 (2009): 25-47.
20. Tyagi, Sanjay. "Imaging Intracellular RNA Distribution and Dynamics in Living Cells." *Nature Methods* 6.5 (2009): 331-38.
21. Bonnet, Grégoire, Sanjay Tyagi, Albert Libchaber, and Fred Russell Kramer. "Thermodynamic Basis of the Enhanced Specificity of Structured DNA Probes." *Proceedings of the National Academy of Sciences* 96 (1999): 6171-176.
22. Armitage, Bruce. "The Impact of Nucleic Acid Secondary Structure on PNA Hybridization." *Drug Discovery Today* 8.5 (2003): 222-28

23. Varani, Gabriele. "Exceptionally Stable Nucleic Acid Hairpins." *Annual Review Biophysics and Biomolecular Structure* 24 (1995): 379-404.
24. Egholm, Michael, Ole Buchardt, Leif Christensen, Carsten Behrens, Susan M. Freier, David A. Driver, Rolf H. Berg, Seog K. Kim, Bengt Norden, and Peter E. Nielsen. "PNA Hybridizes to Complementary Oligonucleotides Obeying the Watson-Crick Hydrogen-bonding Rules." *Nature* 365.6446 (1993): 566-68.
25. Kushon, Stuart A., Jason P. Jordan, Jennifer L. Seifert, Henrik Nielsen, Peter E. Nielsen, and Bruce A. Armitage. "Effect of Secondary Structure on the Thermodynamics and Kinetics of PNA Hybridization to DNA Hairpins." *Journal of the American Chemical Society* 123.44 (2001): 10805-0813.
26. Kolpashchikov, Dmitry M. "A Binary DNA Probe for Highly Specific Nucleic Acid Recognition." *Journal of the American Chemical Society* 128 (2006): 10625-0628.
27. Goedert, Michel, and Ross Jakes. "Mutations Causing Neurodegenerative Tauopathies." *Biochimica Et Biophysica Acta (BBA) - Molecular Basis of Disease* 1739.2-3 (2005): 240-50.
28. Gerasimova, Yulia V., Evan Cornett, and Dmitry M. Kolpashchikov. "RNA-cleaving Deoxyribozyme Sensor for Nucleic Acid Analysis: the Limit of Detection." *Chembiochem* 11.6 (2010 April 12): 811-17.
29. Browne, Kenneth A. "Sequence-Specific, Self-Reporting Hairpin Inversion Probes." *Journal of the American Chemical Society* 127.6 (2005): 1989-994.
30. Nguyen, Camha, Jeffrey Grimes, Yulia Gerasimova, and Dmitry M. Kolpashchikov. "Molecular Beacon-based Tricomponent Probe for Single Base Substitution Analysis in Folded Nucleic Acids." *Chembiochem*. In press.