2022

G-Quadruplex Reporters: Structural Studies and Application for Visual and Fluorescent Detection of Point Mutations in Nucleic Acid Sequences

Ryan Connelly

University of Central Florida

Showcase of Text, Archives, Research & Scholarship

G-QUADRUPLEX REPORTERS:
STRUCTURAL STUDIES AND APPLICATION
FOR VISUAL AND FLUORESCENT DETECTION
OF POINT MUTATIONS IN NUCLEIC ACID SEQUENCES

by

RYAN PATRICK CONNELLY
M.S. University of Central Florida, 2019
B.S. Illinois State University, 2013

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Chemistry
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Spring Term
2022

Major Professor: Yulia Gerasimova

ABSTRACT

DNA-based diagnostics traditionally utilize hybridization probes: strands of DNA complimentary to a target sequence that, upon binding, generate a signal to indicate the presence of the target. The classic hybridization probes are the molecular beacon (MB) and TaqMan probes, both single DNA strands with a fluorophore and quencher at opposing ends. Despite their widespread use in applications such as qPCR due to their ability to multiplex with a variety of bound fluorophores, these probes have several shortcomings: temperature limitations for selective target recognition and differentiating single-nucleotide polymorphisms, intrinsic design limitations to interrogate some target sequences, and a high relative cost for synthesis and purification. A "split probe" design in combination with label-free reporters overcomes these shortcomings. Split probes break the target-recognition sequence into two shorter pieces, each equipped with a portion of a signal reporting unit. The shorter length both allows for the probes to efficiently function at ambient temperatures, opposed to the elevated temperatures generally required for monomer probes, and provides greater ability to discriminate single-nucleotide polymorphisms. These structures will only generate a signal if complete, which may only occur when both halves of the probe are properly bound to the target. By utilizing label-free reporters such as light-up aptamers and/or (deoxy)ribozymes, split probes offer cost-efficiency advantages over other fluorescent probes. In this dissertation, advances in the usage of split probes with G-quadruplex-based signal transducing units are detailed. An alphanumeric display comprised of tandem probes utilizing the peroxidase-like deoxyribozyme for colorimetric output demonstrates the instrument-free usage of these systems. Additionally, the promiscuous activity of the dapoxyl light-up aptamer with a variety of arylmethane, offers a label-free fluorescent split probe that is capable of discerning single-nucleotide polymorphisms without expensive chemical modifications.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF FIGURES

data are average values for three independent measurements with standard deviations as error bars. Inset: Image of the tubes containing the same samples. C) Analysis of "LOW input" and "HIGH input" states of G4-OR gate in non-denaturing 15% PAAG. The lane contents are shown under the gel image. "ssL" and "dsL" designate single-stranded and double-stranded DNA ladder, respectively. The sizes of the ladder fragments are indicated to the sides of the gel. The band corresponding to OR-A/OR-B complex is indicated by a green arrow.

Figure 24: G4-YES gate. A) In the absence of RIF[S] input, YES-A and YES-B strands acquire stem-loop conformations preventing their re-association into a G4 structure and, therefore, ensuring low background of the "LOW output" (left). Addition of the input triggers association of YES-A and YES-B by hybridizing to the input-recognition fragments of the strands (shown in dark grey) and formation of the catalytically active G4 peroxidase responsible for green color of "HIGH output" state (right). The structures are shown as predicted by Mfold[213] or Nupack[111] software. Nucleotides of the gate strands that are not complementary to the input sequence are shown in light grey. Nucleotides involved in the formation of G4 structure are in green. Position of a point mutation in the input is in cyan. Triethylene glycol linkers between the output-generating and input-recognizing modules of the gate strands are shown as black dotted curves. B) Structure of YES-A/RIF[R] complex predicted by Nupack.[111] C) Secondary structure of an alternative YES-A sequence without a G>A substitution, as predicted by Mfold.[213] The substitution position is underlined. D) Absorbance at 420 nm of the samples containing YES-A and YES-B in the absence ("No input") or presence of either RIF[S], RIF[R] or RIF[R]-G inputs. A threshold between low output and high output absorbance is represented by green dashed line. The data are average values for three independent measurements with standard deviations as error bars. Inset: Image of the tubes containing the same samples. E) Analysis of "LOW input" and "HIGH input" states of G4-YES gate in non-denaturing 12% PAAG. The lane contents

are shown under the gel image. "ssL" and "dsL" designate single-stranded and double-stranded DNA ladder, respectively. The sizes of the ladder fragments are indicated to the sides of the gel. The band corresponding to YES-A/YES-B/input complex is shown by a green arrow.

Figure 25: G4-NOR gate. A) In the absence of an input, NOR-A and NOR-B strands associate into a catalytically active structure of G4 peroxidase, thus producing "HIGH output" (left). 3'-Terminal overhang of NOR-A serves as a toehold for an input (e.g. RIF[S]) to displace NOR-A from the NOR-A/NOR-B complex and switch the gate into the "LOW output" state (right). The structure are shown as predicted by Mfold[213] or Nupack[111] software. Nucleotides of the gate strands that are not complementary to the input sequence are shown in light grey. Nucleotides involved in the formation of the G4 structure are in green. Position of a point mutation in the input is in cyan. B) Absorbance at 420 nm of the samples containing NOR-A and NOR-B in the absence ("No input") or presence of either RIF[S], RIF[R] or RIF[R]-G inputs. A threshold between low output and high output absorbance is represented by green dashed line. The data are average values for three independent measurements with standard deviations as error bars. Inset: Image of the tubes containing the same samples. C) Gel analysis of "LOW input" and "HIGH input" states of G4-NOR gate. The lane contents are shown under the gel image. "ssL" and "dsL" designate single-stranded and double-stranded DNA ladder, respectively. The sizes of the ladder fragments are indicated to the sides of the gel. The bands corresponding to NOR-A/NOR-B and NOR-A/input complexes are indicated with green and cyan arrows, respectively.

Figure 26: Logical negation of RIF[R] (NOT[R] function). A) Symbol for NOT[R] logic function obtained by mixing NOR and YES[S] gates. The NOR/YES[S] combination produces an output as if the outputs of NOR and YES[S] were recognized as inputs of an OR gate. B) Truth table for NOT[R] logic function as a sum of NOR and YES[S] logic functions. High output is highlighted in green. C) Absorbance at

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

4WJ   Four-way Junction

A   Adenine

ABTS   2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)

AO   Auramine O

BF   Basic Fuchsin

BG   Brilliant Green

C   Cytosine

CD   Circular Dichroism

CRISPR   Clustered Regularly Interspaced Short Palindromic Repeats

CV   Crystal Violet

DMSO   Dimethyl Sulfoxide

DNA   Deoxyribonucleic Acid

DSA   Dapoxyl Sulfonic Acid

DSEDA   Dapoxyl Sulfonyl Ethylenediamine

DSF   Dapoxyl Sulfonyl Fluoride

DST   Drug Susceptibility Test

FLAP   Fluorescent Light-Up Aptamer

FRET   Förster Resonance Energy Transfer

G   Guanine

G4   Guanine Quadruplex

HEPES N-(2-Hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)

I   Inosine

IAV   Influenza A Virus

INH   Isoniazid

LAMP   Loop-mediated Amplification

LNA   Locked Nucleic Acid

LOD   Limit of Detection

LOQ   Limit of Quantification

Mabs   *Mycobacterium abscessus*

MB   Molecular Beacon

MG   Malachite Green

MK   Michler's Ketone

MTB   *Mycobacterium Tuberculosis*

MTC   *Mycobacterium Tuberculosis Complex*

NASBA   Nucleic Acid Sequence Based Amplification

NMM   N-methylmesoporphyrin IX

PAAG   Polyacrylamide Gel

PCR   Polymerase Chain Reaction

PDz   Peroxidase-like Deoxyribozyme

PPIX   Protoporphyrin IX

PR   Pararosaniline

Q4   Guanine Quadruplex

RIF   Rifampin

RNA   Ribonucleic Acid

RPA   Recombinase Polymerase Amplification

S/B   Signal-to-background Ratio

SDA   Split Dapoxyl Aptamer

SELEX   Systematic Evolution of Ligands by Exponential Enrichment

SF   Selectivity Factor

SNP   Single Nucleotide Polymorphism

SNS   Single Nucleotide Substitution

sPDz   Split Peroxidase-like Deoxyribozyme

T   Thymine

TB	Tuberculosis
TFT	Thioflavin T
TNA	Threitol Nucleic Acid
U	Uracil
UV	Ultraviolet
VB	Victoria Blue B

CHAPTER ONE: INTRODUCTION

Origins of Antibiotic Resistance

Bacterial infections have historically been among the leading causes of human death. Even today, according to the World Health Organization, tuberculosis and the diarrheal diseases (rotavirus, *E. coli*, and others) remain within the top ten causes of death worldwide.[1] In 1928, Alexander Fleming discovered the antibacterial properties of penicillin; since, the rate of death due to bacterial infection has steadily been reduced as a vast number of antibiotics have been discovered to treat these infections.

As early as 1945, Fleming and others issued warnings about the possibility of an infection developing resistance to the drug.[2] A common mechanism of antibiotic resistance is a change in the structure of the protein that the antibiotic works on such that the antibiotic may no longer bind to it. This structural alteration is induced by a single-nucleotide change in the sequence of the protein, which may be caused by random mutation. Other mechanisms of antibiotic resistance exist, such as expulsion of the antibiotic by way of efflux pumps, disabling of the antibiotic via chemical modifications, or downstream modifications such as methylation of the target enzyme.[3-5] These changes, too, may often be attributed to a single-nucleotide mutation; some environmental stresses may cause an increase in expression of either efflux pumps or enzyme modifications, however, and bacteria may also transfer genes between themselves to share resistance. The growing demand for antibiotics for agriculture and medicine has led to an increase in the resistant phenotypes in bacteria. Between 2013 and 2019, the Centers for Disease Control and Prevention noted an increasing number of antibiotic-resistant infections, with the number of deaths as a result rising from 23,000 to 48,000.[6] As a consequence, there is a growing imperative for rapid testing of antibiotic resistance for bacterial infections.

The first widely-used tests for antibiotic resistance testing involved culturing the bacteria and either subjecting the culture to a paper disc coated in the antibiotic of choice or diluting the culture in an antibiotic-containing agar.[7] Culture-based assays directly verify the efficacy of a given antibiotic against the phenotype of the infective agent. These tests featured several key shortcomings, however, including a requirement for specific biosafety settings, laborious test administration requirements, long inoculation times for culturing such as the approximate two week doubling time for some *Mycobacterium*, and results that rely on the discretion of the user, limiting reproducibility and consistency.[8, 9] For clinical testing, these requirements hinder the widescale utilization of culture-based assays on-site; further, the increased test-to-result times may increase patient drop-out and lead to difficulties in limiting further transmission. As such, a push began for better clinic-side tests that overcome these constraints.

Hybridization Probes

The discovery of the polymerase chain reaction (PCR) in 1985 led to a revolution in medical diagnostics. Nucleic acid-based diagnostics are founded in the sequence-recognition abilities of oligonucleotides.[10-13] Probes recognize a chosen target sequence using Watson-Crick base pairing, whereby complimentary pairs are formed between either guanine and cytosine or adenine with either thymine or uracil. These sequences are long enough to be stable at a given running temperature, but short enough that the stability is compromised in the presence of a single mismatch. Probes may target DNA or RNA sequences, and may be made of DNA, RNA, or unnatural nucleic acid sequences such as TNA or LNA.[14, 15]

There are two types of probes that have served as the industry standard since the 1990s: molecular beacon probes and TaqMan probes. Molecular beacon probes are hairpins that feature a fluorophore and quencher on opposite ends of the sequence (Fig. 1A).[16] When unbound in solution, fluorescence is minimized due to the proximity of the quencher to the fluorophore; however, upon hybridization with the target sequence, the fluorophore and quencher are separated in space, allowing for an increase in fluorescence. The flexibility in fluorophore choice allows for the utilization of molecular beacon probes for multiplex detection, as a specific spectral signature can be tied to a single desired target.[17, 18] TaqMan probes also are comprised of a single strand of nucleic acids with fluorophore and quencher labels (Fig. 1B).[19] Here, the fluorophore and quencher are placed on nucleotides in proximity in the sequence, opposed to relying on the formation of the stem to bring them together. The 5'->3' exonuclease activity of Taq polymerase in the PCR assay enables the excision of the fluorophore and quencher, allowing them to separate in solution, ultimately leading to the increase in fluorescent signal. A primary strength of the molecular beacon probe and TaqMan systems is the ability to utilize multiple fluorophores to differentiate between targets in the same sample, which has been referred to as multiplex detection.

Molecular beacons do not preclude the use of the need for target amplification, however, which is true of all commonly utilized nucleic-acid-based diagnostic tests. A variety of isothermal amplification techniques have been suggested for use with molecular beacons for clinical applications, reducing the need for expensive thermocyclers or laborious usage of controlled incubators or water baths.[20] These amplification techniques generally utilize relatively lower temperatures, such as 37 °C, where molecular beacon probes may exhibit insufficient selectivity against single-nucleotide polymorphisms. This loss in selectivity is owed to the partial stability of the near-matched complex.

3
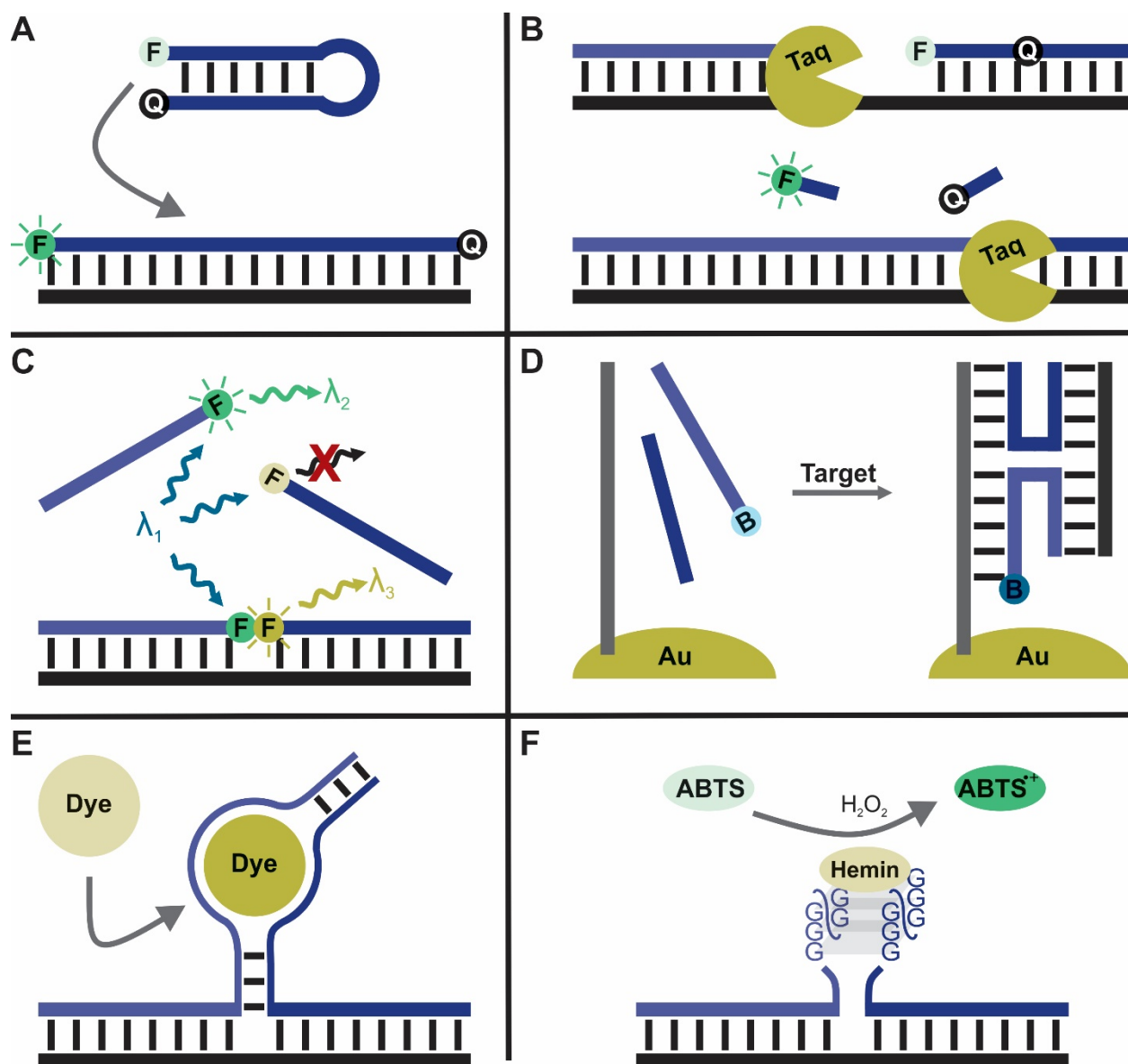
Figure 1: Nucleic-acid-based diagnostic tests. A) Molecular beacon probes exhibit an increase in fluorescence upon spatial separation of the fluorophore and quencher upon hybridization to the target. B) TaqMan probes use the 5'->3' endonuclease activity of *Taq* polymerase to degrade a probe while extending a primer in a PCR reaction, thereby separating the fluorophore and quencher. C) Adjacent Probes. A split probe system using FRET will only exhibit a signal increase at λ_3 if the two fluorophores are brought together by binding to the target. D) A 4WJ is assembled upon the introduction of the target sequence, bringing methylene blue into proximity of a gold nanoparticle or gold electrode and demonstrating a change in the voltametric signal. E) A fluorogen in solution exhibits an increase in fluorescence upon binding to a split-FLAP. F) A split-G4 with the cofactor hemin enables the peroxidation of a colorless substrate such as ABTS, generating a significant increase in absorbance.

As antibiotic resistance is often owed to single-nucleotide substitutions, this can create difficulties in interpreting the results of genotype-based drug susceptibility tests (DST). Nucleic acid hybridization utilizes flexible strands that "zip" together, whereby the difference in free energy between the matched and single-mismatched target are relatively small compared to the antibody-antigen.[21] A shorter target-recognizing sequence will increase the relative difference in stabilities between the matched and mismatched complexes but will also lower the general stability of the matched complex; as such, there is an inverse relationship between sensitivity, owed in part to the stability of the matched complex, and selectivity, owed to the relative difference in stabilities between the matched and mismatched complexes.

A compromise is found in splitting the target-recognizing sequence between two strands, such that the overall stability across the two strands is high, but a mismatch in either strand will heavily destabilize the complex.[22] The standard fluorophore-and-quencher signal transduction system needed to be adjusted for these probes, as the length is insufficient to fully disable quenching when compared to molecular beacons. Resonance energy transfer techniques such as FRET were suggested, whereby only when a donor and acceptor fluorophore are able to interact while the donor is excited may the desired emission signal be measured (Fig. 1C).[23] Several other fluorescence-generating systems that may undergo self-assembly have been utilized, such as excimers and lanthanide complexes.[24, 25]

Molecular beacon probes are sequence-specific to their target, which adds to the general cost due to the need for conjugating both the fluorophore and quencher to the probe being developed for every new target sequence, as well as the need to purify at each step of the process. However, it is possible to offset this requirement of the molecular beacon probe directly binding to the target by using a split probe approach, allowing for a single molecular beacon sequence to be used with a variety

of targets by customizing unmodified probes that act as intermediaries between the given target and the conserved molecular beacon sequence. In particular, a four-way junction design has been suggested for this purpose.[26] This design has been utilized with several readout systems, such as FRET and electrochemical outputs (Fig. 1C,D).[27-33] In all of these listed cases, however, covalent modifications, or labels, are required for signal generation to occur. This adds to the overall cost-per-test of the assay, which can be problematic for some applications or in resource-limited environments.

Label-Free Signal Transduction

Due to the relatively high cost of covalent-attachment for fluorophores and quenchers, there is a desire for label-free signal transduction methods. Light-up aptamers are one such signal transducer. An aptamer is a nucleic acid sequence that binds a specific, generally non-nucleic acid target molecule.[34] In this way, they are similar to antibodies; however, unlike antibodies that are produced *in vivo*, aptamers are discovered *in vitro* using nucleic acids through the systematic evolution of ligands by exponential enrichment (SELEX).[35-38] SELEX utilizes changes in affinity to isolate nucleic acid sequences from a random library that bind the target molecule to create a new library generation, which will then undergo greater scrutiny until a product with the desired properties is reached. Further, the use of nucleic acids allows for more consistent refolding, adding to long-term stability through refreeze cycles.[39]

Certain molecules exhibit significant increases in fluorescent yield upon binding to these nucleic acid structures.[40] These environmentally-sensitive chromophores are referred to as fluorogens.[41] Upon excitation, a fluorogen will fluoresce if its preferred methods of relaxation are denied; such methods include relaxing a highly dipolar excited state through solvent interactions and

internal rotation as seen in molecular rotors. A series of aptamers have been discovered that target several of these molecules, allowing the molecules to "light-up" in the presence of the aptamer, commonly referred to as Fluorescent Light-Up Aptamers or FLAPs.[42-56] FLAPs have several key advantages over more traditional fluorophore approaches. The fluorescent dyes are not covalently bound to the target, so expensive conjugation chemistry and purification are unnecessary. Further, the structures are self-assembling, allowing them to be utilized with a split probe system (Fig. 1E). Additionally, FLAPs have been selected for using either RNA or DNA; while RNA offers greater functional flexibility due to the presence of its 2'-OH, DNA offers greater chemical stability and a lowered production cost; however, even RNA is cheaper than employing conjugation chemistry.[44] These advantages provide a path for label-free detection of single-nucleotide polymorphisms through the use of split-FLAPs.

Many aptamers feature a G-quadruplex (G4) scaffold in order to provide the correct geometry for a functional binding pocket.[57-59] A G4 is a four-stranded helix comprised of two-or-more G-tetrads, each comprised of four guanine residues engaged in four hydrogen bonds to create a square-like plane (Fig. 2A,B). While a G4 may be composed of a single oligonucleotide string, they may also be formed of two-or-more individual pieces that will refold into the G4 when brought into close proximity of each other. There are several ways that a G4 may fold. If all four strands are in the same 5-to-3 direction of propagation, it is considered parallel; if two strands are in each direction it is antiparallel, and three strands in a single direction with one in the other is considered hybrid.[60] Preference for a sequence to fold into one structure over another is influenced by the types of monovalent ions present, with high potassium concentrations favoring parallel and hybrid conformations and high sodium concentrations favoring antiparallel formations.[61, 62] Further, the length and nucleotide composition of
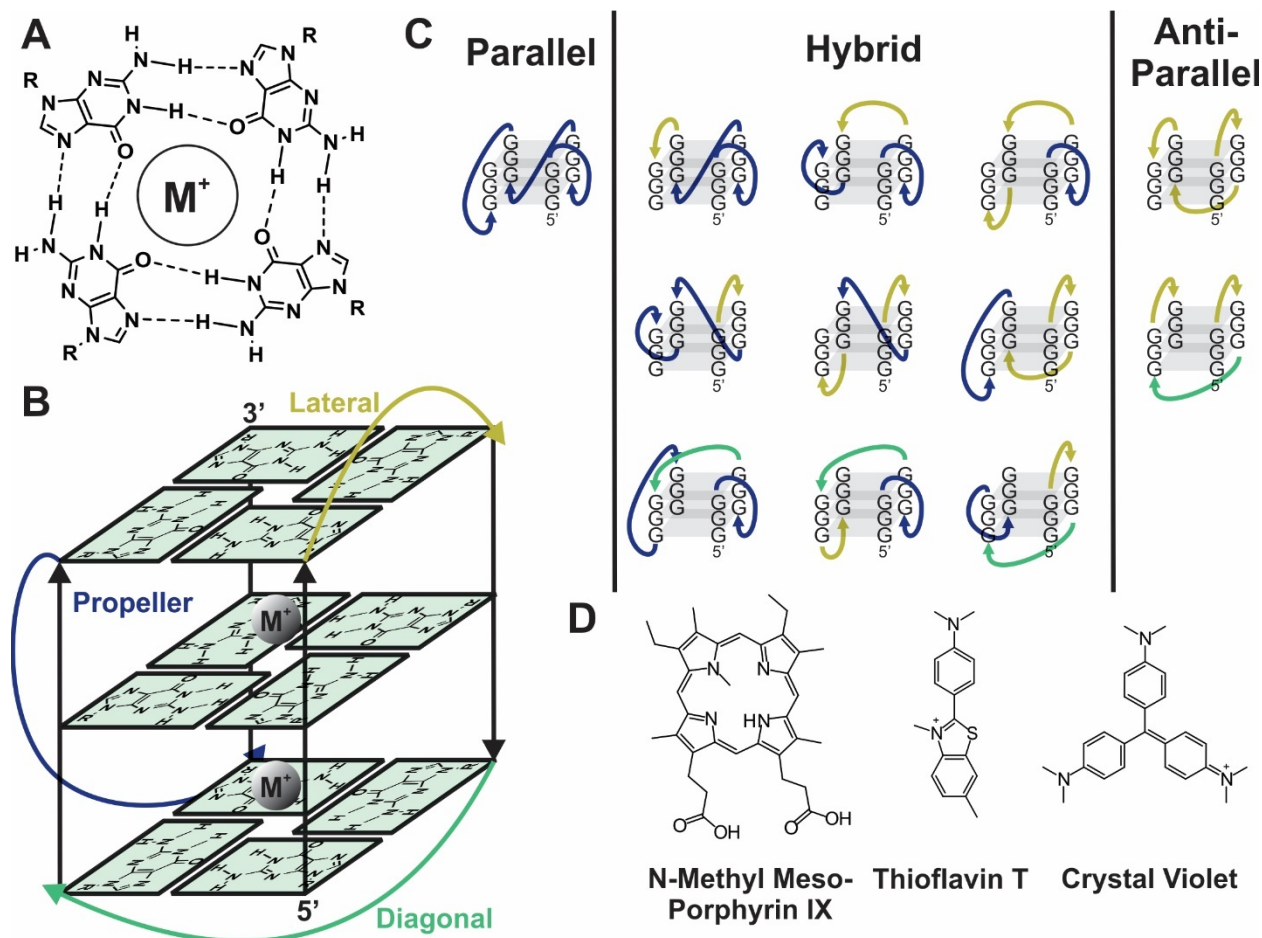
Figure 2: G-quadruplex structure. A) A G-tetrad is formed via hydrogen-bonding between two edges of each of four guanines. B) Multiple G-tetrads stack, producing the G-quadruplex. The inclusion of metal ions such as sodium or potassium are required to counterbalance the electronegativity of the carbonyls. Nucleic acid residues between guanine runs are referred to as "loops." These loops may either offer a continuation of the direction of propagation of the helix (Propeller) or offer an inversion in the direction (Lateral and Diagonal). C) Combining these three loop types provides a variety of G4 structures. Parallel G4s contain all four strands in the same 5'->3' orientation. Anti-parallel G4s have two strands in each direction. Hybrid G4s have three strands in one direction and one in the other. D) *N*-methyl mesoporphyrin IX (NMM) and thioflavin T (TFT) exhibit increased fluorescence in the presence of non-anti-parallel G4s. Crystal Violet (CV) exhibits increased fluorescence when bound to antiparallel G4s.

the sequences between G-rich runs may favor certain folding patterns. Longer fragments have a tendency to fold into lateral loops which provide an inversion of the direction of propagation, whereas shorter fragments favor propeller-like loops that maintain the direction of propagation (Fig. 2B). These factors lead to numerous molecular scaffolds onto which a ligand may bind (Fig. 2C).

G4s have been shown to bind a variety of light-up dyes such as thioflavin T (TFT), crystal violet (CV), and N-methylmesoporphyrin IX (NMM) (Fig. 2D).[63] The specific turn-on properties of these dyes vary with the structural folding of the G4, with NMM preferring parallel structures and CV preferring antiparallel ones.[64-66] A split probe design with G-rich sequences extending from the abutting ends of the two probes, paired with one of these light-up dyes, has been previously utilized for SNP detection.[67-70] This split-signal-transducer design is not limited to simple G4s; FLAPs have been successfully split for this purpose as well.[71-74] While the additional sequence reporter added to the target-recognizing probe is generally shorter for G4-based probes than split-aptamer probes, they may struggle to interrogate SNPs in C-rich sequences. Further, as many light-up split-aptamer probes include G4s as structural elements, many of them may interact with the G4-specific dyes. As such, light-up aptamers that contain a G4 structure may be inappropriate to include with a split-G4 as a multiplex system.

G4s have also been shown to exhibit peroxidase-like activity when paired with the cofactor hemin (Fig. 1F).[75, 76] This allows a G4 to act as a functional mimic of horseradish peroxidase, generating a color-based signal with any of the colorless small molecules that horseradish peroxidase has been utilized with, such as 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS). A split-G4-based sensor has been previously utilized in several works, offering a colorimetric signal indicative of a given nucleic acid input.[76-78] While ABTS is the most commonly used chromogenic substrate with the

G4:hemin complex, others like 3,3',5,5'-tetramethylbenzidine and 3,3'-diaminobenzidine have seen usage.[79, 80] Further, this peroxidation reaction has been paired with luminol as a chemiluminescence assay.[81] It is important to note that the G4:hemin complex will indiscriminately act on these substrates if multiple are present; that is to say, it is as-of-yet impossible to produce differentiable colorimetric signals from multiple G4 probes.

Research Questions

In these works, we sought to advance the development of signal transducers for the detection of antibiotic-resistance-conferring single-nucleotide polymorphisms. To this end, we focused on two systems. First, we examined the utility of the instrument-free, colorimetric capabilities of the split-G4:hemin complex. We sought to improve upon the consistency of the system through the use of multiple sensors for the target, as well as open a pathway for more complex analysis of multiple target sequences. As such, we set out to combine a series of sensors into molecular logic gates to provide complex analysis. Our works on the utilization of these molecular logic gates are described in within Chapter 2.

Further, we wished to optimize the G4:hemin system. For split-G4-based probes, the optimal splitting pattern had only been determined using structureless targets. We sought to determine the optimal splitting pattern for the G4 through the use of medically-relevant targets that often fold into stable secondary/tertiary structures. Additionally, studies into the modification of the loop sequences were performed. The results of these studies may be found in Chapter 3.

The second label-free system we studied was a DNA FLAP, DAP-10-42, originally selected to enhance the fluorescence of dapoxyl sulfonyl dyes. In our works with this aptamer system, we

discovered that it demonstrates promiscuity with a variety of arylmethane dyes. This property of the system allows for a choice of a fluorogen depending on the required spectroscopic properties. It also offers a route towards utilizing one aptamer with one of several dyes to be paired with a secondary FLAP for multiplexing. We used DAP-10-42 split aptamer as a signal transduction unit for molecular logic gates, providing a fluorescent readout that could be seen when excited with a handheld UV light source. Our studies into this promiscuity and its application to split aptamer probe-based nucleic acid analysis can be found in Chapter 4.

We further hypothesized that nucleotide substitutions in the aptameric sequence may optimize the binding efficiency of DAP-10-42 with these dyes, and therefore the dyes' fluorescent turn-on. Further, we hypothesized that some of these changes would offer greater selectivity in which dyes exhibit enhanced fluorescence upon binding. Our exploration into sequence modifications that optimize the fluorescent activity of DAP-10-42 with a variety of dyes, can be found in Chapter 5.

CHAPTER TWO: SPLIT G-QUADRUPLEX SENSORS WITH TURN-ON / TURN-OFF CAPABILITIES

Introduction

In this chapter, we present our work into the utilization of split G-quadruplex-based sensors for selective detection of nucleic acid targets. Here, we sought to create an assay for the detection of rifampin resistance in *Mycobacterium tuberculosis* based on the presence or absence of a SNS in an 81-bp region of the *rpoB* gene. While *tuberculosis* infections are primarily found in developing regions with limited resources, current detection methods for antibiotic resistance are either prohibitively slow and have extensive biosafety requirements in the case of standard culture or require expensive instrumentation in the case of qPCR. To solve these issues, we devised an assay that utilizes the peroxidase-like activity of the G-quadruplex:hemin construct to generate a colorimetric signal that can be read by the unaided eye. Some sensors were designed to provide a strong color in the presence of a specific DNA target (a "turn-on" sensor), while others were to provide a strong color in the absence of specific targets (a "turn-off" sensor).

We combined these turn-on and turn-off sensors by arranging them into a 15-segment array that would produce an alphanumeric character depending on the input. In this way, the results of multiple sensors would be determined as a single combinatory output. For multiplex detection, it would be ideal to have different colors indicate different possible outputs, and a variety of peroxidation targets exist that exhibit one of several possible colors; however, there is currently no method of linking a given sensor to a specific color without expensive conjugation chemistry. The utilization of logic functions with sensors arranged into a 2D array was hypothesized to be a viable method for combining multiple signals into a single output. The paper below details our findings.

Alphanumeric Visual Display Made of DNA Logic Gates
for Susceptibility Testing of Pathogens

Devices that use individual molecules as computational elements have been under experimental validation since 1993.[82] DNA-based biocomputing devices are promising for the analysis of nucleic acids.[83-91] However, to the best of our knowledge there is no broad-scale use of molecular computers or logic circuits. At the same time, biomedical diagnostic still relies on the manual or electronic computer-based analysis of outputs from simple digital sensors with YES (digital 1) or NO (digital 0) response.[92] We hypothesize that analyses can be performed by molecular logic circuits/processors, which would provide a user-friendly output. This study is devoted to the development of such an output/interface. The alphanumerical display reported here takes advantage of logic circuits made of DNA to simplify medical diagnostics, thus making it accessible by a general user. We demonstrated the capability of the DNA logic circuits to perform sophisticated analysis of a point mutation at room temperature. One key advantage of the system is the improvement in reliability conferred by the analysis of each analyte by several sensors, as detailed below. The display produces a user-friendly visible output that can be read by the naked eye.

In this proof-of-concept study, we report on the design of an array of DNA logic gates producing an alphanumerical character as an output, and its implementation for analysis and drug-susceptibility testing of *Mycobacterium tuberculosis complex* (*MTC*). *MTC* is a group of bacteria that are the causative agent of tuberculosis (TB), a potentially lethal infectious disease. The treatment of TB is complicated due to the development of resistance to one or several drugs.[93, 94] Rifampin (RIF) is a first-line antitubercular agent targeting bacterial RNA polymerase. Point-mutations in an 81-bp "hot

spot" region (aa 507-533) of the *rpoB* gene are known to render MTC resistance to RIF. The most prevalent drug-resistance causing point-mutation is a single-nucleotide substitution (SNS) at codon 526.[95-97] Therefore, in this study, we tailored DNA logic gates to respond to DNA inputs RIF[S] and RIF[R], which correspond to a fragment of the *rpoB* gene containing codon 526 from RIF-susceptible and RIF resistant MTC, respectively (App. B: Table 5).

The logic gates were arranged into a 3×5 display that would produce an alphanumerical character – the letter "O" in case of n<u>O</u> TB outcome; and either the letter "P" or "F" if the test for RIF susceptibility <u>P</u>assed or <u>F</u>ailed, respectively (Fig. 3A, left images). To achieve this pattern, the wells *d3*, *e2* and *e3* should stay colorless when samples are TB-positive; well *c2* should turn ON in response to either RIF-sensitive or RIF-resistant MTC; and wells *b3* and *c3* needed to turn OFF only in the case of the RIF-resistant bacteria (Fig. 3B). Wells *a1*, *a2*, *a3*, *b1*, *c1*, *d1* and *e1* of the display (Fig. 3B) should remain colored, while wells *b2* and *d2* stayed colorless disregarding the presence or absence of TB. This display set up required the following logic gates: PASS 1; PASS 0; OR; NOR and NOT[R], with the truth tables as shown in Figure 3C.

As a PASS 1 logic gate, which produces a high signal (digital output 1) disregarding the presence of a DNA input, we used a peroxidase-like deoxyribozyme, PW17.[98-101] This G-rich DNA oligomer forms a G-quadruplex (G4) structure stabilized by cations, such as K^+ (Fig. 4A, left).[98] In the presence of hemin, PW17 demonstrates a catalytic activity for peroxidation of a colorless organic compound (e.g. ABTS, 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid)) into a colored product (Fig. 4A, right). The catalytic activity of PW17 did not depend on the presence of RIF[S] or RIF[R], thereby the wells containing PW17 were colored both in the absence and in the presence of the
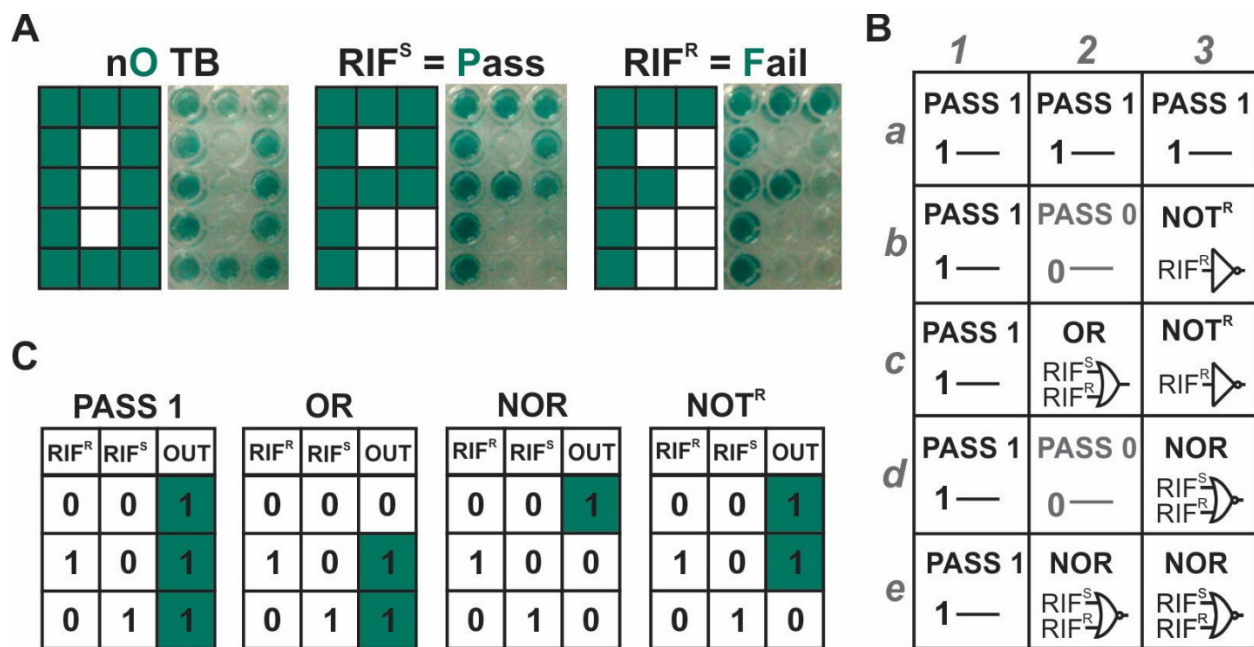
Figure 3: The design and performance of alphanumerical sensor display. A) Theoretical (left) and experimental (right) observation of the display in the absence of MTC DNA ("no TB"); in the presence of DNA inputs corresponding to MTC with either RIF-sensitive or RIF-resistant phenotype ("RIF^S" or "RIF^R", respectively). PASS 0 gate wells contained just a working buffer with no gate strands. PW17 at 0.2 μM served as PASS 1 gate; OR and NOR strands were at 1 μM and 0.5 μM, respectively; and for NOT^R function, YES^S and NOR strands were at 0.8 μM and 0.4 μM, respectively. B) Arrangement of G4 DNA logic gates into a 3×5 graphical display. The logic gate names and symbols are indicated corresponding to the content of each well. The two wells containing a mixture of YES^S and NOR gate working in parallel are marked with a symbol for NOT^R gate, since this combination computes logic negation of RIF^R input (see App. B for detailed explanation). C) Truth tables for the G4 logic gates used in the display. High output is highlighted green.

inputs. The lack of PW17, or any other DNA logic gate strands, resulted in PASS 0 logic – a low output (absence of color) would be detected regardless of input.

The signal-generating modules of other DNA logic gates designed in this study were based on PW17. For this purpose, the deoxyribozyme sequence was split into two subunits (strands A and B), each of which contained two of the four G-triplets of PW17. Each subunit was elongated with the sequence constituting a communicating module that contained the input-recognition site. Association of strands A and B enabled formation of the catalytically active G4 peroxidase resulting in generation

of green color (high output). Using this "split" approach, we designed YES^S and OR gates (Fig. 4B), which turned-on in response to input(s).[77, 102, 103] In contrast, for a NOR gate, addition of input(s) resulted in dissociation of the G4 structure (Fig. 4C).

The input-recognition site of a two-input G4-OR gate was complementary to the 31-nt 5'-terminal fragment of RIF^S and RIF^R, which was identical in both inputs (App. B: Fig. 18A, Table 6). Therefore, high output was expected in the presence of either RIF^S or RIF^R, but not in the absence of either inputs (Fig. 4A). The proper logical behavior of the individual gate was achieved with clear difference in color intensity between the absence and the presence of either RIF^S or RIF^R input (App. B: Fig. 18B). Formation of the high output DNA complex (OR-A/OR-B/Input) was also demonstrated by the appearance of a low mobility band in native PAAG (App. B: Fig. 18C).

The G4-YES^S gate computing the identity function of RIF^S input was designed similar to G4-OR gate. The only difference was in the input-recognition site of the gate; it was complementary to a 28-nt fragment in the middle of the input sequence containing a C>T substitution in RIF^R input (App. B: Fig. 19A). YES^S-B strand was chosen to have a short 8-nt RIF^S-recognition region to prevent its binding to the SNS-containing RIF^R input. YES^S-A strand was designed long enough to promote efficient binding of RIF^S input to the gate. This feature of the strand enabled it to bind RIF^R as well (App. B: Fig. 19B). At the same time, the combination of long YES^S-A with short YES^S-B prevented efficient association of YES^S-A/YES^S-B/RIF^R, thus ensuring low output in response to RIF^G. We also introduced a G15A substitution in the input-recognition domain of YES^S-A to sequester G4-forming fragment into a stable stem structure in its dissociated state (App B: Fig. 19, compare structures in panels A and C), which helped to minimize background signal. Overall, formation of high input
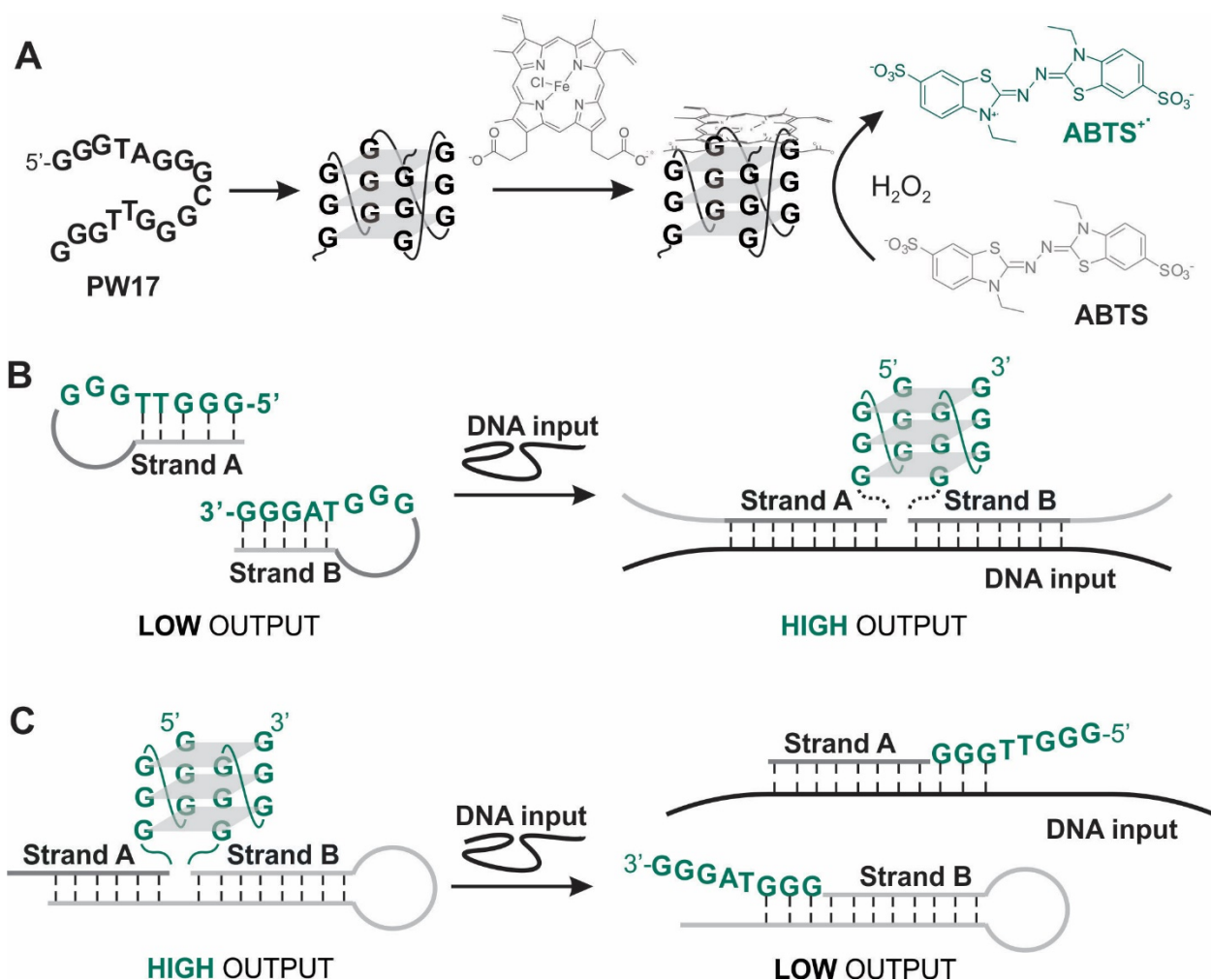
Figure 4: Design of G-quadruplex (G4)-based DNA logic gates with visually detected output. A) PW17[98] folds into a parallel G4 structure with the affinity to hemin. The PW17-hemin complex catalyses peroxidation of a colourless organic substrate ABTS into a blue-green ABTS⁺•. PW17 will produce high output independent on the analyte presence, which corresponds to PASS 1 logic. B) YES and OR gates. Strands A and B of the gate associate into a catalytically active G4 structure ("High output") only in the presence of a specific DNA input. OR gate targets the input fragment outside an SNS position and thus responds to either RIF^S or RIF^R input. YES gate targets the SNS-containing input fragment and selectively binds to RIF^S only. C) G4-based NOR gate. In the absence of a DNA input (either RIF^S or RIF^R), strands A and B are bound to each other forming active G4 peroxidase structure ("High output"). The input binds to strand A, thus displacing it from the complex and turning off the peroxidase activity ("Low output").

complex and proper digital behavior of YES^S gate was observed as expected only in the presence of RIF^S, but not RIF^R (App. B: Fig. 19, panels D, E).

The G4-NOR gate was designed to exhibit high signal only in the absence of inputs, while the presence of either RIF^S or RIF^R would destabilize the G4 structure resulting in low output (Fig. 3C, Fig. 4A). The NOR-B strand contained a 5'-terminal domain complementary to the 3'-terminal part of the NOR-A strand to enable association of the strands into a catalytically active G4 structure in the absence of inputs (App. B: Fig. 20A). At the same time, the 3'-end of the NOR-A strand consisted of the input-recognition site, with the first 4 nt being complementary only to the input sequence, but not to NOR-B. This 4-nt fragment served as a "toehold" for input-mediated displacement of NOR-A from the NOR-A/NOR-B complex, which turned the gate off. Expectedly, since NOR-A bound to the input outside the mutation site, the gate did not discriminate between RIF^S and RIF^R (App. B: Fig. 20B). The formation of the NOR-A/NOR-B complex in the absence of the inputs, as well as the NOR-A/Input complex in the presence of either RIF^S or RIF^R was confirmed by gel-electrophoresis (App. B: Fig. 20C).

Finally, we mixed YES^S and NOR gates to obtain the NOT^R logic function (App. B: Fig. 21A) – the presence of RIF^R input (binary input 1) resulted in low output (absence of the color), while high output was observed either without the inputs or in the presence of RIF^S (Fig. 4A). YES^S and NOR gates competed for the input but otherwise did not interfere with each other, as can be seen by the proper pattern of color development and by gel analysis (App. B: Fig. 21, panels B and C). It should be noted that the capability of the display to differentiate between RIF^S and RIF^R inputs relies on the specificity of the NOT^R logic function. It was demonstrated that RIF^R input triggered a 2-3-fold

increase in absorbance compared to the response to RIF^S input, which enabled reliable discrimination of inputs even by the naked eye (App. B: Fig. 21C).

The high output signal for the YES^S, OR, and NOR gates was generated within the first 10 min after addition of hydrogen peroxide and was stable at least for 60 min (App. B: Fig. 22). An input concentration equimolar to the concentration of the gate strands yielded a signal maxima, as expected. For the NOR gate, the signal plateaued at the RIF^S concentration of 0.5-2 μM, while for the YES^S and OR gates, the signal decreased when RIF^S was in excess over the gate strand, due to a heightened probability of forming catalytically inactive complexes between the input strand and one of the gate strands. As little as 30-70 nM of RIF^S was sufficient to statistically differentiate between the low and high output states of the YES^S, OR, and NOR gates. Visual differentiation can be achieved at the input concentration of 0.2-0.5 μM, which can be attained after isothermal amplification of the pathogen nucleic acid (e.g. using LAMP, RPA, NASBA).

Having confirmed proper digital behavior of all G4-gates, we arranged the five gates into a 3×5 display (Fig. 3B). The array wells developed color in the course of analysis. The correct pattern of colored wells in the presence of either RIF^S or RIF^R input, or in the absence of any MTC-specific input, was experimentally observed with no instrument required to read the display output (Fig. 3A, right images). In addition to RIF^R input, which corresponded to CAC>TAC substitution in *rpoB* codon 526, we tested the response of the display to input RIF^R-G (App. B: Table 6), an analog of RIF^R, containing CAC>GAC substitution, which is known to be the second most prevalent drug-resistance causing point mutation in *rpoB* codon 526.[95-97] The stable secondary structure of this input (App. B: Fig. 22A) made a constraint for its hybridization to the input-recognition region of the OR gate. This resulted in a weaker response of G4-OR gate to RIF^R-G than to the other two inputs (App. B: Fig.

18B). Despite this constraint, the display output triggered by the presence of RIF^R-G could be unambiguously read as "F", indicating a RIF-resistant phenotype. This highlights the advantage of logic gate-based display sensors: several logic units acting in parallel on a given input allow for unequivocal target classification. A similar principle is used in DNA microarrays, where multiple probes differing by one or two nucleotides interrogate the same target sequence to improve the test reliability.[104]

The power of multiple individualized reactions for the analysis of virus signature sequences has been recently demonstrated using a fluorescent display made of deoxyribozyme-based logic gates.[83, 90] The presence of a particular virus-related input was read as the first letter of the virus name upon visualization of the fluorescent read-out using a UV-light box after overnight incubation. In our present work, the displayed character can be read by the naked eye almost immediately after addition of the input. This is achieved due to the use of the peroxidase-like deoxyribozyme PW17 that meditated color change. The split design of the logic gates enabled high selectivity towards point mutations even at room temperature, according to the concept developed by us earlier.[26] This differentiation power capability of the logic gate allowed for the input classification according to MTC RIF-susceptibility genotype. Moreover, the multi-stranded design of DNA logic gates enabled fine-tuning of the gate performance by independent optimization of the design and concentration of each strand.[22] For example, having long input-recognition regions of one or both strands allowed efficient hybridization, even to folded inputs with stable secondary structures.[105]

Constructs with different splitting modes that reconstitute into active G4-peroxidase have been reported earlier.[76, 106-108] For example, a PW17-based probe with the G4-forming sequence split asymmetrically (with one strand containing three G-triples and another having one G-triplet) was used

20

to detect point mutations in the *rpoB* gene of rifampin-resistant *M. tuberculosis*.[106] In this case, an additional competition probe sequestering a mutation-containing sequence was required for the reliable identification of the wild-type sequence, and vice versa. The use of several logic constructs responding to the same input sequence, as reported in the present work, not only differentiates point-mutations, but also provides a user-friendly read-out.

In conclusion, we have developed an array of DNA logic gates based on G4 peroxidase, which displayed a specific alphanumerical character in accordance with the sequence of a pathogen-related DNA input. Importantly, the display did not require any instrumentation for the character read-out, which could be done with the naked eye merely minutes after adding the input. In this proof-of-concept work, the logic gates of the display responded to a fragment of the *M. tuberculosis* complex *rpoB* gene to unambiguously distinguish the absence of the input from the presence of either wild type (classified as RIF-sensitive bacterial phenotype) or SNS-containing input (classified as RIF-resistant phenotype). Each analyte was interrogated by several gates (YES^S, OR, NOR), increasing the reliability of detection. Even though this proof-of concept study reports preliminary results, it highlights the advantages of logic gate-based "intelligent" sensing devices for autonomous and less ambiguous detection of infectious diseases and contributes to the repertoire of instrument-free molecular graphics processing units, which may stimulate their industrial uses.

Experimental Section

DNA oligonucleotides were purchased from IDT, Inc. (Cornville, IA, USA) and used without purification. Hemin, ABTS and H_2O_2 were purchased from Sigma-Aldrich (St. Louis, MO, USA). For a visual (colorimetric) assay, samples containing strands of DNA logic gates (0.5 μM for NOR gate and 1 μM for YES^S and OR gates) were mixed with DNA inputs (1 μM) in a colorimetric buffer (50

mM HEPES-NaOH, pH 7.4, 50 mM $MgCl_2$, 20 mM KCl, 120 mM NaCl, 1% DMSO, and 0.03% Triton X-100) at room temperature followed by addition of ABTS (1 mM), hemin (0.375 μM) and H_2O_2 (1 mM) to observe the development of a blue-green colour. For quantification purposes, the absorbance of the samples at 420 nm was measured using a Thermo Scientific NanoDrop One[C] UV-Vis Spectrophotometer (Waltham, MA, USA). For gel electrophoresis analysis of low output and high output complexes of the gates, the samples containing gate strands in the absence or presence of inputs were mixed at concentrations used for the assay, loaded on 12% or 15% PAAG supplemented with $MgCl_2$ (10 mM) and KCl (10 mM) and run for 2 h at 80V. Alphanumerical display was made by pipetting the DNA logic gate strands into designated positions of a well plate, each of which contained the colorimetric buffer and one of the inputs. The character corresponding to the input sequence was observed upon addition of hemin, ABTS and H_2O_2.

Further Works on DNA Logic Gates
for Susceptibility Testing of Pathogens

The paper above focused on the most prevalent mutation in the aa 507-533 hotspot of the *rpoB* gene.[95-97] As the 526 CAC>TAC mutation represents only a plurality of rifampin-conferring mutations in the gene, we sought to expand our array by incorporating additional mutation sites. The 531 and 516 codons are the most-common non-526 mutations in the hotspot.[109, 110] As the 516 site is readily accessible in a suggested NASBA amplicon for the hotspot region, we first focused on the 516 GAC>GGC mutant in order to evaluate the viability of our colorimetric, alphanumeric array for multi-codon assessment.

Using the previously reported logic system (Fig. 4, panels B and C), output for the wildtype selective logic (YES^S) would need to yield an elevated signal only in the case that both codons 516 and 526 were of the wildtype sequence, which corresponds to logical conjunction – AND logic gate. This prevents the inclusion of two simple YES^S probes (YES^S-516A, YES^S-526C) to make the required AND logic, as the presence of a single-mutant sequence would still yield a high-output overall. Accordingly, the logic system was heavily rewritten to incorporate either NOT-S or YES^R logic opposed to YES^S logic (Fig. 7, left panel). This necessitates the inclusion of a YES^R probe for each given mutant sequence. The new system incorporated PASS-1 and PASS-0 gates as before, with a non-selective OR2 gate, non-selective NOR2 gate, and a YES^R system that combined YES^R-516G and YES^R-526T (Fig. 5B). The final array was designed to read "P" for "Pass," implying that there is an *MTB* infection and rifampin may be used, "R" for "Resistant," implying that there is an *MTB* infection but it will resist rifampin, or "F" for "Fail," the result that will be present if the test failed to identify a *MTB* infection (see below).
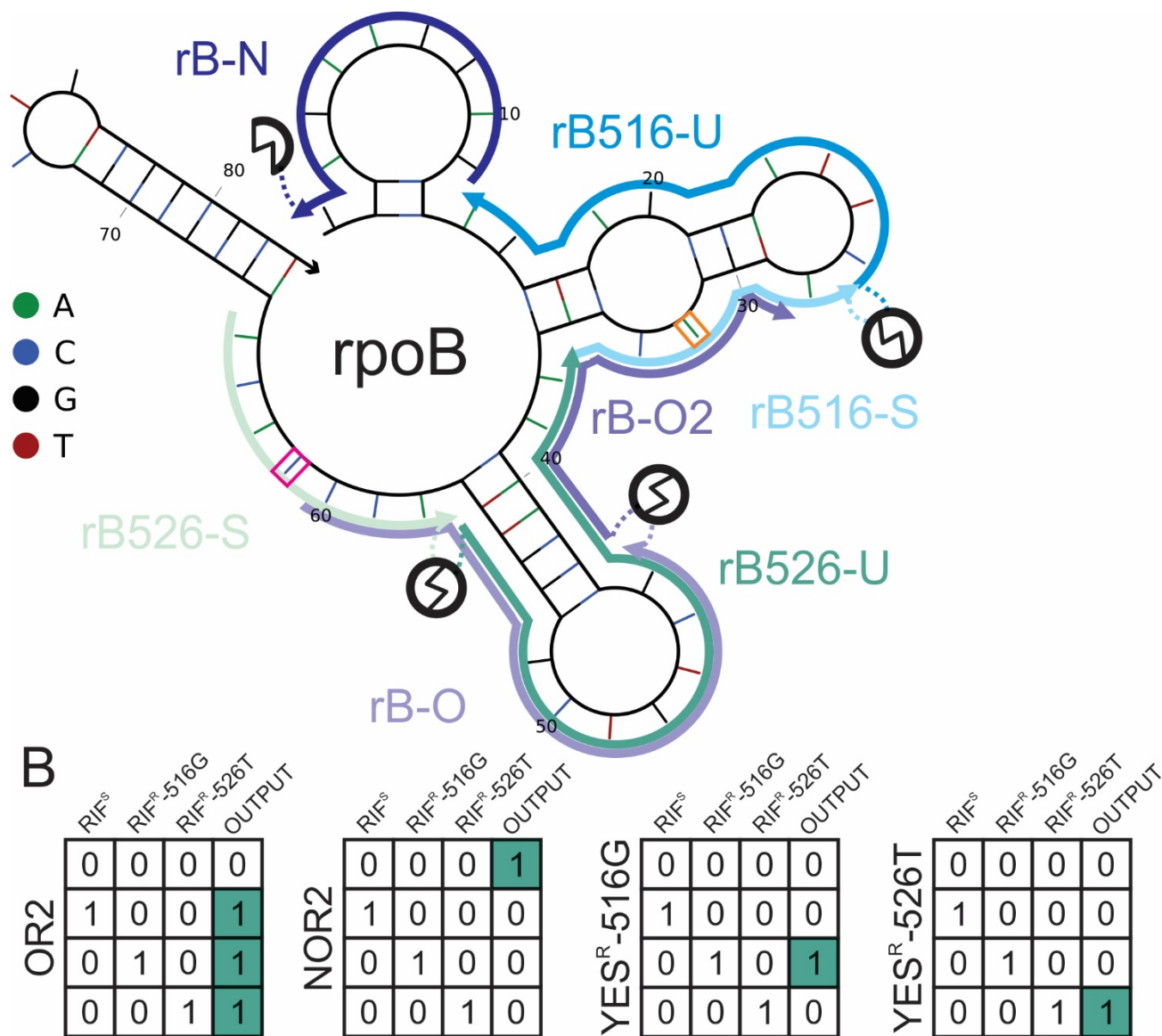
Figure 5: Layout of the *rpoB* fragment interrogated by the 516/526 sensors. A) The target synthetic *rpoB* sequence includes a purine rich region resulting from NASBA amplification. Structure of the synthetic target was predicted using NUPACK software.[111] To create a more universal turn-off sensor, rB-N was designed to target this sequence (NOR2 gate). The OR2 gate (rB-O and rB-O2) was designed to cover a region large enough that the sensor would not be susceptible to either mutation site studied. YES[R]-516G (rB516-U and rB516-S) and YES[R]-526T (rB526-U and rB526-S) were designed such that one probe would be unstable when hybridized with non-specific targets and no overlapping regions were included. B) Truth tables for the four basic logic gates used in this study. For the first three columns, a 0 indicates the absence of the listed analyte and a 1 indicates the presence of that analyte. For the fourth column, a 0 indicates a low output signal, while a 1 indicates a high output signal, as designated by the green color.
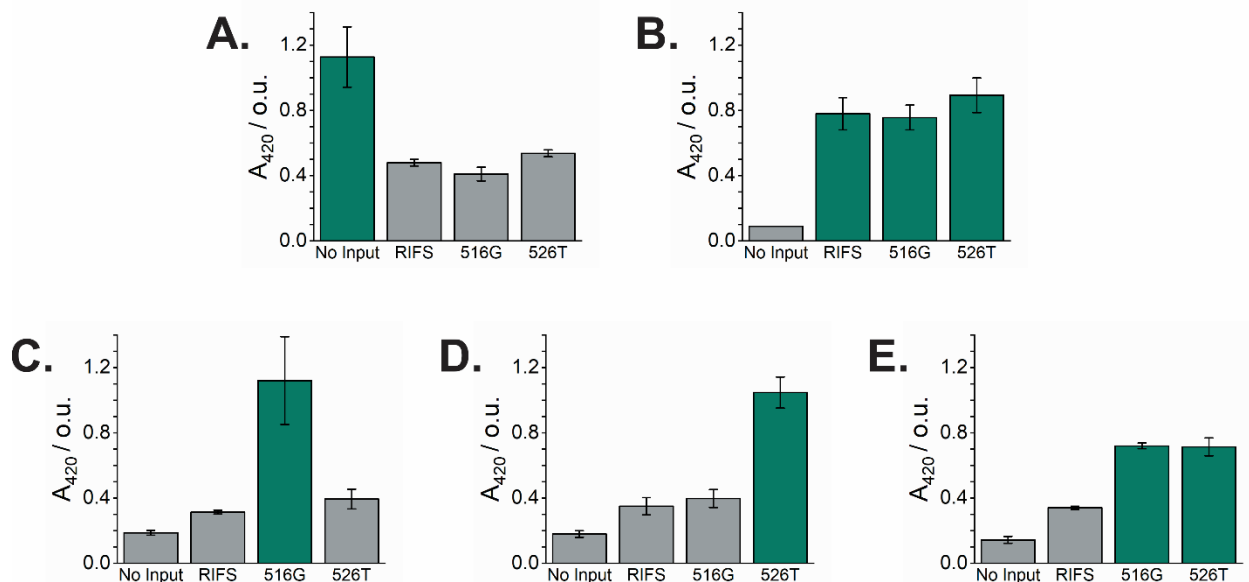
Figure 6: Response of NOR2 (**A**), OR2 (**B**), YES[R]-516G (**C**), YES[R]-526T (**D**), and YES[R] (**E**) gates in the absence ("No Input") or presence of RIF[S], RIF[R]-516G, and RIF[R]-526T *rpoB*. Green bars indicate a color sufficient to visualize with the unaided eye, whereas grey bars indicate insufficient reaction to differentiate from the No Input samples without instrumental assistance. The data are average values for three independent measurements with standard deviations as error bars.

A series of new logic gates were designed (App. B: Table 6). The new NOR2 gate was designed as before (Fig. 4C) using strands rB-N and rB-N2 (App. B: Table 6), but with the target-binding region covering the 5' end of a synthetic target mimicking a NASBA amplicon (Fig. 5). The YES[R]-526T gate comprised of strands rB526-U and rB526-S (Figure 5, App. B: Table 6) targeted the same region as the previously reported YES[S] gate, excepting that 526TAC was targeted opposed to 526CAC. The YES[R]-516G gate, consisting of strands rB516-U and rB516-S, spanned the gap between the NOR and YES[R]-526T gates, with the short rB516-S sequence abutting the rB526-U sequence (Fig. 5, App. B: Table 6). The OR2 gate spanned the region covering aas 517-524 with the strands rB-O and rB-O2 (Fig. 5, App. B: Table 6). As the OR2 gate and YES[R] gates would not be included in the same reaction vessel, overlaps of the target fragment interrogated by the strands of these logic gates could be
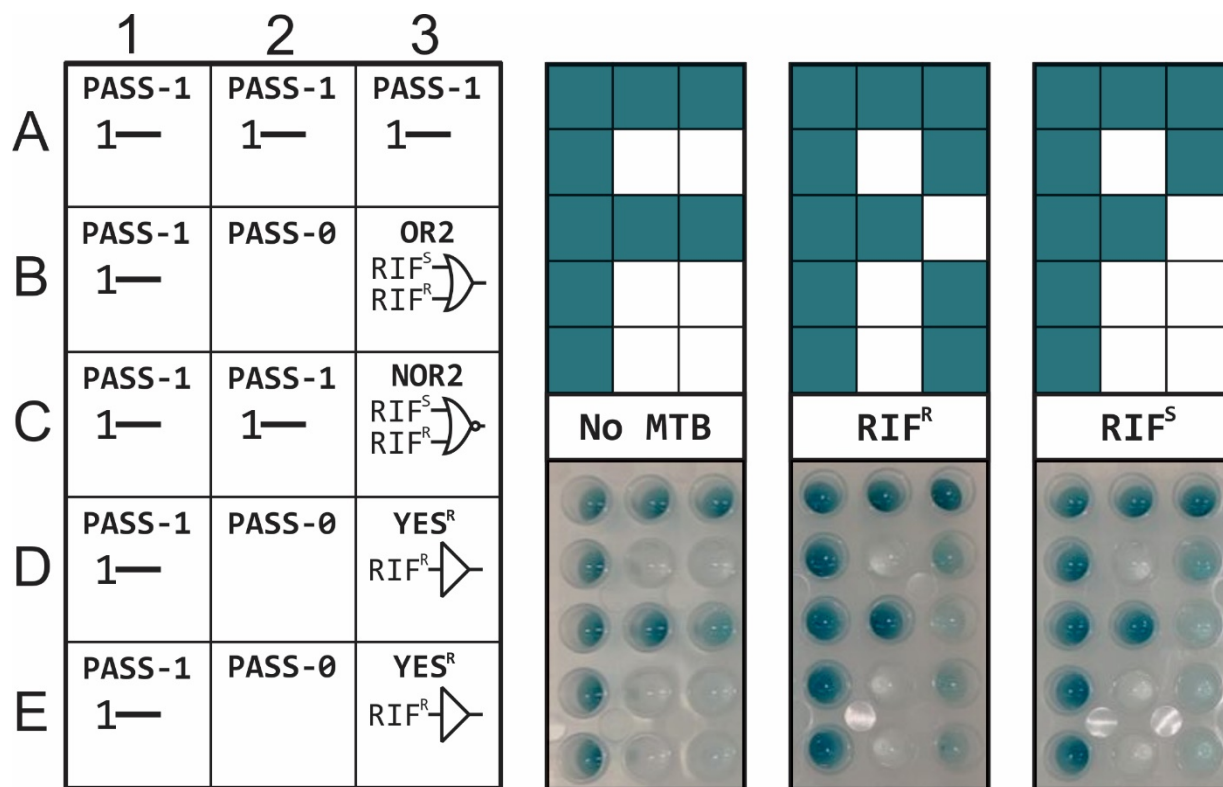
Figure 7: Layout of sensors required for array for the detection of RIF[R]-conferring mutations in codons 516 and 526 of *rpoB*. An "F" for "fail" appears if no *rpoB* is detected. An "R" for "resistant" appears in the presence of either a 516 A>G or 526 C>T mutation is present. If the RIF[S] genotype is present, a "P" for "pass" appears. RIF[R] image is indicative of the performance of the sensors in the presence of RIF[R]-516G.

tolerated. The OR2 gate was designed to be sufficiently long to be indifferent to mutations at either the 516 or 526 site at ambient temperature.

The performance of each gate was tested independently as described previously. At a target concentration of 1 μM, the four gates (NOR2, OR2, YES[R]-516G, YES[R]-526T) all demonstrated satisfactory signal differentiation between turn-on and turn-off conditions quantitatively via UV-Vis measurements (Fig. 6). While the combined YES[R] sensor set demonstrated a lower overall turn-on signal (0.72 ± 0.017 AU for RIF[R]-516G, 0.71 ± 0.055 AU for RIF[R]-526T) compared to the

independent sets (1.1 ± 0.27 AU for YES^R-516G, 1.05 ± 0.095 AU for YES^R-526T), the turn-off signals in the presence of the wild type were similar (0.19 ± 0.02 AU for YES^R-516G, 0.18 ± 0.02 AU for YES^R-526T, 0.14 ± 0.02 AU for YES^R).

The sensors were combined into an array to ensure high/low signal in the wells to make a character P in the presence of the wild type sequence of the target, to make a character R ("resistance") in the presence of either *rpoB*-516GGC or *rpoB*-526TAC, and a character F in the absence of the *Mtb*-related nucleic acid sequence (Fig. 7). While the performance of the NOR2 was satisfactory, as seen in well C3, the OR2 gate in B3 demonstrated a lower-than-desired turn-on. Further, the YES^R gate's lower overall turn-on signal was notable, as can be seen in wells D3 and E3. This would prevent differentiation between rifampin-resistant cases of tuberculosis and tuberculosis-negative cases.

While there are several possibilities for improving the performance of this system, one option is to improve the catalytic efficiency of the split-G4 probes. Improved peroxidation activity from the split probes would allow for a lower limit of detection; further, reduced interaction between probes would allow for a greater number of probes to be utilized inside a single gate system. As such, we turned our attentions towards the optimization of the split-G4 as a signal reporter.

CHAPTER THREE: OPTIMIZATION OF SPLIT G-QUADRUPLEX PROBES

Introduction

While numerous publications had reported the usage of split-G4 probes as signal transducing units, there was a dearth in information regarding optimization of the split-G4's performance towards a given target sequence. In particular, designing split-G4 probes, especially for the discrimination of a C>T mutation that is prone to wobble base mismatches, may be challenging.

Most split-G4 probes in literature report utilizing 12 guanine residues in the formation of the G4 unit. However, a variety of splitting patterns have been reported, ranging from an even distribution of 12 guanines (six in one strand of the probe and six in the other) to a 11:1 split.[107] Further, it was found that the traditional 3:9, 6:6, and 9:3 splits offer greater stability than more uneven splits such as 5:7 or 2:10.[112] These studies compared the performance of these splits against a single, usually otherwise structureless, target. Here, we designed and compared the performance of probes for several targets related to clinically relevant genes with varying secondary structure and G/C content and studied the optimal splitting patterns for each target. Using this information, we constructed a general flowchart that can be used to narrow down the options for splitting the G4-based signal transducer to the most optimal design based on the sequence around the chosen mutation site.

Additionally, we began studies on modifying the loop composition of the G4 to optimize catalytic efficiency of G4-hemin peroxidase-like deoxyribozyme, which enables the colorimetric signal for split-G4 probes. Ultimately, a monomer G4 sequence was identified that provides approximately a 70% increase in absorbance of the oxidized organic indicator ABTS compared to the most common sequence of the G-rich reporter utilized for split-G4s.

Towards a Rational Approach to Design Split G-Quadruplex Probes

Hybridization probes use synthetic oligonucleotide strands that specifically bind to nucleic acids targets through Watson-Crick base pairs. They have been used in PCR, microarrays, and fluorescent *in situ* hybridization, to name a few, for diagnosis of infectious diseases and human genetic disorders. Along with diagnostic applications, understanding of nucleic acid hybridization is important for such technologies as PCR, antisense, siRNA, and CRISPR/Cas. Optimization of hybridization probes has been a subject of extensive research, which lead to establishing the nearest neighbor hybridization parameters for determination of binding affinity and melting temperature of the probe-analyte hybrids.[113] Optimization of the hybridization selectivity of DNA microarray probes has been extensively studied and optimized.[114, 115] Soon after introduction of nanostructured molecular beacon probes, a series of studies on the probe optimization has revealed the relationship between probe performance and the length of its stem and loop parts, and, in general, the effect of the degree of the conformational constraint on the probe's selectivity.[116-118] These studies established the guidelines for the molecular beacon probe design, which contributed to both widespread use of the probes, as well as our general understanding of hybridization thermodynamics.[119] Recent developments of molecular-beacon-based split (or multicomponent) probes demonstrated how some of the probe limitations8 can be overcome.[26, 105, 120-122] The multicomponent probe approach promises to add an important advantage to the existing hybridization tools by enabling recognition of structured nucleic acids with high affinity and selectivity.[22, 123] Systematic investigation on the design of multicomponent probes can facilitate their practical applications.

One of the advantages of split hybridization probes is their ability to reliably discriminate single nucleotide substitutions (SNS) in nucleic acid targets under ambient conditions.[22, 26, 120, 123, 124] Analysis of SNS patterns in human genes is relevant to the diagnostics of genetic diseases,[125] determining people's predisposition to diseases,[126] or response to therapy,[127] and can assist in studying the evolution of human populations,[128] as well as in forensic applications.[129] In addition, SNSs can serve as important markers for genetic studies of human disease vectors,[130] and for prediction of pathogen drug-resistant phenotypes.[3] The potential of hybridization probes for SNS analysis has been widely explored in genotype-based drug susceptibility testing of pathogenic bacteria, which is considered a faster alternative to routinely used phenotypic tools.[131] Most molecular approaches for drug susceptibility testing rely on fluorescent labels covalently attached to the probes to reveal the presence of a genetic signature of interest. In addition to the high cost of fluorescent probes, such tests require instrumentation to read the signal, which limits their use to laboratory settings. The use of color change as a signal readout instead of fluorescence would make molecular drug susceptibility testing more affordable and suitable for point-of-care settings (e.g. clinics or doctor's office), since in this case the signal can be read by the naked eyes without the need for instrumentation.[132] This would make molecular drug susceptibility testing more affordable to the patients, which would allow for more efficient and timely treatment and prevention of antibiotic misuse.

Optical sensors based on peroxidase-like deoxyribozymes (PDz) have been wildly explored for the detection of various analytes including nucleic acids.[76-78, 102, 103, 106, 108, 133-142] The color is generated due to the catalytic activity of G-quadruplex (G4) DNA,[75, 143-145] which uses hemin as a cofactor to catalyze peroxidation of colorless organic molecules to produce a colored oxidation product. Previously, several probes utilizing split PDz (sPDz) for colorimetric nucleic acid detection have been
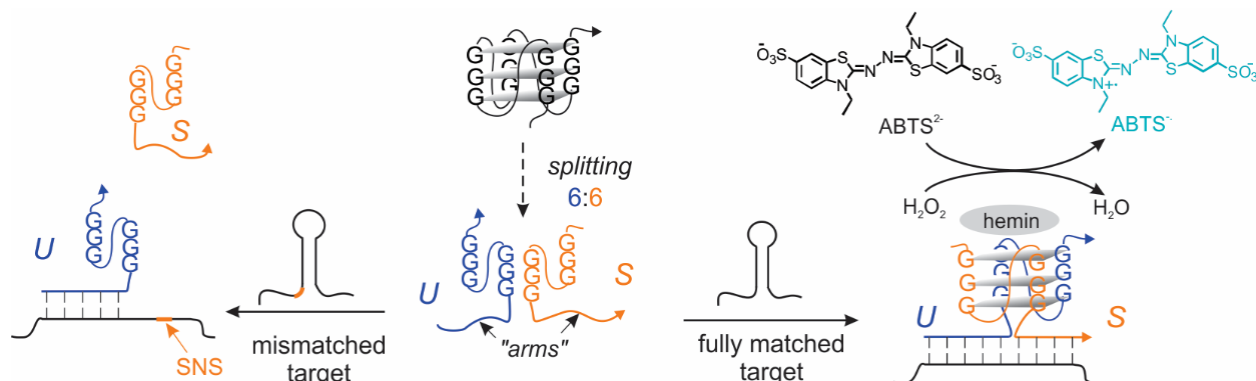
Figure 8: Design of split peroxidase-like deoxyribozyme (sPDz) probe. The G4 core is split into two halves, with the target recognition sequences ("arms") added to each half to constitute strands U and S. The color change is triggered only in the presence of the fully matched target due to the formation of the catalytically active G4-hemin complex promoting peroxidation of colorless substrates (e.g. $ABTS^{2-}$) to colored products ($ABTS^{+\bullet}$).

reported, including those targeting SNS sites.[76, 77, 102, 103, 106, 108, 140-142] However, to the best of our knowledge, nucleic acid detection tests based on PDz probes have yet to reach commercial success. High background signals were previously reported as a major deterrent,[107, 146] and such issues are further investigated in this study.

In the PDz design, twelve guanine residues can be distributed between the probe's strands either symmetrically (each strand contains two G-triads, 6:6)[77, 102, 103, 140-142] (Fig. 8, middle) or asymmetrically (one strand includes a single G-triad, while another contains three triads, 3:9 or 9:3).[76, 106, 108] The two strands are equipped with DNA sequences complementary to adjacent positions on the target ("arms" in Fig. 8). In the presence of the complementary DNA or RNA target, the G4 structure is formed due to the proximity of the two PDz subunits resulted from their hybridization to the target (Fig. 8, right). In the case of mispairing between the target and strand *S* (e.g. in case of SNS) the catalytic G4 structure fails to assemble, and the signal remains low (Fig. 8, left). The differential binding capabilities of the probe components allow for their fine-tuning and ensure maintaining both high

affinity and selectivity – the combination of properties generally unattainable by conventional hybridization probes.[21, 22]

Despite an abundance of data reflecting the high selectivity of the sPDz probes interrogating various targets,[76, 77, 103, 106, 108, 142] systematic investigation on the probe design is still lacking. Previously,[107] the effect of the G4-splitting mode on the ability of the probe to bind a fluorogenic hemin analogue, PPIX, has been investigated. It has been demonstrated that the background caused by symmetric splitting of the G4-forming sequence is higher than that for asymmetric splitting. This conclusion was, however, made based on results with one model target only. At the same time, primary and/or secondary structure of the target may have a significant effect on the probe performance. Unintended interactions of the G4-forming nucleotides with the target, especially if the target is C-rich, may be more pronounced for certain modes of G4 splitting, thus governing the probe design. Indeed, for the reported use of the sPDz probe,[77, 106, 142] thorough optimization of the probe's design and/or addition of the competition probe was required to achieve optimal SNS discrimination, especially for the targets with high GC content. In this study, we demonstrated the effect of the target sequence (the presence of cytosine clusters and type of SNS) on the signal turn-on ratio and SNS discrimination ability of the probe. We also explored a repertoire of strategies to overcome target-associated challenges and highlighted the strategies for the design of efficient sPDz probes for SNS discrimination analysis. The findings are summarized into empirical guidelines for rational design of the sPDz probes.

Results and Discussion

Target selection and probe design. Generally, design of a split hybridization probe for discrimination of SNS is straightforward: the sequence of a signal reporter is divided into two halves, which are elongated with the target-complementary fragments to ensure formation of the signal-

generating structure in the probe-target complex.[22] The sequences of the signal reporter fragments are kept unchanged for all the targets to be interrogated, with only the target-interacting fragments tailored to the target's sequence. In case of the sPDz probes, there is some degree of freedom in the sequence of signal-generating fragments: the twelve guanine residues of the G4-forming sequence can be distributed equally to both strands of the probe, or one of the probes may have more G-triplets than another. This G-rich sequence can abruptly interact with the target to disturb the probe's performance, especially if a target is G- and/or C-rich, which is often the case for targets of bacterial origin.[147] Varying the number of G-triplets in the probe's strands can provide a simple solution for the improved performance of the sPDz probes. Here, we focused on the targets with two or more consecutive C residues in the fragments complementary to one or both strands of the sPDz probe.

Four practically significant model nucleic acid targets were chosen (App. C: Table 7). Three target sequences (T1-T3) corresponded to the gene fragments of mycobacterial species (*M. tuberculosis* and *M. abscessus*) with point mutations conferring antibiotic-resistant bacterial phenotypes. Bacterial species belonging to *M. tuberculosis* (*Mtb*) complex are the main causative agent of tuberculosis. Species of *M. abscessus* (*Mabs*) complex can cause pulmonary disease, and skin and soft tissue infections.[148] Targets T1 and T2 corresponded to the fragments of the *katG* and *rpoB* genes of *Mtb*, respectively. Point mutations in an 81-bp fragment of the *rpoB* gene are responsible for 95% cases of resistance to a first line anti-tuberculosis drug, rifampin.[149] Point-mutations in the *katG* gene contribute to resistance to another first line antibiotic – isoniazid.[150] T1 contained the "wild type" sequence of codon 315 of the *katG* gene (a genotype causing isoniazid-susceptible phenotype). T1C was a "mutant" corresponding to isoniazid-resistant phenotype with an AGC>ACC SNS.[150] T2C contained a genome fragment from rifampin-susceptible bacteria (with CAC sequence at codon 526), while T2T

represented a sequence with a CAC>TAC substitution (App. C: Table 7).[150] Targets of the T3 group contained a sequence of the *M. abscessus rrl* gene with four possible nucleotides at each of the positions 2058 and 2059 (*E. coli* numbering). Target T3 contained adenine residues at both positions and represented the drug-susceptible genotype. Any nucleotide substitution at each potion renders bacteria resistant to a macrolide antibiotic clarithromycin.[151] Six targets (T3GA, T3TA, T3CA, T3AG, T3AT, or T3AC) corresponding to all possible SNS-containing drug-resistant genotypes were used in our experiments (App. C: Table 7). The last target corresponded to a fragment of segment 7 from the genome of the Influenza A virus (IAV). T4 mimicked H1N1 subtype of IAV, which is one of the most prevalent virus genotypes infecting humans.[151] H3N2, the second most common IAV subtype circulating in humans,[152] was represented by T4T. T4A corresponded to a fragment from H5N1 genome (App. C: Table 7). In this case, differentiation between the SNSs in target T4 would allow IAV subtyping rather than drug susceptibility testing.

The targets were selected based on the presence of three or more consecutive cytosine residues, in or near the site complementary to the arm of strand *S*, as well as the type of SNS (App. C: Table 8). (i) Target T1 had no apparent clusters of cytosine (C) residues, and there was a G>C substitution in T1C resulting in a C-C mismatch in the probe-T1C complex. We anticipated no problems with the probe's selectivity and/or turn-on properties. (ii) Target T2 contained a C-cluster in the middle of the fragment complementary to strand *S*, and the SNS site was within the cluster. Taking into account the G-rich sequence of the sPDz probe, we hypothesized interfering target-probe interactions involving the C-cluster. In addition, a C>T SNS in a mismatched target T2T would cause a G-T wobble base formation between the mismatched target and strand *S*, thus compromising the probe's selectivity. (iii) Target T3 had a C-cluster flanking the fragment interrogated by strand *S*

(shown in grey italics in App. C: Table 7). It could allow an extended hybrid between the target and a G-triplet of strand *U*, affecting signal intensity and/or selectivity. There were two adjacent SNS positions, with adenine changed into all three possible SNSs in each of the two positions of mismatched targets. Along with a challenge of discriminating the SNSs at both position using the same probe, two mismatched targets (T4GA and T4AG) would form G-T wobble base pairs with strand *S*, which would compromise the probe selectivity. (iv) Finally, target T4 also did not contain apparent C-clusters with no problems with the probe performance envisioned. It was mainly used as a "case study" to experimentally verify an algorithm for the probe design suggested based on the findings from the data analysis for T1, T2, and T3.

Four sets of sPDz probes targeting T1, T2, T3, and T4 were designed. The probes interrogated 26-34 nt target fragments and contained the arms of strands S and U fully complementary to their correspondent targets. To ensure SNS discrimination, the target recognition fragment of the SNS-discriminating strand *S* (Fig. 8) was designed to be 8-9 nt long (T_m =19-36 °C). This is long enough to enable the formation of a stable complex between the sPDz probe and the fully complementary target at room temperature. At the same time, if a single nucleotide in the target was mis-paired with the probe, strand *S* would fail to bind to the target, thus preventing the G4 structure from assembling (Fig. 8, left).

The arm of strand *U* was chosen to be longer (17-24 nt, T_m =56-66 °C) to tightly bind to the target and assist in unwinding of its secondary structure (App. C: Fig. 30). For successful design of strand *U*, it should be complementary to a single-stranded target fragment, which serves as a "toe-hold" for strand *U* binding to the target. In all sPDz designs that recognize the same target, the arms were the same (App. C: Table 8-11).

Table 1: Splitting of the G4 core in the sPDz probes[a]

Splitting mode	Number of G-triplets in the signal transducing elements	
	Strand U	**Strand S**
asymmetric, 3:9 ($P^{3:9}$)	One G-triplet (^3U)	Three G-triplets (^9S)
symmetric, 6:6 ($P^{6:6}$)	Two G-triplets (^6U)	Two G-triplets (^6S)
asymmetric, 9:3 ($P^{9:3}$)	Three G-triplets (^9U)	One G-triplet (^3S)

a The notations used for the probes and probe strands are given in parenthesis. Full sequences of the probe strands are listed in App. C: Table 9.

(1) Mode of the G4 core splitting. The sequence of the G4 signal transducer was distributed between strands U and S in three different ways (Table 1): (i) each strand of the probe contained two G-triplets (6:6 split, probe designation $P^{6:6}$); (ii) one G-triplet was placed in strand U, while strand S had three triplets (3:9 split, probe designation $P^{3:9}$); or (iii) there were three and one G-triplets in strand U and strand S, respectively (9:3 split, probe designation $P^{9:3}$). Correspondent strands U and S are designated according to the number of the G-residues they contain indicated in superscript (Table 1). For most of the probes, we explored all three G4-splitting patterns. Depending on the number of consecutive C residues in the target, the probes with the same arms but different splitting of the G4 core differed not only in the background and signal intensity but also in selectivity. Therefore, on a quest for the most optimum sPDz probe, it is worth considering different ways of G4 core splitting, and either symmetric or asymmetric splitting mode can be advantageous depending on the target sequence.

(2) Linkers between the G4-forming and target-recognizing fragments. The most cost-efficient option is to use unmodified oligonucleotides as probe components. In this case, the two elements of each probe strands would be connected via a phosphodiester bond, with an option to include a bridging nucleotide as a spacer. At the same time, a flexible non-nucleotide linker between these

elements can aid to the formation of a stable G4 structure, while having the probe tightly bound to the target.[77] This can result in higher signal intensity, but compromise SNS discrimination ability. Here, we experimentally verified this hypothesis.

(3) Conformational constraint in strand *S*. One strategy to increase the selectivity of a hybridization probe is to introduce a conformational constraint in the form of a stem-loop on one or both strands of the probe.[116, 119] For "signal-on" probes, such a constraint may improve selectivity and reduce the background signal, but may also reduce the signal in the presence of the desired target. For the sPDz probes, the conformational constraint, if not formed intrinsically, can be introduced by adding a C-rich "tail" to the arm(s) of the strand(s). It would help sequester one or more of the G-triads into a stem-loop structure in the absence of the target. In this work, the advantageous effect of the "tail" in strand *S* was experimentally tested.

Performance of the sPDz probes was evaluated based on the color intensity for the samples containing the probe components in the absence of a target (blank samples, background intensity), or in the presence of either a fully matched (specific signal intensity) or SNS-containing (non-specific signal intensity) target. The color intensity was visually monitored and quantified by measuring the absorbance of the samples at 420 nm (one of the maxima for ABTS$^-$. absorbance).[153] The following characteristics of the probes were determined. Signal-to-background ratios (S/B) were calculated by dividing the absorbance for the target-containing sample by the blank absorbance (sPDz in the absence of the target). Higher S/B values predispose the probe to have lower detection limits. In addition, selectivity of the probes was tested by comparing the signal triggered by the fully complementary target with that by an SNS-containing target. A selectivity factor (SF) was calculated using Equation 1:

$$SF = \left(1 - \frac{A_{SNS} - A_0}{A_{wt} - A_0}\right) * 100\%, \tag{1}$$

where A_{wt}, A_{SNS}, and A_0 are values for absorbance at 420 nm for the samples containing the probes in the presence of the fully complementary target, SNS-containing target, and in the absence of the target, respectively. Based on the equation, the selectivity factor can be between 0%, when a probe fails to discriminate SNSs, and 100%, when the SNS-containing target does not trigger the signal above the background. By comparing the absorbance values for the signal of the sPDz probes with visual observation of the color intensity, we empirically derived an S/B≥3 and SF≥90 as criteria which would indicate an excellent probe performance. For clear visual observation of the color change, absolute absorbance of the target-containing and blank samples should be higher than 0.8 and less than 0.3 o.u., respectively. We further referred to the probes with an S/B≥3 and SF≥90 as "well-performing" ones (Tables 2 and 3, values in bold).

Target without apparent cytosine clusters. The sequence of the target T1 representing a fragment of the *Mtb katG* gene does not have more than two consecutive C residues in the probe-interrogated region (App. C: Table 8). Therefore, we expected the simplest design for the PDz probes (without non-nucleotide linkers and conformational constraints in the probe strands) being efficient in interrogating the target and discriminating the G>C substitution. As expected, the P1 probes exhibited excellent selectivity disregarding the G4 core-splitting mode (Figure 9, panel a, Table 2). For the three probes, the signal triggered by SNS-containing T1C was at the background level. The excellent SNS discrimination ability of the P1 probes (with SF of 98-100%) is contributed to by the changes in the secondary structure in the fragment complementary to strand *S* caused by the G>C substitution (App. C: Fig. 30). In T1C, the SNS site is in a stem, which further stabilizes the dissociated

state of the target in comparison with the associated state when the target is in complex with strands U and S.

Even though P1$^{6:6}$ produced the lowest signal, all three probes enabled reliable visual detection of T1 (Figure 9, panel a), with S/B in the range of 4-6 (Table 2). Somewhat lower intensity of the T1-triggered signal for P1$^{6:6}$ and P1$^{9:3}$ relative to P1$^{3:9}$ can be attributed to an extended hybrid between the target and strand ^{6}U1 or ^{9}U1 beyond their arms, which is absent in case of ^{3}U1 (Figure 9, panel b). This partially sequesters the G4-forming nucleotides and likely interferes with T1-strand S interaction, thereby decreasing the target-induced signal. Indeed, annealing of strands U and S with T1 prior to the peroxidation reaction resulted in the same signal intensity for all three probes regardless of the G4 splitting mode (Figure 9, panel c). The same effect was observed in the case of prolonged incubation of strands U and S with T1 prior to the peroxidation reaction (data not shown), as opposed to immediate signal generation (Figure 9, panel a). Our data agrees with similar signals observed for symmetrical and asymmetrical designs of the split G4-based probes reported previously.[107] The extended hybrid formation can be impeded if a triethylene glycol (*teg*) linker is placed between the G4-forming sequences and "arms" of the probes (Figure 9, panel d, top). We hypothesized that the linker would prevent the formation of a T-G wobble base-pair, which otherwise contributes to the extended hybrid stabilization. Thus, it would be easier for strand S to interact with the target complexed with *teg*-containing rather than *teg*-free strand U. Indeed, for all *teg*-containing probes, the T1-induced signal was more intense than for their linker-less counterparts (Figure 9, panels a and d). This agrees with the previously reported improvement of the target-specific signal for the 6:6 sPDz probe upon *teg*-linker incorporation.[77] For the P1$^{3:9}$-*teg* probe, however, improvement of the T1-induced signal was not significant, while the non-specific signal in the presence of T1C increased to the point that the
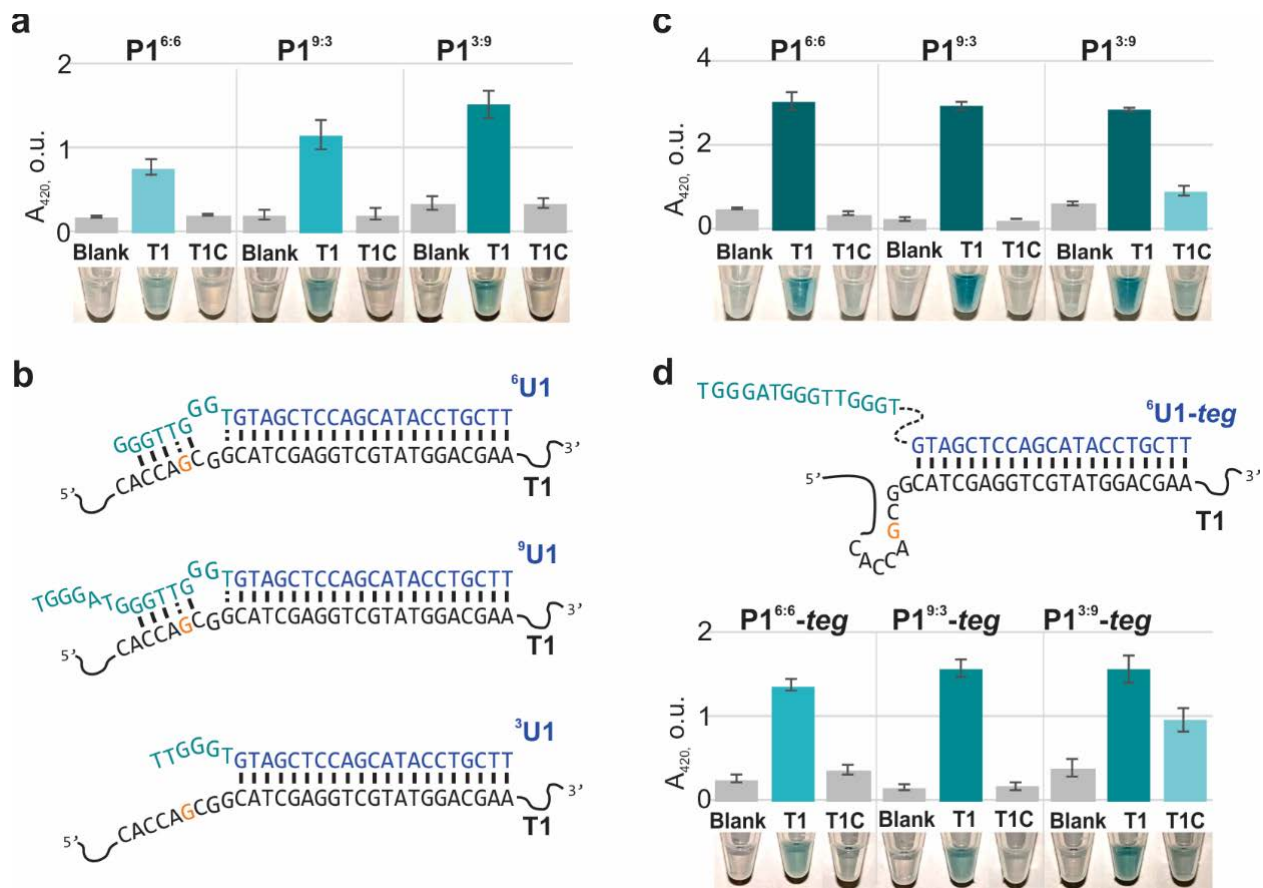
Figure 9: Performance and selectivity of the sPDz probes interrogating a fragment of *Mtb katG* gene. a) Probes P1$^{6:6}$, P1$^{9:3}$ and P1$^{3:9}$ in the absence of the targets, or in the presence of the specific target T1 or SNS-containing T1C. b) Proposed interactions between T1 and either ^{6}U1, ^{9}U1, or ^{3}U1. The structures are drawn as predicted by NUPACK (https://www.nupack.org). The G4-forming nucleotides are shown in green, and the arm of strand *U* is in blue; the SNS position is highlighted in orange. c) Probes P1$^{6:6}$, P1$^{9:3}$ and P1$^{3:9}$ pre-annealed in the absence or presence of T1 or T1C before the signal generation step. d) Complex of T1 with ^{9}U1-*teg* (top), and performance of P1$^{6:6}$-*teg*, P1$^{9:3}$-*teg* and P1$^{3:9}$-*teg* in the absence of presence of T1 or T1C (bottom).

color was clearly observed (Figure 9, panel d), which resulted in about 2-fold drop in the selectivity

factor (Table 3).

Probe P1$^{3:9}$ had higher background than the other two probes. This most likely reflects the

folding of strand ^{9}S into a G4-structure with at least two planes of G-tetrads (App. C: Fig. 31, panel

a), which is supported by the circular dichroism (CD) spectroscopy analysis (App. C: Fig. 31, panel b).

The background can be decreased by introduction of a conformational constraint in the form of a stem-loop (sl) into one or both strands of the probe.[124, 154, 155] Therefore, 4-5-nt C-rich "tails" complementary to a portion of the G4-forming sequences were introduced next to the arm of strands U and S to make probes sl-P1$^{6:6}$, sl-P1$^{3:9}$, and sl-P1$^{9:3}$ (App. C: Table 8). As expected, the background for sl-P1$^{3:9}$ decreased to that same level as was observed for the probes P1$^{6:6}$ and P1$^{9:3}$ (App. C: Fig. 32, compare with the data in Figure 1, panel a). At the same time, the stabilization of the dissociated state of the probe compromises the interaction between the probe and desired target. Indeed, decreased signal was observed for all constrained probes (sl-P1$^{6:6}$, sl-P1$^{3:9}$, sl-P1$^{9:3}$), with sl-P1$^{9:3}$ showing no signal higher than the background (App. C: Fig. 32). The signal could be partially restored if the conformational constraint introduced together with a *teg*-linker between the G4-forming fragments and arms (App. C: Fig. 32, sl-P1u:s-*teg* probes).

In several previously reported sPDz probes,[76, 103, 106, 141] the target-recognizing elements were designed to leave a one-nucleotide gap separating them. Therefore, we tested if this gap affects the performance of P1$^{9:3}$ probe. For this purpose, strand U was "moved" one nucleotide away from strand S, so that the probe-target complex had a C residue in the target free from interaction to either of the strands (App. C: Fig. 33). The gap resulted in a slight drop in the specific target-induced signal, while the high selectivity of SNS differentiation was retained (App. C: Fig. 33, panel c).

In summary, for a targeted nucleic sequence without three or more subsequent C-residues and with an SNS other than C>T or A>G, any mode of the G4 splitting would produce a well-performing *teg*-free sPDz probe. The presence of a cluster of two C residues in the target's fragments complementary to strand S may increase the signal intensity for P$^{3:9}$ both in the absence and presence of the target due to folding of strand ^{9}S into a catalytically active two-tiered G4 stricture. The *teg*-linker

Table 2: Performance of the sPDz probes of the P1 and P2 sets[a]

Probe	P1 set					
	$P1^{6:6}$	$P1^{3:9}$	$P1^{9:3}$	$P1^{6:6}$-*teg*	$P1^{3:9}$-*teg*	$P1^{9:3}$-*teg*
S/B	**4.26±0.05**	**4.5±0.1**	**5.8±0.1**	5.42±0.04	4.0±0.1	**10.47±0.06**
SF[b]	**98**	**100**	**99**	90	51	**98**
Probe	P1 set with a stem-loop					
	sl-$P1^{6:6}$	*sl*-$P1^{3:9}$	*sl*-$P1^{9:3}$	*sl*-$P1^{6:6}$-*teg*	*sl*-$P1^{3:9}$-*teg*	*sl*-$P1^{9:3}$-*teg*
S/B	1.82±0.03	**4.63±0.05**	1.31±0.03	**3.84±0.06**	**6.78±0.08**	3.32±0.04
SF[b]	100	**99**	69	**97**	**91**	95
Probe	P2 set					
	$P2^{6:6}$	$P2^{3:9}$	$P2^{9:3}$	$P2^{6:6}$-*teg*	$P2^{3:9}$-*teg*	$P2^{9:3}$-*teg*
S/B	1.22±0.03	0.9±0.1	1.21±0.02	1.31±0.08	ND	1.92±0.05
SF[b]	65	66	100	30	ND	66
Probe	P2 set with a stem-loop					
	sl-$P2^{6:6}$	*sl*-$P2^{3:9}$	*sl*-$P2^{9:3}$	*sl*-$P2^{6:6}$-*teg*	*sl*-$P2^{3:9}$-*teg*	*sl*-$P2^{9:3}$-*teg*
S/B	1.49±0.04	0.8±0.3	1.28±0.07	**6.95±0.08**	1.1±0.4	3.22±0.08
SF[b]	24	0	16	**73**	0	70

[a] The values for S/B and/or SF indicating poor performance of the probe are in grey. Acceptable probes are bold. "ND" stands for "not determined";

[b] In case of a mathematical value for SF>100%, the SF was made equal to 100%. When the signal triggered by the mismatched target was higher when the signal in the presence of the fully complementary target, the SF was 0.

improves the S/B through background reduction but increases the commercial cost for the probe strands. In the case of T1 analysis, probe $P^{9:3}$ seems to provide the optimal cost-efficiency. Similar designs have been used previously for visual detection and/or discrimination of point-mutations in

nucleic acid targets. In the reported cases, however, probe selectivity down to single mismatch was achieved only due to the addition of competition sequences.[76, 106]

SNS site within a cluster of cytosine residues. In some cases, the presence of several consecutive C residues in the probe-interrogated fragment of the targeted nucleic acid is unavoidable. For example, when an SNS site is within or next to a C-cluster. To demonstrate approaches for the sPDz design in this case, we used a model target T2 with a CCC-sequence and C>T substitution of the last nucleotide of the C-triplet in its SNS-containing target T2T (App. C: Table 7). The SNS site was located next to a stable stem (App. C: Fig. 30), which is generally considered a challenging target for interrogation by hybridization probes.[105, 120] To be able to unwind the target, strand *U* was designed to have a fragment complementary to the stem-loop forming nucleotides of T2, as well as beyond the stem toward the 5'-end of the target (App. C: Fig. 30). In this case, the intramolecular interactions in the arms of strand *U* (App. C: Fig. 34, panel a) may hinder its intermolecular interactions with the target.[120] At the same time, minimum energy secondary structure of strand *U* reveals an intrinsic conformational constraint that can help in reducing the background. To preserve the conformational constraint while mitigating undesirable intramolecular interactions, we introduced a G>A substitution in strand *U* (App. C: Table 9 and Fig. 34, panel b), which would not significantly compromise the strand-target hybridization due to the length of the hybrid. The same substitution was preserved strands *U* of all the probes in the P2 set. Because of the intrinsic complementarity between the arms and G4-forming fragments, even the "open" sPDz probes (designated as P2$^{6:6}$, P2$^{3:9}$, P2$^{9:3}$, P2$^{6:6}$-*teg*, P2$^{3:9}$-*teg*, or P2$^{9:3}$-*teg* in App. C: Table 9) utilized a conformationally constrained strand *U*.
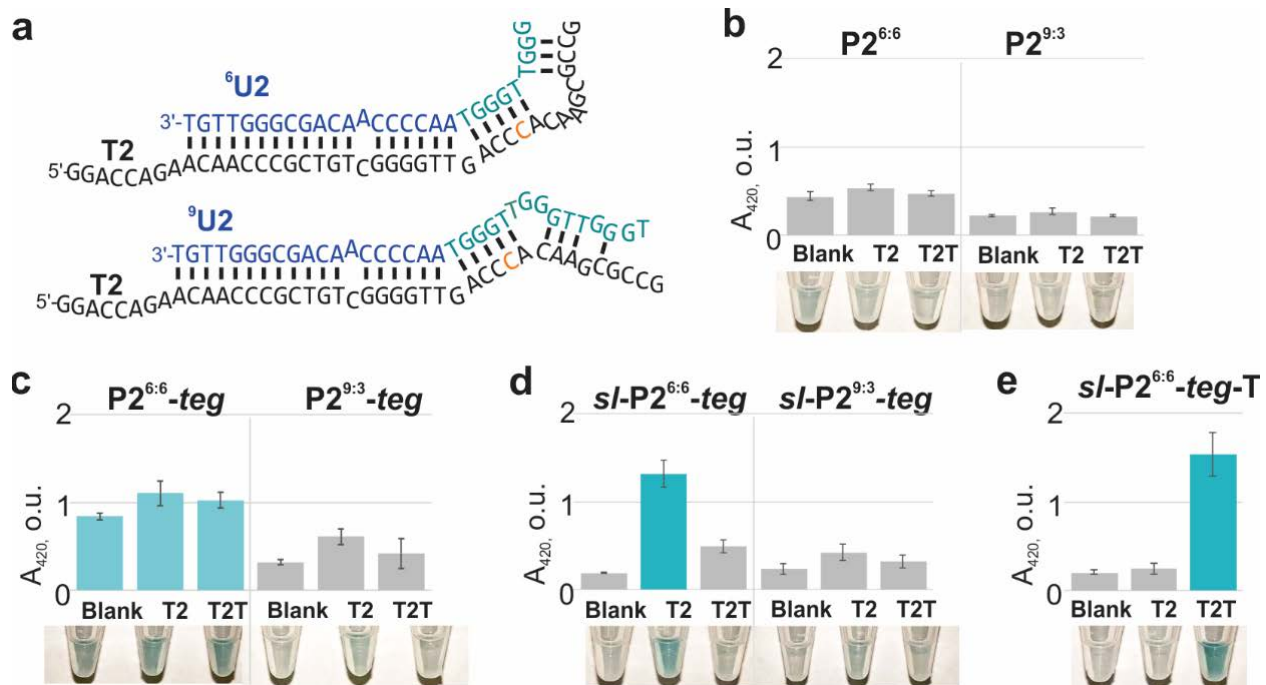
Figure 10: Performance and selectivity of the sPDz probes interrogating a G/C-rich fragment of *Mtb rpoB* gene. a) Proposed interactions between T2 and either ^{6}U2 or ^{9}U2 strands. The structures are drawn as predicted by NUPACK (https://www.nupack.org). The G4-forming nucleotides are shown in green, and the arm nucleotides of strand *U* are in blue; the SNS position is highlighted in orange. b) Probes P2$^{6:6}$ and P2$^{9:3}$ in the absence of the targets ("Blank"), or in the presence of the specific target T2 or SNS-containing T2T. c) Probes P2$^{6:6}$-*teg* and P2$^{9:3}$-*teg* in the absence of the targets, or in the presence of the specific target T2 or SNS-containing T2T.D. Probes *sl*-P2$^{6:6}$-*teg* and *sl*-P2$^{9:3}$-*teg* in the absence of the targets, or in the presence of the specific target T2 or SNS-containing T2T. e) Probe *sl*-P2$^{6:6}$-*teg*-T designed to specifically recognize T2T in the absence of the targets, or in the presence of either T2 or T2T.

The presence of a C-triplet in the fragment of T2 interacting to strand *S* (App. C: Table 7) can cause several issues in the design of the sPDz probe. First, for the P23:9 probe, due to an additional G-triad in the arm (complementary to the CCC-stretch of T2), folding of strand 9S2 into a unimolecular G4 structure is expected, which, in case of parallel G4 topology (App. C: Fig. 35, panel a), would cause high signal of the probe disregarding the target presence. This hypothesis was experimentally confirmed by CD analysis (App. C: Fig. 35, panel b) and in the colorimetric assay for G4 peroxidase-like activity (App. C: Fig. 35, panel c). If "closing" of ^{9}S2 is implemented to over

compete the G4 formation, a very stable stem-loop structure would be required, which may compromise probe interaction with the target. Previously,[156] a stem-loop structure with $\Delta G \sim$ -12 kcal/mol was needed to inhibit a thermodynamically favored unimolecular G4 structure. Second, the presence of the C-triplet in T2 makes the formation of a target-strand *U* extended hybrid highly probable for both $P2^{6:6}$ and $P2^{9:3}$ probes (Figure 10, panel a). This hybrid would mitigate the target-specific signal for the indicated probes, which was experimentally confirmed (Figure 10, panel b). Indeed, no visible difference between the color intensities of the samples containing the probe alone or in the presence of either T2 or T2T target was observed for either probe. As expected, introduction of a *teg*-linker resulted in an increase of the signal intensity for all the samples (blank, specific and non-specific signal) with $P2^{6:6}$ and $P2^{9:3}$ probes (Figure 10, panel c), resulting in the SF of 30 and 66, respectively. For both high intensity of the target-specific signal and acceptable selectivity, simultaneous presence of the non-nucleotide linker and the "closing tail" was required (Figure 10, panel d). In this case, only the $P2^{6:6}$ probe exhibited an appreciable color change in the presence of the specific target (Figure 10, panel d). Due to the presence of a wobble base-pair in the T2T-probe complex, a slight increase in absorbance triggered by T2T was observed, resulting in a SF of 75 (Table 2). Nevertheless, visual differentiation between T2 and T2T was achieved (Figure 10, panel d). In the case of the *sl*-$P2^{6:6}$-T probe specific to T2T, the specific target triggered the signal of S/B=7.6 ± 0.1, while the non-specific signal was at the background level, resulting in a SF of 97 (Figure 10, panel e). Target T2 corresponding to a fragment of the *rpoB* gene from *Mtb* is among the most challenging targets to interrogate with the sPDz probe. Previously,[106] discrimination between SNSs in *rpoB* targets has been achieved only in the presence of the competition probe that would bind to the mismatched target, thus mitigating the non-specific signal.

To conclude, for a target with an SNS within or immediately next to a C-cluster, the sPDz probe with strand S containing three G-triplets will likely fold into a G4 structure, thus generating high background that could not be mitigated by introduction of a stem-loop constraint into the strand. Such catalytically active monomolecular G4 structure can be also formed if G-triplets are brought in proximity as a result of a stem-loop formation in a strand, which has been previously demonstrated.[146, 157, 158] A non-nucleotide linker will be beneficial to prevent formation of a too stable extended hybrid between strand U and the target. For the probe with symmetric G-core splitting, it is less probable that the extended hybrid formation completely prevents the target to bind to both probe strands instead of just strand U, but this would depend on other nucleotides in the sequence of the fragment interrogated by strand S.

Clusters of cytosine residues near the SNS site. In the structure of T3 (App. C: Table 7), there is a long (4 nt) stretch of C residues downstream the SNS sites. It is expected that the stretch is too far from the junction between the strands of the sPDz probe to cause the formation of a monomolecular G4 structure by [9]S3 strand. At the same time, the presence of the stretch can affect the performance of the probes containing strands [6]U3 and [9]U3 with the G4-forming fragments long enough to reach to the C4 stretch on either the fully complementary or SNS-containing target (App. C: Fig. 36, panel a). This extended hybrid is less likely in case of strand [3]U3, so the P3[3:9] probe is expected to have poor selectivity, which was experimentally verified (App. C: Fig. 36, panel b). Between the probes P3[6:6] and P3[9:3], the extended hybrid of T3 with [9]U3 is expected to be slightly more stable than that with [6]U3 due to the presence of additional T and G residues of strand U that can base-pair with hanging C and G residues of the target (though not predicted by NUPACK[111] algorithm). Therefore, the signal (both specific and non-specific) for P3[6:6] should be higher than that for P3[9:3],
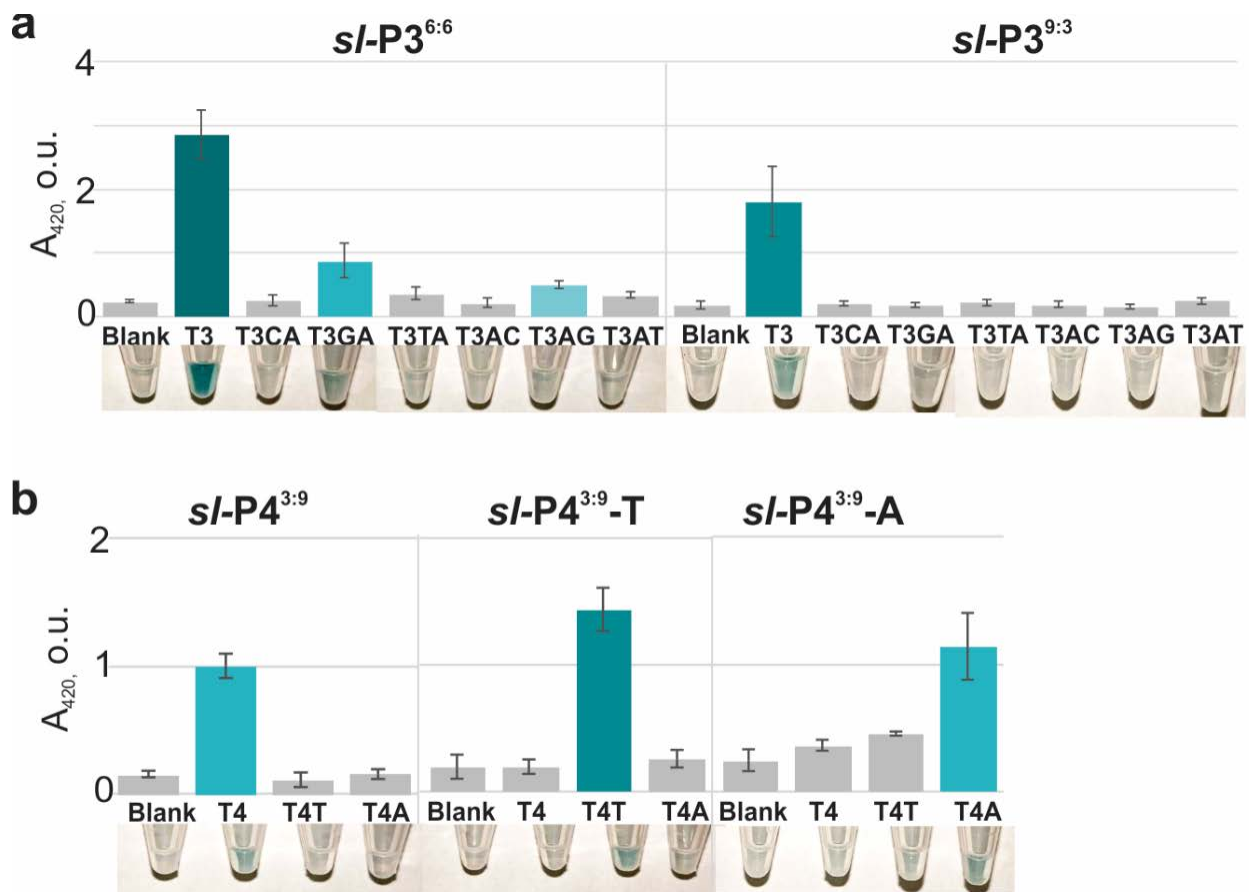
Figure 11: Performance and selectivity of the conformationally constrained sPDz probes interrogating targets T3 and T4. a) Probes *sl*-P3$^{6:6}$ and *sl*-P3$^{9:3}$ in the absence of the targets ("Blank"), or in the presence of the specific target T3 or SNS-containing targets indicated. b) Probes *sl*-P4$^{3:9}$, *sl*-P4$^{3:9}$-T, and *sl*-P4$^{3:9}$-A are designed to be complementary to the targets T4, T4T, and T4A, respectively. The targets represent genomes of IAV subtypes H1N1, H3N2, and H5N1, respectively. Absorbance at 420 nm and tube images are shown for the samples containing the probes in the absence of the targets ("Blank"), or in the presence of the correspondent specific target or one of the SNS-containing targets indicated. Images of the sample tubes are taken immediately before absorbance measurement.

experimentally confirmed (App. C: Fig. 36, panel b). This significantly compromised selectivity of both P3$^{6:6}$ and P3$^{9:3}$. Out of six SNS-containing targets used (three for each of the two SNS positions – nt 2058 and nt 2059), only T3CA and T3AC did not stabilize the catalytic G4-core well enough for visually observed color even for more selective P3$^{9:3}$ (App. C: Fig. 36, panel b). This can be attributed to even more stable hybrid between strand *U* and T3CA or T3AC than between strand *U* and the fully

complementary T3 (App. C: Fig. 36, panel a). By introducing a 5-nt stem in strand ^{1}S3, it was possible to achieve excellent selectivity and high S/B value simultaneously (Figure 11, panel a, Table 3). For the conformationally constrained probe sl-P3$^{6:6}$, though, even the presence of the "tail" to stabilize the probe's dissociated state did not help discriminate an A>G mismatch (Figure 11, panel a), due to the wobble-base in the S-target complex.

Characterization of the most optimal designs of probe sets P1 and P2 was performed in terms of the analysis of the signal intensity dependence on the target concentration, calculation of the limit of detection (LOD) and limit of quantification (LOQ), as well as linear dynamic range (App. C: Fig. 37, panels a-d). The signal intensity increases with the concentration of the probe-specific target and then slightly declines after it exceeds the concentration of the probe. This phenomenon can be explained by the increase in abundance of the catalytically inactive complexes of the target with only one strand of the probe. The linear dynamic range was determined to be 0-500 nM, with LOD and LOQ in the range of 65-75 nM and 195-206 nM, respectively (App. C: Fig. 33, panels b and d). This concentration range is not low enough for the SPDz probe to be able to detect targets in biological fluids without target amplification. Successful use of the sPDz probes in combination with PCR[103, 106] or isothermal amplification[108] has been previously reported.

Empirical Guidelines for the Design of the sPDz probe. Based on our results and literature data, we summarize possible complications in the probe performance depending on the presence of clusters of G- or C-residues in the interrogated target fragment (Figure 12, panels a-c). Previously, DNA duplex-assisted formation of a G4 structure,[157, 158] as well as G4 tolerance to the presence of additional nucleotides and even duplexes in the G4 loop portions,[146] have been demonstrated. In case of split sPDz probes, if a target has a cluster of three or more G-residues in the

Table 3: Performance of the sPDz probes of the P3 set[a]

Probe	S/B	SF[b]					
		SNS at nt2058			SNS at nt2059		
		G	T	C	G	T	C
P3[6:6]	5.8±0.3	26	39	87	50	24	86
P3[3:9]	5.2±0.2	0	0	11	2	3	10
P3[9:3]	7.5±0.2	72	80	100	87	71	100
P3[6:6]-*teg*	5.3±0.1	0	ND	ND	0	ND	ND
P3[3:9]-*teg*	11.81±0.05	3	ND	ND	0	ND	ND
P3[9:3]-*teg*	10.3±0.2	22	ND	ND	24	ND	ND
sl-**P3**[6:6]	12.1±0.2	75	95	100	90	96	100
sl-**P3**[3:9]	11.48±0.08	10	ND	ND	20	ND	ND
sl-**P3**[9:3]	**9.5±0.3**	**100**	**98**	**99**	**100**	**97**	**100**
sl-**P3**[6:6]-*teg*	8.4±0.2	89	ND	ND	95	ND	ND
sl-**P3**[3:9]-*teg*	16.94±0.04	43	ND	ND	65	ND	ND
sl-**P3**[9:3]-*teg*	11.6±0.2	97	ND	ND	100	ND	ND

[a] The values for S/B and/or SF indicating poor performance of the probe are in grey. Acceptable probes are bold. "ND" stands for "not determined";

[b] In case of a mathematical value for SF>100%, the SF was made equal to 100%. When the signal triggered by the mismatched target was higher when the signal in the presence of the fully complementary target, the SF was 0.

fragment complementary to strand *S*, binding of the target to strand *U* can form a duplex with overhang G4 structure (Figure 12, panel a). As a result, even an SNS-containing target would support G4 formation resulting in high signal and lack of SNS discrimination. Targets with clusters of three or more C-residues can also trigger complications in the probe design. For example, if a probe's strand recognizing the C-rich fragment of the target is designed with three G-triplets, it would fold into a

unimolecular G4 structure (Figure 12, panel b), regardless the target's presence. Extremely high background for such probes was observed in this work (App. C: Fig. 35). We also showed that even in case of two consecutive C-residues in the target, the complementary strand would fold into either a two-tiered or incomplete three-tiered G4 (App. C: Fig. 31) with catalytic activity compromising the background (Figure 9, panel a). Finally, formation of an extended hybrid between a C-rich target and strand U beyond its arm is very likely, especially if strand U has two or three G-triplets (strands ^6U or ^9U, respectively) (Figure 12, panel d). This may decrease the signal triggered by both matched and mismatched target.

Taking into account the above-mentioned complications, we suggested an algorithm for the selection of the G4-splitting mode while designing the sPDz probe (Figure 12, panel d). The algorithm takes into account the presence of two or more consecutive C or G residues (C- or G-clusters, respectively) in the probe-interrogated target fragment. Even though in this work we studied the sequence effect for the targets with at least three C–residues, we included the cases of two consecutive C-residues in the algorithms, since folding into a parallel G4 structure with an incomplete third G-tetrad has been demonstrated. For the probe recommendation indicated with an asterisk (*), the algorithm suggests the most optimal G4 split, even though alternative splitting modes can give acceptable performance. For targets with no indicated sequence features, all modes of G4 splitting are acceptable (indicated with double asterisk, ** in Figure 12, panel d).
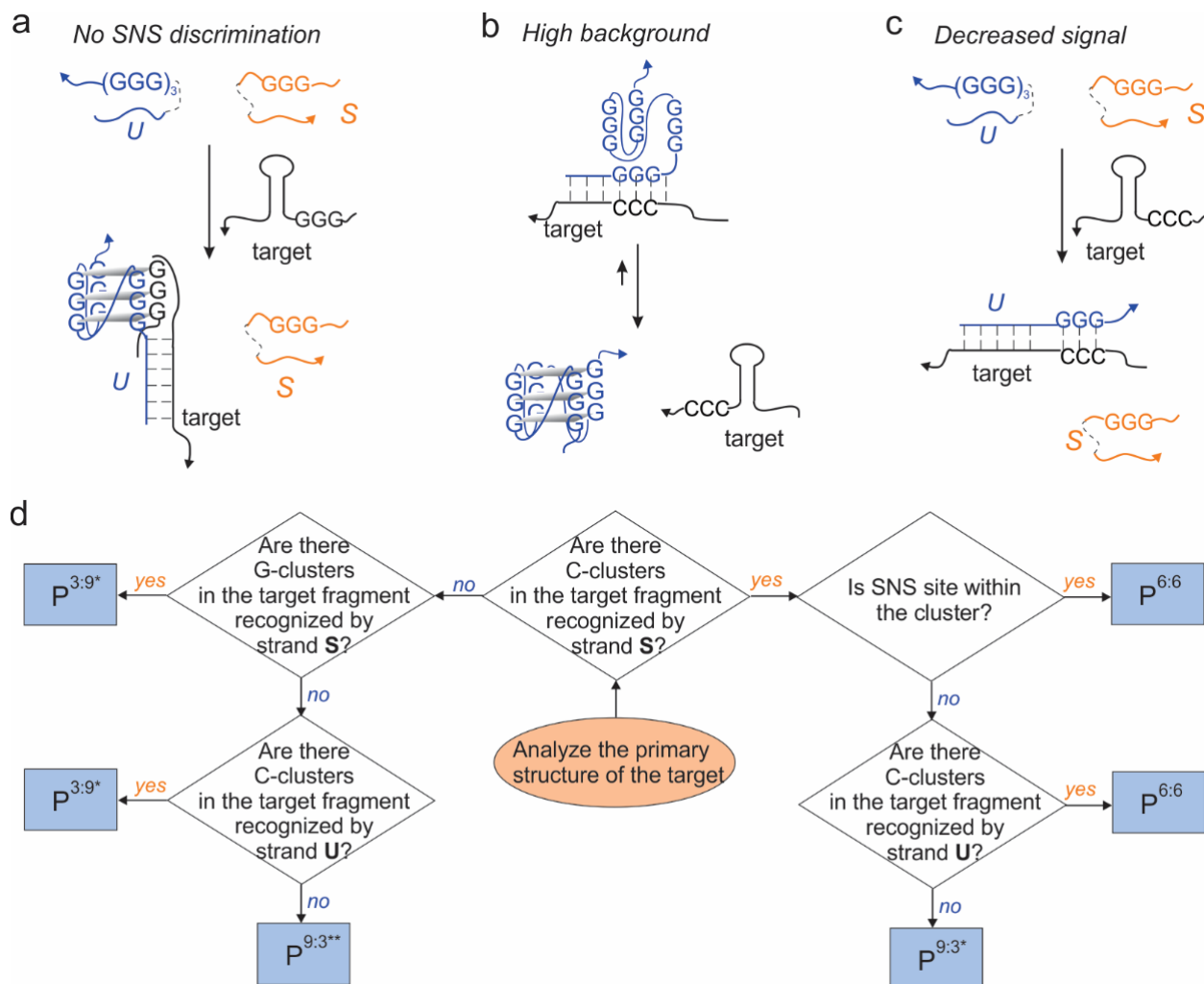
Figure 12: Possible caveats and recommended strategy for the sPDz probe design. a) Formation of the catalytically active G4 structure by target-strand *U* interactions; b) Formation of the catalytically active G4 structure by one of the probe's strands alone; c) Formation of the extended hybrid between the target and strand *U* resulting; d) Proposed algorithm for target-dependent selection of the G4-splitting mode for the probe design. In some cases (indicated by the asterisk,*), sub-optimal P6:6 design could be also acceptable. For targets with no complications (indicated by **), all G4 splitting modes can give acceptable probe performance.

This algorithm was run through the sequences interrogated by the sPDz probes reported in the literature (App. C: Table 12). For most of the sequences, the G4-splitting mode the authors selected is in agreement with one recommended by the algorithm. In one study,[76] a target with C-clusters in target fragments recognized by both strands *U* and *S* was interrogated with the $P^{9:3}$ probe

instead of recommended $P^{6:6}$ to prevent folding of the strands into full or partial unimolecular G4 structures. Since in the reported case, strand U contained two G-doublets in addition to its G4-forming sequence, only an incomplete G4 structure was possible, which would not significantly compromise the probe's performance. In another study,[106] the $P^{9:3}$ instead of recommended $P^{3:9}$ was used to interrogate a target with a 4-nt G-cluster complementary to strand S. Therefore, DNA duplex-assisted G4 formation in target-strand U complex is expected, leading to low selectivity. To alleviate this problem, the authors used a competition probe complementary to the SNS-containing target and, at the same time, to a portion of the G4-forming sequence of strand U.[106]

Further suggestions and troubleshooting. Through careful selection of the G4 splitting mode, it is possible to design a well-performing sPDz probe with no additional modifications. However, in some cases, the performance of this probe is still far from ideal. To optimize the probe, additional tools exist that can assist in improving the ability of the probe to discriminate SNS and/or increasing the intensity of the target-specific signal. A list of suggested troubleshooting options is as follows.

1. Low signal. Nucleotide or non-nucleotide (e.g. *teg*) linkers between the arms and the G4-forming fragments of the probe may decrease the steric hindrance inherent to multi-stranded complexes, such as the probe-target complex of the sPDz probe. This would facilitate G4 formation, thereby increasing overall signal. However, the loss of such hindrance may reduce selectivity.

2. High background may be reduced by extending probes with C-rich sequences that are not complementary to the analyzed target. These sequences will form stem-loop structures with the G4-forming regions, thus serving as conformational constraints that stabilize the dissociated state.

This would prevent association of G4-forming elements of strands *U* and *S* in the absence of the target.

3. Low selectivity. Constraining the dissociated conformation of the strands due to a stem-loop formation described for the previous problem of high background can be helpful in mitigating low selectivity. The length of the stem in the conformational constraint can act as a fine-tuning nob to maximize signal while maintaining selectivity. If the target features a G-rich region, such conformational constraint may be intrinsic to the target-binding region, thereby compromising signal intensity. Another tool is to take advantage of an intrinsic extended hybrid between the target and strand *U* to optimize the ability of the probe to discriminate SNSs.

"Case study" to design an efficient sPDz probe for an arbitrary sequence was performed to check if the above guidelines are sufficient to design an optimal probe for SNS discrimination with minimal optimization of the probe structure. We used an arbitrary target T4, which corresponds to a fragment of IAV genome (subtype H1N1). The 5'-terminal fragment of the target contains several single nucleotide positions that are substituted in the genomes of subtypes H3N2 or H5N1 (App. C: Table 8, targets T4T and T4A, respectively). These SNSs can be interrogated for IAV genotyping and subtype identification, which is important for epidemiological purposes.[153] For the probe design, we selected the SNS site closest to a fragment constant between the IAV subtypes (App. C: Table 7, the targeted SNS position is in red). In this case, it is possible to use the same strand *U* (the strand that is not designed to interact with the SNS) for all virus subtypes, and subtype identification would rely on the target interaction with a set of subtype-specific strands S.

The selected target T4 folds into a secondary structure with two hairpins and the probed SNS site in the middle of a single-stranded region (App. C: Fig. 30). To ensure unwinding of the secondary

structure, we designed strand U4 to bind to a loop region and one of the stem strands. The 5'-terminus of the arm of strand U4 is complementary to a portions of a complementary stem strand to increase the probe-target affinity. The arm of strand S was designed complementary to the single-stranded fragment of T4 containing the SNS site.

There are no C- and/or G-clusters in the strand S-complementary fragment of the target. However, there is a cluster of three cytosine residues in a T4 fragment interacting with the 5'-terminal portion of strand U4. The algorithm suggests the use of the $P^{3:9}$ probe as a starting point. We also tested $P^{9:3}$ and $P^{6:6}$ designs to confirm that the algorithm selection was indeed the most optimal design. By analyzing the folding of just the arm of strand U4, it can be predicted that strand $^9U4$ would likely form a monomolecular G4 structure due to the presence of a stem-loop element (App. C: Fig. 38, panel a). This structure itself could trigger the color change, thus causing too high background signal (App. C: Fig. 38, panel b). Both strands $^6S4$ and $^9S4$ would have an intrinsic stem-loop formed between the arms and G4-forming nucleotides of the strands, but the conformational constraint on $^9S4$ is more stable (App. C: Fig. 38, panel c), which would result in a less intensive signal but lower background and higher selectivity for $P4^{3:9}$ relative to $P4^{6:6}$, which was proven experimentally (App. C: Fig. 38, panel d). Therefore, based on the rational considerations according to our summarized guidelines, probe $P4^{3:9}$ was proven to have the most optimal design for a T4-specific sPDz probe. Similar 3:9 G4 splitting-based design was proven efficient for the T4T- specific sPDz probe (App. C: Table 13). To discriminate A>G SNS in case of T4A-specific sPDz probe (Fig. 12B), "closing" of strand S into a stable ($\Delta G= -3.8$ kcal/mol) stem-loop structure (App. C: Table 13) was required.

Conclusions

Our findings indicate that there is no particular design of split G4 peroxidase-like deoxyribozyme probes that ensures optimal probe performance disregarding the target sequence. The most challenging targets to be interrogated are those with clusters of cytosine residues. Varying the distribution of G-triads may prevent excessive binding of the G4-forming fragments to the analyzed target, thus increasing the signal. Introducing non-nucleotide linkers for better flexibility may allow for an increase in signal in a similar fashion. Closing the G4-forming fragments into latent stem-loop structures reduces background and improves selectivity. Modification of the interior-loop sequences of the G4-forming regions may help reduce unwanted binding of one of the probe strands to the target, and therefore acts as a workaround for otherwise challenging targets. An algorithm for the selection of the G4-splitting depending on the target sequence features for optimal signal and SNS discrimination is suggested and supplemented with tips on how to achieve a well-performing probe using additional design elements. This algorithm can be used, in least at the initial step of the probe design, to eliminate the probe variants that would *a priori* cause either high signal or low selectivity. With more data on interrogation of specific targets with the sPDz probes, the algorithm should be further adjusted to strengthen the rational element in the probe design.

Methods

Materials. DNA oligonucleotides were purchased from IDT, Inc. and used without purification. Oligonucleotides were dissolved in DNase-free water to prepare ~100 μM stock solutions. The concentration of the stock solutions was corrected based on the absorbance of the solutions at 260 nm and the extinction coefficients provided by the vendor. ABTS and H_2O_2 were purchased from Sigma-Aldrich.

Testing and characterization of the sPDz probes. Samples containing two probe strands (1 μM) in the absence or presence of correspondent DNA targets (1 μM) were prepared in a colorimetric buffer (50 mM HEPES-NaOH, pH 7.4, 50 mM $MgCl_2$, 20 mM KCl, 120 mM NaCl, 1% DMSO, and 0.03% Triton X-100) at 22 °C. Alternatively, probe strands were annealed to the target by heating at 95 °C for three minutes and slowly cooling to room temperature overnight. Hemin (375 nM), ABTS (1 mM), and H_2O_2 (1 mM) were added either immediately or following an overnight incubation at 22 °C to observe the development of a blue-green colour. Color intensity was quantified by measuring absorbance at 420 nm using a Thermo Scientific NanoDrop OneC UV-Vis Spectrophotometer (Waltham, MA, USA). Tube images were captured using a cell phone camera. For the experiments to determine limit of detection (LOD), limit of quantification (LOQ) and linear dynamic range for the probes, the concentration of DNA targets was varied (0-2 μM).

Circular dichroism spectra were measured with a Jasco J-810 Spectropolarimeter. Oligonucleotides (2 μM) were prepared at room temperature in the colorimetric buffer. Measurements were conducted in a 4 mm-path cell at 22 °C, and data was obtained with a 2 nm slit width from 350 to 235 nm at 1 nm intervals. CD spectra were averaged over five scans.

Further Works on G-quadruplex Peroxidase Optimization

When utilizing a G4 as a signal reporter for nucleic acid targets, the nucleotide sequence of the reporter may have partial complimentary to the target to be interrogated, which would compromise the probe's performance. This necessitates changing the loop nucleotides of the G4 to avoid, or at least destabilize, the intramolecular interactions. For example, the most commonly used TT loops may be changed to either AT or TA. Such changes may affect the catalytic properties of the G4:hemin complex. Here, we explored the effects of loop composition on the peroxidation performance of single-stranded G4.

Peroxidase activity of the G4:hemin complex is optimized when the G4 has undertaken a parallel conformation.[75, 107] Several factors contribute to the preferred folding pattern of a G4-containing sequence, including salt conditions, length of loops between guanine runs, and the sequence of these loops.[62, 65, 159] We compared the peroxidase-like activity of G4 sequences with varying length and nucleotide composition of the G4 loops. All experiments were performed in the same buffer to exclude the effect of salt composition and concentration.

First, sequences with an increasing number of thymine residues per loop were studied (App. C: Table 14). Performance was monitored by evaluating the maximum color change in the solution upon peroxidation of ABTS as measured through absorbance at 420 nm. Performance for G4 with loops of length 1, 2, and 3 nucleotides was found to be comparable, without a clearly favored distribution; however, upon extending to loop length 4, no color change was observed (Fig. 13A). It has been shown that CD spectra of G4s can be used to approximate the folding pattern of the structure, with a negative 265 nm value indicating an antiparallel conformation, while positive 265 nm values indicate parallel structures when paired with a lower 290 nm value and hybrid when paired with
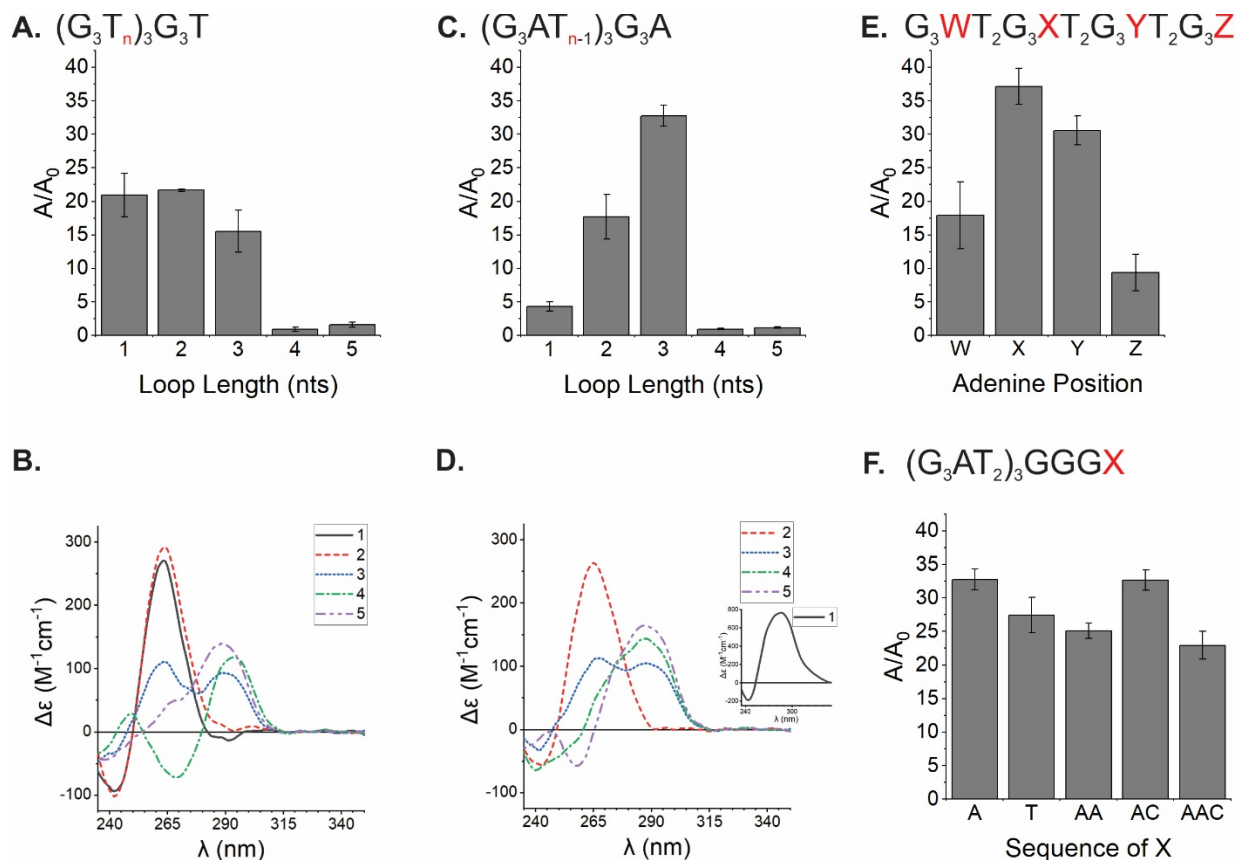
Figure 13: Sequence-based performance of G4:hemin constructs. Signal to Background (A/A_0) ratio is the absorbance of the complex with ABTS divided by the absorbance of ABTS in the presence of hemin alone. All measurements were taken after 15 minutes of incubation at room temperature using the pedestal function of a NanoDrop One[C] (ThermoFisher). Values represent the average of at least three trials. Circular dichroism spectra were collected using a J-810 spectropolarimeter (Jasco). Spectra were collected as the average of ten scans at ambient temperature in the absence of hemin. Spectra shown are after background subtraction and 11-point Savitzky-Golay smoothing. Sequence modifications included varying the length of an oligo-T loop between G-triplets (A, B), varying the length of the loop with the first thymine replaced by an adenine (C, D, inset), varying the location of an adenine in a sequence containing 3-nt loops (E), and varying the terminal nts for the sequence containing ATT loops (F).

a higher 290 nm value.[160] The CD spectra for sequences of loop length 4 and 5 demonstrated a significant reduction in CD at 265 nm and an increase at 290 nm when compared to the shorter loops, indicating an antiparallel conformation for loop length 4 and hybrid for length 5, conformations which have been previously shown to not exhibit strong peroxidation activity with hemin (Fig. 13B).

Previously, it had been suggested that adenine may help coordinate peroxide with the G4:hemin complex, thereby increasing catalytic activity.[161-163] To assess the effect of the adenine position on the peroxidase-like activity of G4, the first thymine in each 1-5nt loop sequence was substituted with an adenine (App. C: Table 14). The 4-5nt loops once more showed a loss of activity compared to the shorter loops (Fig. 13C). The single A loop demonstrated an 80% decrease in intensity compared to the single T variant. The CD spectra for the single A loop (Fig. 13D, inset) does not exhibit a standard quadruplex fold, with a larger CD signal than other variants. However, the similar structure (GGA)$_5$ had been previously shown to fold into a quadruplex.[164] The spectra shown here are in the absence of hemin; it is possible that a tetramer G4 is forming, seeded by hemin; however, further experimentation would be required to verify this hypothesis. The AT loop sequence exhibited nearly unchanged signal compared to TT; the ATT loop sequence, however, demonstrated a 110% increase in signal intensity over the adenine-less 3-nt sequence, and a 51% increase over the highest adenine-less intensity, TT (compare Fig. 13A and Fig. 13C). Interestingly, the CD spectra indicated a more hybrid structure for the 4-nt loop, closer to both the 5-nt loop with and without A; ultimately, this did not lead to an increase in catalytic performance (Fig. 13D).

As the performance of the ATT loop is significantly higher than any other option so far, modifications to this sequence were chosen to study. Single adenine substituted sequences were chosen to see if a single adenine is responsible for the increase in signal over the TTT loop sequence

(Fig. 13E). A T>A substitution at the beginning of the second loop (position X) led to an increase in S/B of from 15.5 ± 3.2 with a thymine to 37.1 ± 2.7 with an adenine; position Y raised A/A_0 to 30.6 ± 2.1. A substitution in the first loop (position W) had no statistically significant effect compared to the TTT sequence while replacing the terminal thymine (position Z) with an adenine led to a 40% loss of signal. It had been previously suggested that the terminal nucleotide(s) of the monomer G4 sequence can play a heavy role in the overall catalytic performance, emphasizing once more that adenine and also cytosine may help with the coordination of the peroxide.[163, 165] The ATT sequence with a terminal AA, AC, or AAC sequence never surpassed the simple A sequence; a terminal T also did not significantly hinder the performance (Fig. 13F).

Overall, it was found that, among all sequences studied, a sequence with three-thymine loops, with the second loop replacing the first thymine with an adenine, offered the greatest overall signal. The increase was substantial, offering a 71% increase over any sequence without adenine. However, this analysis is as-of-yet incomplete. Further testing on utilizing 1-2nt loops with this single adenine inclusion must be done, including variable length loops. Further, these modifications must be tested in a controlled environment for the effect on split probes, as dimeric G4 folding may not directly correlate to monomeric folding. The optimal sequence found here provides immediate uses for any sensing system that would wish to utilize a monomer G4:hemin construct as the output, such as the cascade sensor or ion-detecting sensors.[166, 167]

CHAPTER FOUR: A LIGHT-UP APTAMER FOR SPLIT-PROBE SIGNAL READOUT

Introduction

There is a growing need for low-cost, multiplex DNA-based diagnostic assays. Standard qPCR utilizes fluorophore-labeled molecular beacon or TaqMan probes, which have the advantage of allowing for a variety of fluorophore labels with varying spectral profiles to be incorporated, although the need of covalent conjugation of the oligonucleotide sequence with the fluorophore and quencher moieties and purification of the conjugate increases the overall cost per test. Light-up aptamers interact with fluorogenic dyes non-covalently, and intrinsic fluorescence of the dyes is low enough to eliminate the need for quenchers. As a result, probes utilizing light-up aptamers as signal reporters are cheaper per test, which makes them a promising substitute for the currently used fluorescent hybridization probes, especially in developing areas.

We discovered that the DAP-10-42 aptamer demonstrated a remarkably high proficiency to not only light up its specific fluorogens (dapoxyl sulfonyl dyes) but also other fluorogenic molecules that traditionally light-up in the presence of G4 structures such as NMM, TFT, and CV. As CV demonstrated a light-up capacity much higher than that for the CV-specific aptamer, we set out to test a variety of dyes with similar structure to CV to determine the limits of this capability. As a result, a series of viable dyes have been found that will light-up in the presence of the aptamer. For a given FLAP, it may be possible to tailor the reporter utilized with DAP-10-42 to the spectral characteristics required for usage with another given FLAP, thereby expanding DAP-10-42's utility in multiplex scenarios.

Promiscuous Dye Binding by a Light-Up Aptamer:
Application for Label-Free Multi-Wavelength Biosensing

Aptamers are nucleic acid sequences that bind their cognate targets (small molecules, proteins, whole cells, etc.) through stacking interactions, hydrogen bonding, and/or ionic interactions.[168, 169] With the affinities to their specific substrate characterized by dissociation constants (K_d) in the low micromolar to nanomolar range, aptamers are viewed as more stable alternatives of antibodies in a varieties of applications.[169] For example, fluorescent light-up aptamers, which bind fluorogenic dyes and thereby increase their fluorescence manifold, can be used for tracking RNA in live cells or for label-free sensing biological molecules.[40, 48, 71, 73, 74, 170-174] For the latter application, DNA aptamers are preferable due to their improved chemical stability and lower synthetic cost.

Recently, a DNA aptamer DAP-10-42 that binds dapoxyl dyes and enhances their fluorescence has been reported.[47] It has been fused with the ATP⁻ or thrombin-binding aptamer to construct aptasensors to detect the correspondent targets. In addition, DAP-10 aptameric core has been used for the design of split aptasensor for nucleic acid analysis using dapoxyl sulfonyl ethylenediamine (DSEDA) or dapoxyl sulfonyl fluoride (DSF) as a dye.[74] Being a DNA aptamer, DAP-10-42 offers an advantage of greater chemical stability and lower synthetic cost as compared to RNA-based aptamers. Out of all known light-up DNA aptamers,[175] DAP-10-42 is particularly attractive due to its high affinity to the dapoxyl dyes. We believe that DAP-10-42 is an exceptional candidate for a biosensor design to detect biological molecules, including specific DNA and RNA sequences. Indeed, unlike commonly used dye-conjugated fluorescent probes (e.g. molecular beacon (MB) and TaqMan probes), the split

aptameric probe consists of inexpensive label-free DNA oligonucleotides and, as most split (or binary) probes, exhibits well-documented exceptional selectivity towards signal-base substitutions.[22]

Here, we demonstrate promiscuity of DAP-10-42 interactions with thioflavin T (TFT) and arylmethane dyes, which offers an advantage of using a variety of dyes as a signal reporter with emissions at different wavelengths. Moreover, we applied the discovered promiscuity to design a split sensor for recognition of a *Mycobacterium tuberculosis* gene fragment with a drug-resistance causing point-mutation.

The DAP-10-42 aptamer was originally selected to turn-on fluorescence of dapoxyl dyes – DSEDA and dapoxyl sulfonic acid (DSA). We tested the ability of the aptamer to bind other commonly used fluorogenic dyes‡ that contain dialkylaniline moiety present in dapoxyl (Fig. 14A). TFT, N-methyl mesoporphyrin IX (NMM) and arylmethane dyes are known environmentally sensitive dyes, with low intrinsic fluorescence due to rapid de-excitation in aqueous solutions and enhanced fluorescence emission upon changes in polarity and/or viscosity,[41] which makes them promising fluorogens. Previously, the ability of the malachite green (MG) aptamer to interact with closely related triphenylmethane dyes has been demonstrated.[176] Here, we report that, of the twelve dyes studied, TFT, MG, crystal violet (CV), and auramine O (AO) exhibit sub-micromolar dissociation constants and >100-fold fluorescence turn-on when mixed with DAP-10-42 (Fig. 14B). Fluorescence enhancement of CV by DAP-10-42 was even greater than by a CV-specific aptamer CV30S,[42] which could be attributable to higher affinity of DAP-10-42 to the dye ($K_d = 0.21 \pm 0.07$ μM, App. D: Table 16) than reported for CV30S ($K_d = 0.49$ μM[42]). Greater fluorescence enhancement by DAP-10-42 than by its specific DNA aptamer MG1-3 was also observed for MG.[44] Remarkably, DAP-10-42 triggered >1500-fold maximal turn-on of AO fluorescence (Fig. 14B), despite lower affinity of the
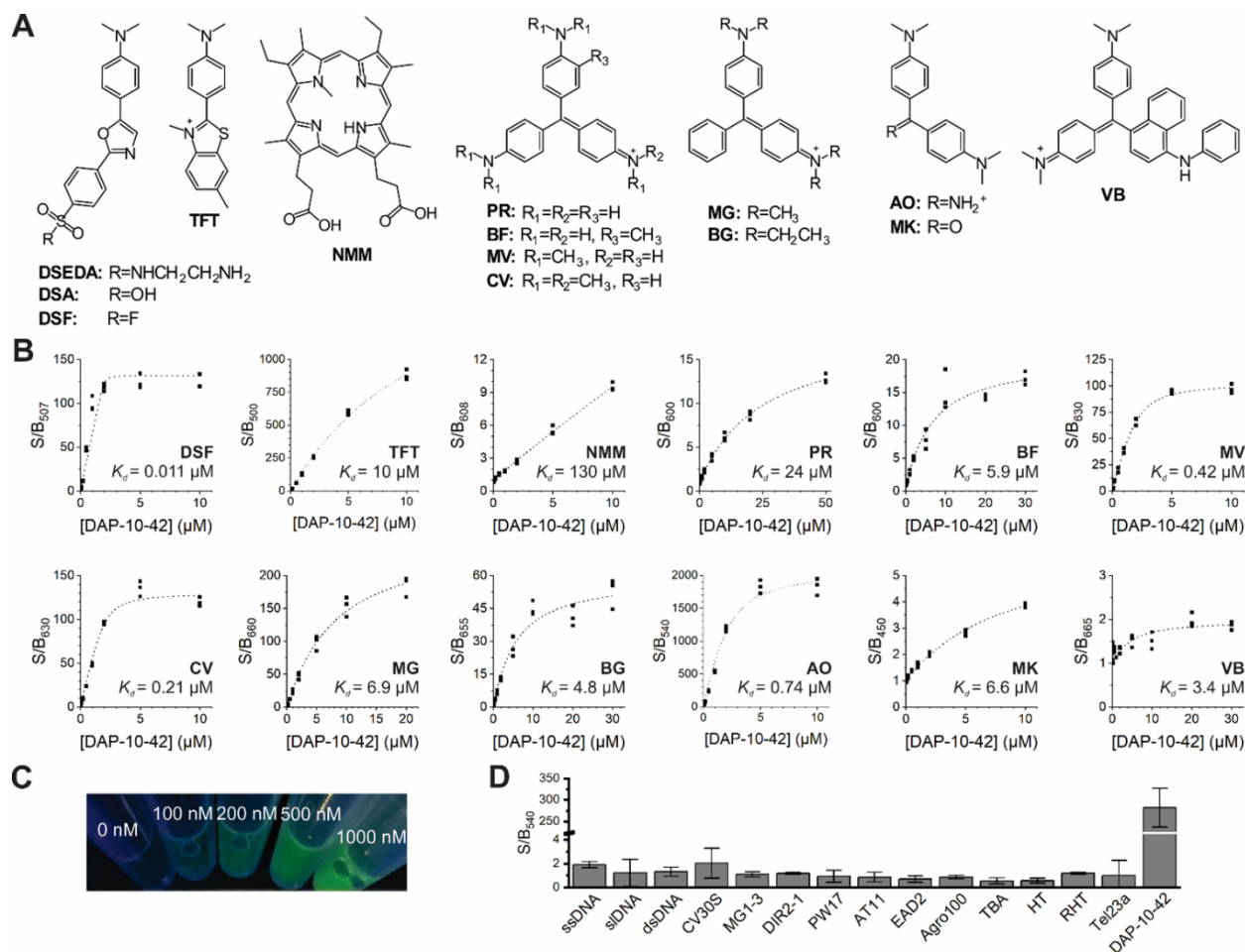
Figure 14: Promiscuous dye binding by DAP-10-42. A. Structures of the dyes tested. B. Fluorescence turn-on (as signal-to-background ratio, S/B) of the indicated dyes by DAP-10-42. C. Fluorescence of AO (2 μM) at the indicated DAP-10-42 concentrations upon excitation with a 365 nm light source. D. Enhancement of AO fluorescence by DAP-10-42 and other oligonucleotides (see Table S1 for sequences). S/B is calculated as fluorescence of the dye in the presence of DAP-10-42 over intrinsic dye's fluorescence. The average values of three independent experiments are shown with error bars as standard deviations.

aptamer to AO ($K_d = 0.7 \pm 0.1$ μM) than for its cognate dye ligand (K_d in the nanomolar range depending on the dapoxyl sulfonyl derivative[47]). The other tested dyes (TFT, MG, BG, BF, PR, MK, VB, and NMM) interacted with DAP-10-42 with lower affinity than AO, as evidenced by >4-times higher K_d (App. D: Table 15). Nevertheless, high fluorescence turn-on was observed for TFT and MG (App. D: Table 15). Victoria Blue B (VB) dye exhibited the lowest fluorescence turn-on by the

64

aptamer, which, along with the emission maximum at a longer wavelength than that for DSF or AO, makes VB an attractive dye target for selection of a VB-specific DNA aptamer. If no cross-reactivity between that aptamer and DAP-10-42 is noticed, the two sequences can be used as aptasensor scaffolds for multiplex biosensing.

AO is an attractive fluorophore for DAP-10-42-based sensors. Unlike dapoxyl dyes, it is an inexpensive commercially available compound. Moreover, AO exhibits negligible background fluorescence, and a quantum yield (φ_F) of 0.00004.[177] Even though the quantum yield of the AO-aptamer complex cannot be precisely determined due to very low absorbance of the dye at the excitation wavelength, it can be said that DAP-10-42 causes at least 12,500-fold φf increase (φf of 0.5 was determined for the dye-aptamer complex, App. D: Fig. 35). For DSA or DSEDA in the absence/presence of DAP-10-42, φf of 0.025/0.866 or 0.00086/0.621, respectively, were previously reported, which corresponds to 35- or 722–fold improvement of the photon emission efficiency.[47] For both DSF and AO, equimolar stoichiometry of binding to the aptamer was observed (App. D: Fig. 36). Fluorescence of AO in the presence of DAP-10-42 is bright enough to be observed with the unaided eye when excited with a broad-spectrum UV lamp (Fig. 14C). Moreover, enhancement of AO fluorescence in the presence of random unstructured or stem-loop folded oligonucleotide, short DNA duplex, other light-up aptamers, or G-quadruplex forming oligonucleotides was negligible (Fig. 14D).

Next, we tested four dyes – DSF, TFT, CV, and AO – as fluorophores for a split dapoxyl aptamer (SDA) probe targeting a *katG* gene fragment from *M. tuberculosis*, point mutations in which have been linked to the isoniazid-resistant phenotype.[178] The probe was designed by splitting DAP-10-42 between nts 28 and 29 (Fig. 15A), which is known to retain the aptamer's affinity to the dapoxyl dyes.[47, 74] The aptamer fragments were connected to the target-binding fragments via a dithymidylate
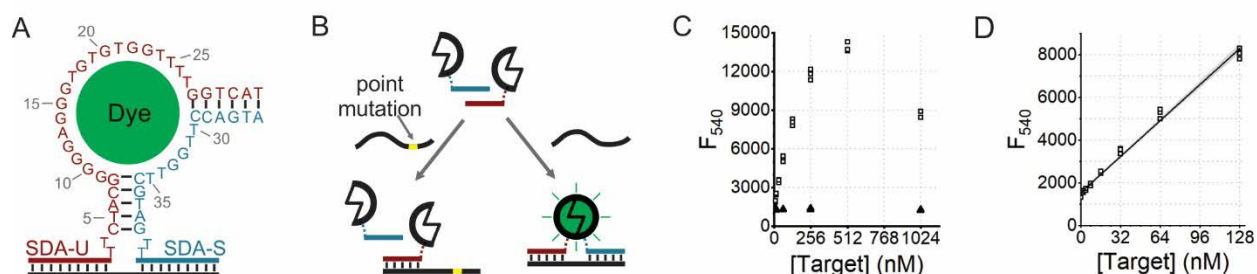
Figure 15: Split dapoxyl aptamer probe. A. Splitting of the DAP-10-42 sequence. Numbers represent nucleotide positions in the intact aptamer. B. Mechanism of the signal generation in response to the fully complementary target (right) and lack of the signal in the presence of a mismatch target (left). C. Response of the *katG*-specific SDA probe to increasing concentrations (0-1024 nM) of *katG* (white squares) or *katG*-G>C (black triangles). D. Linear dynamic range of the probe with AO (2 µM) in response to *katG* (0-128 nM). The line of best fit and 95% confidence interval are shown (black line, gray shading).

linker (to minimize interference between the probe domains) to yield strands SDA-U and SDA-S. The strands were designed to associate in the presence of the fully complementary *katG* target (Fig. 15B, right), but not the mismatched *katG*-G>C (Fig. 15B, left). The target-binding fragment of SDA-S was relatively short (9-nt long) to discriminate the point mutation, while that of SDA-U was 21-nt long to unwind a stable secondary structure of the target (App. D: Table 16, Fig. 37).

Regardless of the dye used, the fluorescent signal increased with concentration of the specific *katG* target in the range of 1-512 nM, while in the presence of the mismatched *katG*-G>C target the signal remained at the no-target level within the concentration range used (Fig. 15C, App. D: Fig. 38A-C). Linear dynamic range was 0-128 nM (Fig. 15D, App. D: Fig. 38D-F). No correlation was found between the affinity of the parent aptamer to the dyes (App. D: Table 16) and the limit of detection (LOD) or limit of quantification (LOQ) of *katG* by the SDA probe (Table 4). However, it was observed that the probe with either TFT or AO, whose fluorescence enhances up to 2000-fold upon saturation with the aptamer, exhibits twofold lower LOD or LOQ than in the case of CV or DSF,

Table 4: Limits of detection and quantifications calculated for the SDA probe targeting *katG* in the presence of various fluorogenic dyes.

Dye	DSF	AO	CV	TFT
LOD, nM	5.4	2.6	4.7	2.8
LOQ, nM	18	8.6	16	9.5

which exhibit more modest fluorescence turn-on (App. D: Table 16). The observed LOD is comparable to that known for the MB probes,[179] which are commonly used in commercial molecular diagnostic assays. Unlike MB probes, the SDA probe offers an advantage of label-free detection. In addition, the probe can be employed at ambient temperatures and still be point-mutation selective.

The LOD values preclude the SDA probe to be used for point-mutation discrimination without amplifying the correspondent mRNA. Here, we used the Nucleic Acid Sequence-Based Amplification (NASBA) method[180] to amplify the targeted fragment of the *katG* mRNA (nts 2868-3001) from a sample of total *M. tuberculosis* RNA using previously optimized primer sequences.[181] Components of the NASBA reaction had little effect on the SDA-AO assay performance, as demonstrated by similar LOD (3.7 nM) and linear dynamic range (0-128 nM) observed when the no-target amplification control (NTC) was spiked with increasing concentrations of the synthetic *katG* target (Fig. 16). Two independent NASBA reactions were performed (App. D: Fig. 39), and concentration of each amplicon (2% v/v) in the assay samples was calculated as 12 ± 4 nM and 19 ± 2 nM (Fig. 16) based on the determined linear dynamic range. When the same amplicons were spiked with the synthetic *katG* target at the final concentration of 32 nM, they triggered a signal of the SDA-AO sensor within the linear dynamic range, which allowed quantification of the spiked samples and yielded 43 ± 5 nM and 50 ± 2 nM, respectively, which corresponds to 2% deviation from the expected value.
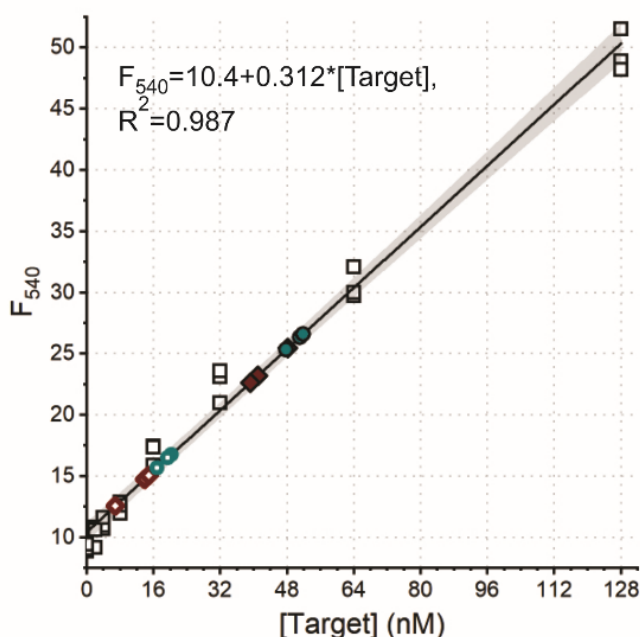
Figure 16: Linear dynamic range of the probe using AO (2 µM) in response to *katG* (0-128 nM) spiked into NASBA NTC sample (2% v/v) performed in triplicate (white squares). The line of best fit and 95% confidence interval are shown (black line, gray shading). Two independent NASBA-*katG* samples were assayed with the SDA sensor at 2% v/v (red rhombus, cyan circle) or 2% v/v spiked with *katG* (32 nM) (filled red rhombus, filled cyan circle).

Therefore, we demonstrated feasibility of the AO-dependent SDA assay for quantitative analysis of

NASBA amplicons without the need of their isolation from the amplification sample.

The reported assay utilizes a label-free probe and thus promises to be cost-efficient in

comparison with conventional dye-modified probes, such as MB or TaqMan probes. A label- free MB

probe has been previously reported.[182] It relies on binding of a fluorophore in place of nitrogenous

bases lacking at the abasic sites of the DNA hairpin, which quenches the dye's fluorescence.[182] While

promising for nucleic acid analysis and genotyping, such MB probe may have reduced shelf life due

to inherent susceptibility to cleavage at the abasic site. More chemically stable examples of label-free

biosensors are based on nucleic acid sequences with a propensity to acquire specific secondary/tertiary

structures, such as i-motifs, G-quadruplexes or DNA triplexes, which exhibit affinities to light-up dyes or nanomaterials.[183-188] In this case, the structure-specific light-up dyes may also bind non-specifically to other similar structures formed with the help of a nucleic acid target or other components of the sample. The use of aptamers minimizes the probability of sequence non-specific dye-nucleic acid binding, thus reducing the background and/or the rate of false-positive responses. In the case of promiscuous dye-aptamer interactions, as it is demonstrated for DAP-10-42 in the present study, the signal wavelength can be selected based on the assay requirements by choosing an appropriate dye from those interacting with the aptamer. Such a property of the aptasensor reported here opens a venue to multiplex applications when the aptameric dye-binding core serves as a signal transducer in combination with another aptamer and/or secondary structure motif interacting with a structure-specific light-up dye.

In conclusion, we have demonstrated that a DNA aptamer DAP-10-42 originally selected to enhance fluorescence of dapoxyl dyes also binds thioflavin T and a few arylmethane dyes. This promiscuity allows to use the aptamer as a scaffold for a multi-wavelength biosensor. A split aptasensor – SDA probe - was designed for label-free nucleic acid detection and point-mutation discrimination. The sensor was shown to respond to an unpurified NASBA amplicon of the *katG* gene from *M. tuberculosis*. The sensor was tested using four dyes of different affinities to the parent aptamer. Excellent selectivity, nanomolar detection limit, and linear dynamic range were observed regardless of the dye used, with no obvious LOD correlation to K_d. However, the LOD seemed to reflect fluorescence turn-on ratios at saturation with the aptamer.

A Logical Array with Alphanumeric Fluorescent Output

Molecular logic gates are molecules, often comprised of nucleic acids, that incorporate computational elements.[83-91] These logic gates take an input (e.g. a nucleic acid input and provide a binary output consisting of either a high-output (digital 1) or a low-output (digital 0).[92] By combining several such molecular logic gates into a single system, it is possible to yield complex analysis of a target while maintaining a binary output. A two-dimensional array of such circuits is capable of providing the creation of alphanumeric characters in response to the presence of specified nucleic acid inputs.[142] When applied to DNA-based diagnostic tests, such a system provides increased reliability owed to the ability to interrogate a target with multiple sensors while also allowing for a simple-to-read output that can be analyzed without instrumental assistance. A logical array for the detection of rifampin-resistance-conferring mutations in the *rpoB* gene of *M. tuberculosis* with alphanumeric and colorimetric output has been reported.[142] Here, a split-DAP-10-42 aptamer has been utilized to create a logical array for the detection of the isoniazid-resistance-conferring mutation in the *katG* gene of *M. tuberculosis* is detailed.

A series of molecular logic gates were designed with fluorescent output owed to the binding of the fluorogenic dye AO to DAP-10-42, utilizing the split system previously described (Fig. 17, App. E: Table 17).[74] Previously, we have utilized the split dapoxyl aptamer to interrogate the *katG* gene of *M. tuberculosis*.[189] There, the short, selective strand of the probe (SDA-S, Fig. 15) consisted of the shorter-half of the split aptamer, allowing the longer strand (SDA-U, Fig. 15) to also host the longer portion of the split aptamer. Combined, these two strands yield a construct with the YES-315S logic.

The probe design for the fluorescent array system took advantage of critical evaluation of the performance of the earlier colorimetric systems, namely, the use of a universal strand across all probes

to minimize the number of unique system components was explored. The greater the number of probes utilized, the greater the complexity of the system; in the case of molecular logic gates utilizing split-transducer units such as split-G4 or split-FLAP units, this can lead to significant increases in the background signal. Split-G4 strands containing 6 or 9 guanine-residues may dimerize to generate an active G4:hemin complex in solution; by adding more sensors to the mix, the probability of dimer formation is increased, and therefore background is increased. As such, asymmetric splitting of the G4 reporter into 9- and 3-guanine containing parts is advantageous; while two 9-guanine strands can readily dimerize into a functional G4, four 3-guanine strands would be required to do so, which is less likely to occur. Further, the possibility of secondary interactions between probes in a complex system increases, which can decrease the maximum signal. While the exact effect of the overlapping interactions can be difficult to predict, minimizing the number of strands utilized can only provide benefits in terms of reducing complexity and cost. As such, for combinatorial logic, it is preferential to have an asymmetrically-split signal transducer that can minimize the utilized concentration of one strand. The split dapoxyl aptamer is an inherently asymmetric signal transducer, with two unique sequences required on abutting ends of the target-recognizing fragments to generate a signal.

Based on these considerations, a system was designed to utilize the long SDA-U strand for all probes (Fig. 17A), intending to prevent probe overload-induced background increases. NOT-R logic may be created via the combination of this YES-S sensor (with high output only in the presence of the "isoniazid-susceptible," INH^S, input sequence) and a NOR sensor (with low output in the presence of either INH^S or INH^R input sequences). To create the NOR3 sensor, a stem-loop-containing construct SDA-N was generated; in the absence of inputs, SDA-N binds SDA-U, yielding a completed aptamer that provides a high output (Fig. 17B, App. D: Table 17). However, the extended toehold on
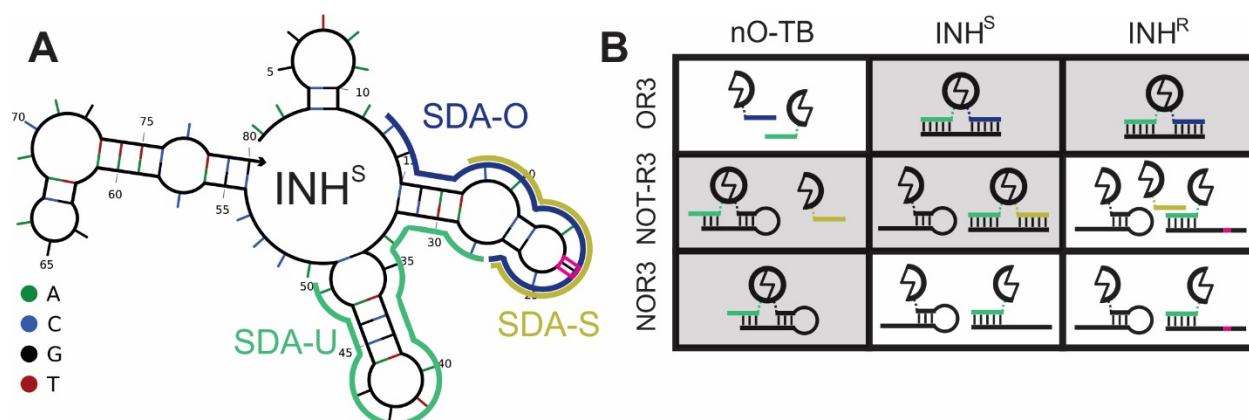
Figure 17: Design of molecular logic gates for utilization in the fluorescent array. A) Secondary structure of INH^s, a synthetic mimic of the gene *katG* from *M. tuberculosis*, as predicted by NUPACK software.[111] The mutation site is highlighted in pink. B) Design of molecular logic gates OR3, NOT-R3, and OR3 and the complexes formed in the absence of TB-related inputs ("nO-TB") or in the presence of the inputs INH^S or INH^R. Colors depicting target recognition sequences match those of SDA-U, SDA-S, and SDA-O in panel A. A grey background indicates a high output signal, whereas a white background indicates a low output signal.

SDA-U allows for the target *katG* (either INH^S or INH^R) to displace SDA-N, yielding a reduction in signal. For NOT-R3, the bound SDA-U then pairs with SDA-S in the presence of the fully-matched target INH^S, returning to a high output (Fig. 17B). Extending SDA-S such that it is no longer selective against the indicated mutation site generates a new probe SDA-O; when paired with SDA-U, this creates the OR3 gate with a high-output in the presence of either the fully-matched or single-mismatched target (Fig. 17). For these three gates, SDA-U, SDA-O, and SDA-S were utilized at 1 µM concentration, while SDA-N was incorporated at 0.5 µM: the NOR3 and NOT-R3 sensors are annealed overnight, allowing for ideal binding conditions that are not available for the turn-on gates, reducing the relative maximum signal of the turn-on gates while administering the assay. As such, a higher concentration of SDA-U, SDA-O, and SDA-S is required to ensure an even signal.

Previously, it was shown that the split dapoxyl aptamer demonstrates activity in the presence of several fluorogenic dyes.[47, 189] The performance of the YES-315S gate was monitored with both
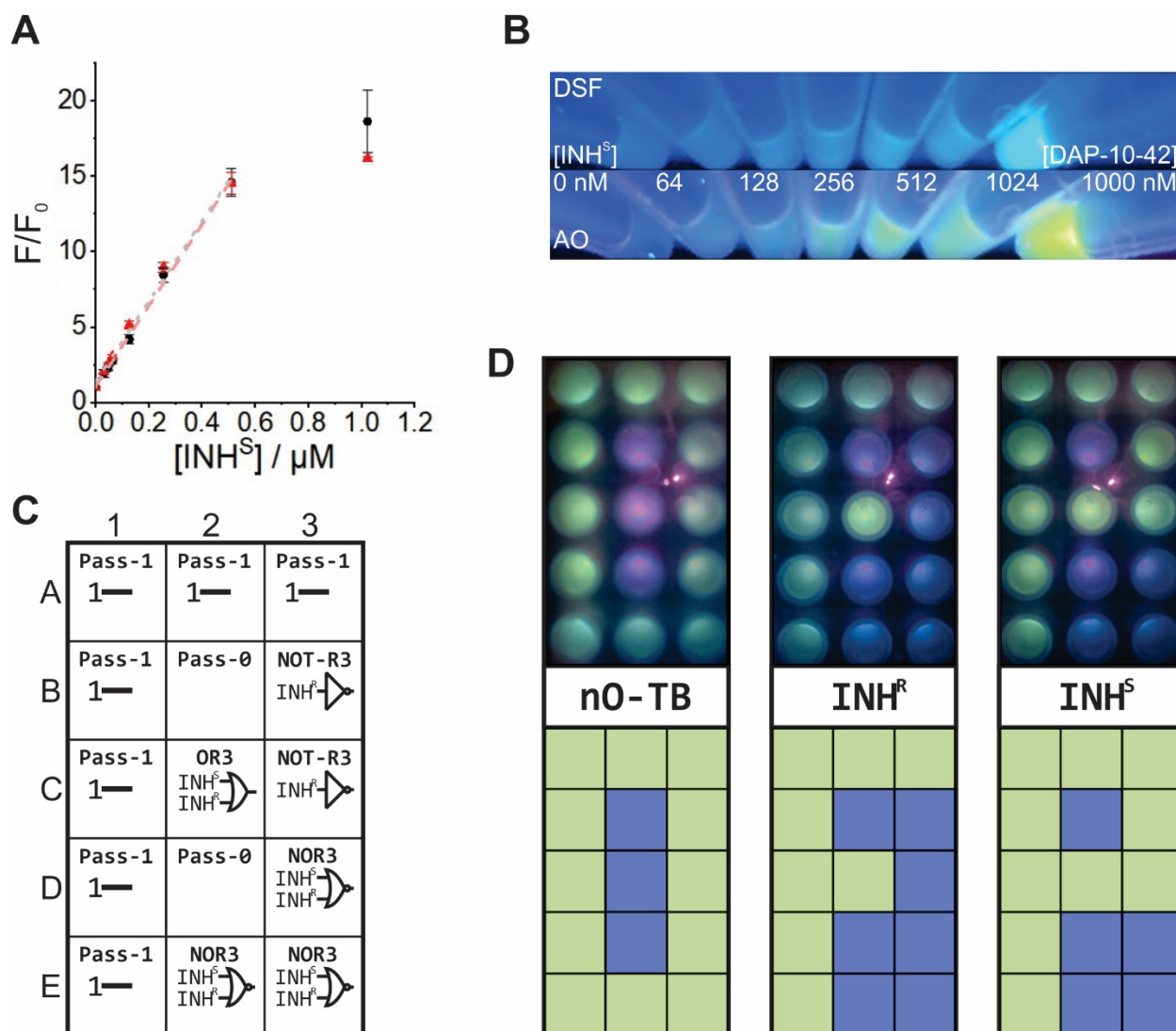
Figure 18: Alphanumeric display of molecular logic gates with fluorescent output. A) Activity of the YES-INH^S sensor with DSF (red) and AO (black) as a function of the S/B on the input concentration, which demonstrates a complete overlap across the linear dynamic range. Dashed lines represent the linear dynamic range of the sensors. DSF measurements were taken with 10/10 slits, AO measurements were taken with 5/5 slits. B) A comparison of DSF and AO fluorescence with YES-INH^S under excitation with a handheld 365 nm UV lamp. C) Layout of the alphanumeric array. D) Expected outputs are displayed below images of the array taken upon excitation with a UV transilluminator. Pictures in panels C and D were taken with a smartphone camera without modification.

DSF and AO, which bind to the aptamer with differing K_ds. (Fig. 18B). When measurements were normalized to signal-to-background ratio, the response across the linear dynamic ranges was found to be independent of the dye utilized (Fig. 18B). This indicates that the assembly of the Target:SDA-U:SDA-S:dye complex is limited by probe hybridization, not dye affinity; for instrumental analysis, there should be only minor changes in the measurement conditions depending on the spectroscopic characteristics used.

For the two fluorogens used, both dye:aptamer complexes may be visualized with a handheld UV lamp (Fig. 18C). Under UV illumination, DSF exhibits a small amount of visible background fluorescence without target present; comparatively, AO shows an almost imperceptible background signal. DSF demonstrates a signal that is readily differentiable from the background at around 256 nM of INH[S]. AO, conversely, exhibits a brightness above the background at 64 nM INH[S] and a green-yellow color of significant intensity at 256 nM. As such, it is apparent that AO is a preferable dye to utilize for detection with a handheld UV lightsource.

The alphanumeric array was assembled in a transparent 96-well plate. The array was designed such that an "O" for "nO TB" would appear in the absence of any variant of *katG,* an "F" for "Fail" would appear if the resistance-causing mutation is present (such as in the presence of the INH[R] input), and a "P" for "Pass" would appear if the fully-matched sequence of the INH[S] input is present (Fig. 18D). Three arrays were assembled; the first without any input added, the second with a final concentration of 1 μM of INH[R], and the third with 1 μM of INH[S] added to each of the 15 wells. After 15-minutes incubation at room temperature, the plate was visualized on a UV-transilluminator (Fig. 18D, top). The low background and bright fluorescence indicated the desired characters and expected behavior (Fig. 18D, bottom) for each array.

More testing is required to fully validate this system, however. While the performance at maximum concentration shown in Fig. 18D indicates a consistent fluorescent readout at saturation, it may be compromised if lower input/target concentration is used. An ideal system will portray consistent fluorescence in all wells at any given concentration that could be found in test conditions. A brighter aptamer is desirable such that a lower concentration of turn-off gates (NOR, NOT-R) may be utilized while maintaining a similarly high signal, allowing for a distinguishable turn-off signal to be present at low concentrations of the target.

CHAPTER FIVE: OPTIMIZATION OF A LIGHT-UP APTAMER VIA NUCLEOTIDE SEQUENCE MODIFICATIONS

Introduction

Aptamers are nucleic acid based "receptors" that are capable of sequence- and/or structure-specific binding of their cognate targets - small molecules, proteins, whole cells.[35, 38, 168] Aptamers promise to advance such traditionally antibody-based applications as biosensing, bioimaging, diagnostics, and therapy.[169, 172, 190] For biosensing and bioimaging, aptamers binding low-fluorescent dyes with enhancement of their fluorescence (fluorescent light-up ap-tamers, FLAPs) can serve as an inexpensive label-free signal transducers. For example, light-up RNA aptamers have been used as label-free reporting systems in living systems to study RNA dynamics[170, 171, 191] or for metabolite and ion detection.[192-195] FLAPs have also found application as signal transducers in *in vitro* analytic assays.[43, 47, 48, 71-74, 173, 175, 189, 196] Success in elucidation of the structure of FLAPs and their complexes with dye ligands stimulated the development of these applications by providing a rational approach for designing the aptamer-based constructs. The use of DNA aptamers instead of their RNA counterparts can contribute to the assay's robustness and affordability due to greater chemical stability and lower synthetic cost of DNA in comparison with RNA oligomers.

An example of DNA aptamers is DAP-10-42, which was originally selected to bind and enhance fluorescence of dapoxyl dyes.[47] We have recently demonstrated that DAP-10-42 acts as a light-up aptamer for other fluorogenic dyes including auramine O (AO) and crystal violet (CV).[189] The aptamer has been explored as a signal transducer for detection of ATP,[47] thrombin,[47] or nucleic acids.[74, 189] A mutational analysis study has been undertaken to explore the source of this promiscuity and discover improvements in the fluorescence capabilities of the FLAP towards each promiscuous target.

Light-Up Optimization Via Sequence Modifications

Aptamers rely on a specific nucleotide sequence in order to acquire the conformation responsible for tight binding of their target.[168, 169] Through SELEX, more stringent sequences that provide tighter binding can be found.[35-38] When selecting for DAP-10-42, two dyes with slightly different structure were used; affinity towards one dye (dapoxyl sulfonyl ethylenediamine, DSEDA) was greater than the other (dapoxyl sulfonic acid, DSA).[47] We hypothesized that the tolerance of the aptamer to both structures is one of the core causes of the promiscuity of the aptamer towards arylmethane dyes that we previously reported.[189] It is possible that by changing the nucleotide sequence in the aptameric core affinity of the modified aptamer to a specific dye could be modified, with or without compromising the affinity to another.

We first analyzed the structure of DAP-10-42 to pinpoint potential mutation sites. Kato et al.[47] proposed that the guanine(G)-rich sequence of DAP-10 may form a G4 structure, which is a common motif for light-up aptamers.[172, 197-200] The predicted minimal secondary structure suggests the duplex/quadruplex fold that is similar to one observed for anti-thrombin DNA aptamers.[201-203]

To confirm the presence of the G4 motif in DAP-10-42, we utilized CD analysis. In G4 structures, each tetrad is formed by four guanines bound to each other via both Watson-Crick and Hoogsteen base-pair interactions; several tetrads stack on each other, thus stabilizing the G4 structure.[204] A variety of topologies are known for G4, which differ in the polarity of the four G-strands, orientation of the loops between the strands, and the number of G-tetrads.[204] The G4 topologies produce characteristic signatures in their CD spectra, which makes CD spectroscopy a technique of choice for studying G4-folding.[160]

CD analysis for DAP-10-42 is complicated by the presence of a duplex between nts 1-8 and 36-42 (Fig. 19A). Double-stranded (ds)DNA is known to exhibit signals in the same region of the CD spectrum as G4 structures.[205] Therefore, subtraction of the duplex contribution from the overall CD spectrum was necessary to reveal a possible G4 folding of the "loop" domain. To ensure associated state of just the stem portion under the analysis conditions (22 °C), the self-complementary fragments of the aptamer were elongated by S1-S8 and S9-S16 nucleotides (DAP-10-58 in Figure 19A). The original DAP-10-42 and extended DAP-10-58 triggered similar fold increase in fluorescence of dapoxyl sulfonyl fluoride (DSF, Fig. 19B), a cognate aptamer ligand (Fig. 19C). Therefore, it was possible to substitute DAP-10-42 with DAP-10-58 in the CD experiments.

A differential CD spectrum (Fig. 19D, solid black curve) was obtained by subtracting the spectrum for the duplex formed by Stem-5' and Stem-3' (App. E: Table 18) from the DAP-10-58 spectrum. It exhibits an intense positive signal with a maximum at 260 nm, a smaller positive peak at ~290 nm, and a negative signal at ~240 nm, which is indicative of a parallel G4 containing propeller loops.[206] Addition of increasing concentrations of DSF results in an increase in the intensity of the positive 290-nm peak with the concomitant decrease at 260 nm, while the negative peak at ~240 nm was preserved (Fig. 19D, see the arrow directions). The CD changes can be explained by the change in the glycosidic bond conformations of some of the stacked guanosine residues from anti-anti in the absence of the dye into syn-anti in the dye-aptamer complex.[206] These changes can be caused by flipping the polarity of the quadruplex-forming G-strands and formation of lateral loop(s), thereby causing a (3+1) hybrid conformation. A hybrid G4 topology has been determined for a human telomeric sequence Tel23a or a fragment of HIV-1 long terminal repeat LTR-III, which exhibit similar CD spectrum as we observed for the G4 motif of the DSF-DAP-10-58 complex (Fig. 19D).[207-209] The

intensity of the CD signal at 290 nm correlated with the fluorescence intensity measured for the aptamer-DSF complex at varied dye concentration (Fig. 19E). Therefore, the dye can act as a chaperon by forcing the aptamer to acquire the hybrid G4 topology allowing for the formation of the dye-binding pocket. This implies that the presence of the G4 is necessary for the function of the aptamer.

A series of mutations were examined with DSF to pinpoint which should be utilized with all dyes for comparison (Fig. 19F, App. E: Table 18). Based on the changes in fluorescence of the dye complexes with aptamer mutants with guanine residues substituted with adenine (A) or hypoxanthine (I) (Fig. 19F), the bases that can be involved in the formation of G-tetrads include guanines at positions 10, 11, 14, 15, 16, 18, 20, 22, 23, 32, and 33. Therefore, most of the substitutions at these residues were not studied for optimizing the aptamer with each promiscuous dye. A loop (composed nts 24-31) encompasses the splitting site on the split aptamer; this site is tolerant to substitutions, deletions, and insertions, and is therefore less likely to affect dye binding.

A series of DAP-10-58 mutations were selected for testing with the dyes that demonstrated promiscuity towards DAP-10-42 (App. E: Table 18). Nucleotides at positions 9, 34, and 35 most likely act as the junction between the G4 and stem; such junctions have previously been shown to be used as binding locations on aptamers.[198, 199, 210] Additionally, mutations at sites 18 and 20 were included; unlike the loop consisting of nts 24-31, the sequence of this aptamer fragment seems to have some effect on its performance. Each sequence was tested using ten dyes that had previously shown fluorescent potential with DAP-10-42: DSF, AO, CV, MG, BG, PR, VB, BF, TFT, and NMM.[189]

TFT and NMM have been previously utilized to detect the presence of G4s with parallel or hybrid conformations.[211] The only modifications at positions 34 and 35 that led to a notable change in fluorescence with TFT were 34X,35X and 34C,35C, yielding an 84% and 65% reduction in
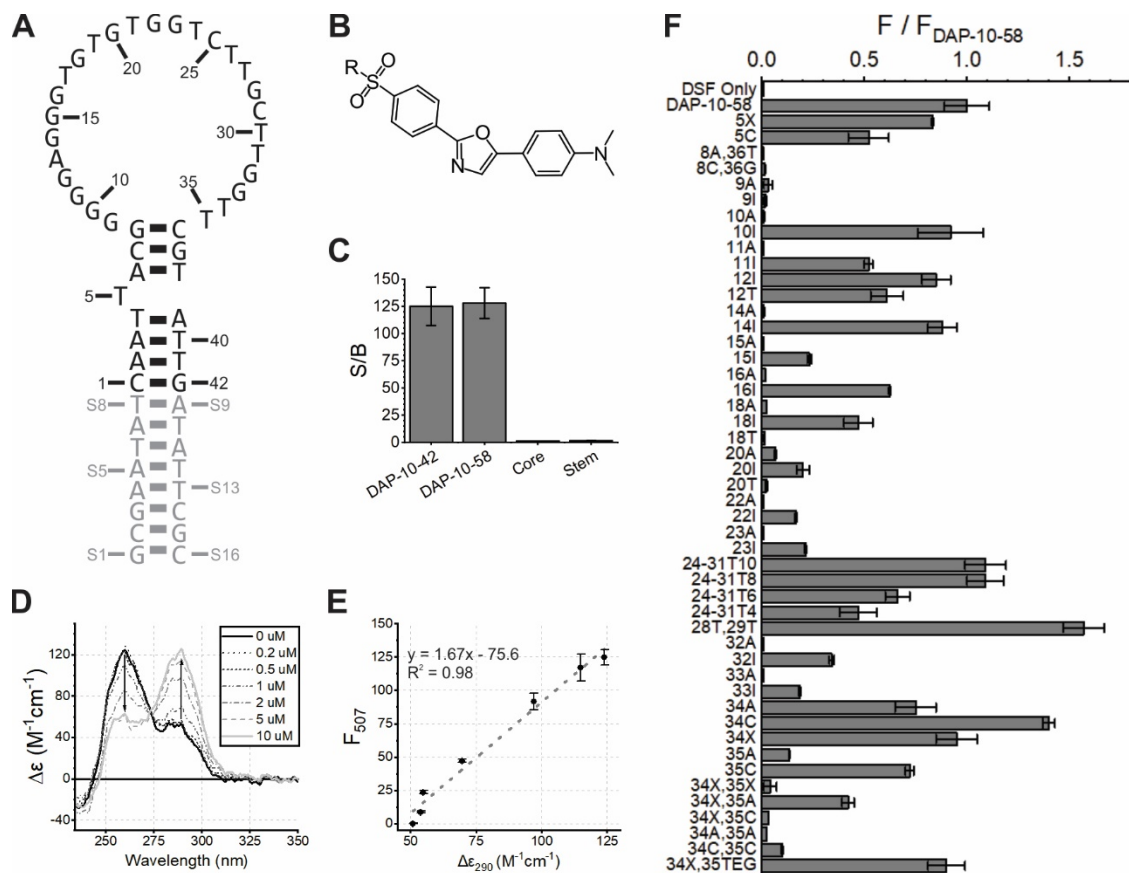
Figure 19: Analysis a potential G4 within DAP-10-58. A) Minimum energy secondary structure of DAP-10-58 as predicted by NUPACK.[111] The nucleotides are numbered according to their positions in the original DAP-10-42, with the added nucleotides labelled in gray (S1-S16). B) DSF structure. C) Fluorescence of DSF (2 μM) in the presence of the indicated sequences (each at 1 μM). The sequences are listed in App. E: Table 18. "Stem" refers to a duplex formed between nts S1-S8,1-9 and 36-42,S9-S16 of DAP-10-58. D) Differential CD spectra for DAP-10-58 in the presence of DSF (0-10 μM) after subtracting the CD spectrum for the stem. E) Correlation of fluorescence intensity at 507 nm for DSF-DAP-10-58 complex with its CD signal at 290 nm. F) Fluorescence intensity of DSF in the presence of the indicated aptamer mutants compared to DAP-10-58 at 507 nm.

fluorescence compared to the unmodified aptamer, respectively (Fig. 20A,B). Further, while 9I yielded

no significant change, 9A led to an 83% loss in fluorescence of mutant-bound DSF when compared

to DAP-10-58. The 9A substitution yielded extreme losses with all dyes except NMM. NMM,

conversely, exhibited its highest signal overall in the presence of the 9A substitution (Fig. 20A-C). A

similar trend can be seen in substitutions at position 20, where a drastic increase in NMM fluorescence

was observed while diminished fluorescence appeared for all other dyes. NMM fluorescence is maximized in the presence of a parallel G4, with reduced fluorescence for hybrid G4 conformations.[212] As such, we hypothesize that these mutations are stabilizing the parallel conformation; however, experimental evidence either in favor or rebuttal of this claim has yet to be obtained.

Several dyes (BF, PR, VB) demonstrated only a slight fluorescence increase, with F/F_0 below 3 in the presence of the unmodified aptamer or any of the mutants (Fig. 20B). Even though fluorescence of the 34X:BF sample was higher than that in the presence of the original aptamer (Fig. 20A), the actual fluorescence turn-on (F/F_0) was only 1.1 ± 0.2 and 1.3 ± 0.2 for BF in the presence of DAP-10-58 and its 34X modification, respectively, rendering them indistinguishable (Fig. 20B). The lowered binding affinity for these dyes previously noted is therefore likely owed to the change in hydrophobicity among the non-methylated aniline moieties of BF and PR when compared to CV, MG, and BG (Fig. 14A).[189] As such, it is unlikely that any single substitution will lead to a notable increase in fluorescence for these dyes.

Most of the remaining mutant:dye pairings yielded a significant drop in fluorescence, although a few interesting cases were present. The 34X,35C variant only yielded a minor (9%) loss in signal for CV, while DSF, AO, and MG yielded 83-98% reductions in fluorescence (Fig. 20A). As of this study, no variant has been discovered that demonstrates both low F/F_0 with CV and high F/F_0 with another dye; such a combination may be useful for multiplex detection.

Mutation 34C exhibited a 41% increase in fluorescence with DSF compared to the unmodified aptamer, the highest recorded fluorescence among mutant:DSF complexes reported in this study (Fig. 20A-B). This mutant demonstrated a loss of signal with AO and CV compared to the unmodified aptamer, although these two dyes still exhibited a 141-fold and 70-fold F/F_0 in the presence of the

dye, respectively. Interestingly, this mutant recorded a 68% loss of signal with NMM. A solution containing 2 μM of each DSF and NMM was tested in the absence and presence of both the 9A and 34C mutants (Fig. 20D). Due to the broad emission spectrum of DSF interfering with the 608 NMM emission maxima, NMM fluorescence was measured at 670 nm, the second highest emission peak, instead of 608 nm used in other NMM-based experiments. While 34C triggers a significantly higher signal at 507 nm than 9A, indicating better activity with DSF, the signal difference at 670 nm is diminished; this may be attributed to the broad emission of DSF, opposed to activity of 9A with DSF.

The deletion at position 34 led to no discernable change in fluorescence for DSF, although an 89% increase in fluorescence was noted for AO, 19% for MG, and 173% for BG (Fig. 20A). While the 34X:BG complex's overall fluorescent turn-on (6.7-fold) is both the highest among studied sequence variations with BG and comparable to some others reported,[42, 44] it is significantly less than the 34X:AO complex's 445-fold fluorescence increase, the highest among all sequence:dye combinations presented here.

A primary goal of this investigation is to determine what sequence:dye combination provides the brightest turn-on when illuminated by a UV lamp. Fluorescence by the highest-turn-on sequence variants for each mutant (Fig. 20B) were compared under illumination from a 365 nm UV handlamp (Fig. 20C). Not all dyes demonstrate significant excitation under UV illumination; MG, BG, BF, PR, and VB all demonstrate little-to-no fluorescence. NMM exhibits little fluorescence, insufficient for the purpose of instrument-free analysis. DSF, TFT, CV, and AO all demonstrate significant fluorescence under 365 nm excitation (Fig. 20C). However, it is unclear if the optimized mutants demonstrate a significant increase in fluorescence observed with the unaided eye under these conditions. The
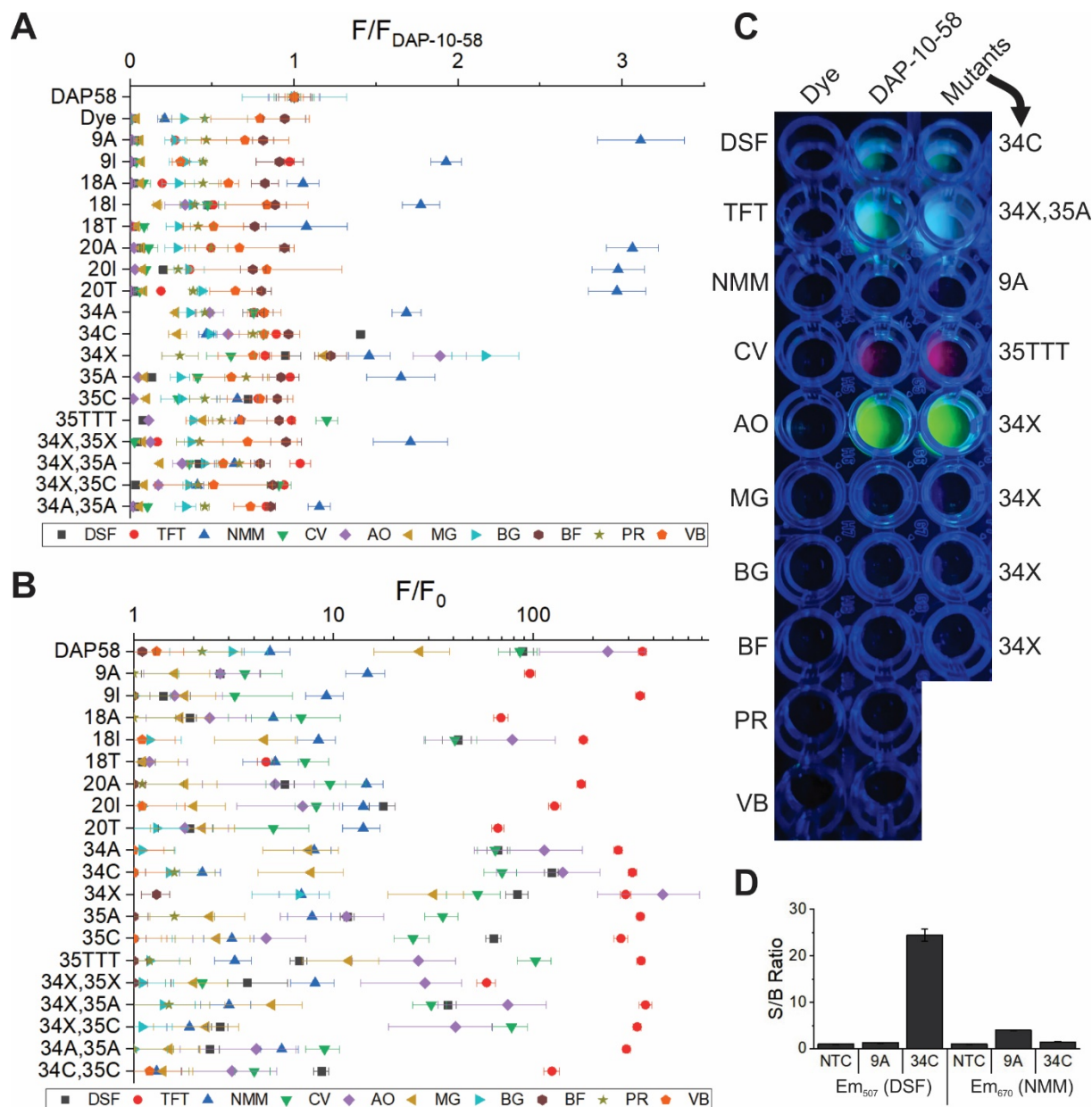
Figure 20: Fluorescent turn-on for the dyes in the presence of DAP-10-58. Mutational analysis of DAP-10-58. A) Relative turn-on values for each dye:mutant complex. B) F/F_0 values for each dye:mutant complex. C) The fluorescence of select mutants when excited with a 365 nm handheld UV lamp. For each dye that exhibited the highest F/F_0 in the presence of a specific DAP-10-58 mutant, that mutant was included in the far-right column. The central column contains DAP-10-58, and the left column is dye only. All DNA is included at 1 μM and each dye at 2 μM, and samples were prepared in the running buffer. Image was captured with a smartphone camera. D) Fluorescence of DSF (at 507 nm) or NMM (at 670 nm) in the absence of DNA or in the presence of either 34C or 9A mutants. The samples contained both dyes (2 μM, each).

expected signal gains were marginal for CV (35TTT: 20% ± 9% gain over DAP-10-58 as measured with a spectrofluorimeter) and insubstantial for TFT (34X,35A: 4% ± 4%). At the same time, mutants containing samples with DSF (34C: 41% ± 9%) and AO (34X: 89% ± 19%) were expected to demonstrate greater brightness (Fig. 20C). The difference in brightness for DSF and AO with their optimized mutants will likely be more observable by the unaided eye at lower aptamer concentrations, where the overall fluorescence is non-saturated.

The 34X:AO, 34C:DSF, 34X,35C:CV, and 9A:NMM complexes represent an improvement in the fluorescent capabilities of the dapoxyl aptamer with each of these fluorophores. The discovery of these higher-performing mutant:dye pairings can provide immediate value for systems with a complete aptamer as a signaling system. However, splitting the aptamer may not yield identical results; the maximum fluorescence of a split DAP-10-42 variant is diminished compared to the monomer aptamer (Fig. 18B). As such, a further study into the optimal sequence for the split dapoxyl aptamer would be required in order to translate these optimizations to utilization in split probes.

CHAPTER SIX: CONCLUSIONS

In this work, we sought to improve upon split hybridization probe technology for the detection of antibiotic-resistance-conferring and species/strain-typing single-nucleotide polymorphisms in nucleic acid targets corresponding to bacterial and viral genomes. To this end, we focused on studying two systems: one that provides a colorimetric output due to the activity of the peroxidase-like G4:hemin complex and another that provides a fluorescent output owed to utilizing a light-up aptamer. For each system, we sought to improve upon the signal generation provided by their respective signal-transducer via sequence optimizations. Further, we sought to explore the viability of label-free split probes for multiplex detection.

In Chapter 2, we sought to explore the utilization of molecular logic gates with colorimetric readout for single-nucleotide polymorphism detection. We created a series of molecular logic gates that exhibit NOR, NOT, OR, PASS-1, PASS-0, and YES logic that were utilized to create a 15-segment display that provides an alphanumeric readout. Interrogating the target sequence through multiple simultaneous probing events provides a level of redundancy to the test, theoretically providing a decrease in the false positive rate when compared to tests without such redundancy. Further, the application of such an assay provides the potential for instrument-free analysis at a clinical level.

In Chapter 3, we sought to determine an optimum split-G4 sequence for split-probes. In our work, we determined that no singular sequence was optimal for all targets. As an example, a 6:6 split of the G4-folding domain was ideal for the C526T mutation of the *M. tuberculosis*'s *rpoB* gene while a 9:3 split was ideal for the G315C mutation in *katG*. We determined that three elements in the target's sequence should be considered when choosing a splitting pattern: the presence of a C-cluster in the

fragment to be interrogated by the SNS-selecting strand of the probe, the presence of G-clusters in this region, and the presence of C-clusters in the region targeted by the non-SNS-selecting strand. We compiled our findings into an algorithm that would dictate what sequences would work best for a given target, then confirmed the algorithm's accuracy using influenza viral strain-typing as a case study. Due to the formation of an unwanted stem loop, the sensor sequence for *rpoB* had to be modified from a TT to a TA loop. Curious on the effects of such modifications, we began a study into loop sequences on monomer G4s to determine the effects on peroxidase activity, where preliminary results indicated that a G4 with loops of TTT, ATT, and TTT provides a significant increase in catalytic activity; however, further testing would be required to determine the best sequence for the dimeric split-G4.

In Chapter 4, we discovered promiscuity in the interactions of DAP-10-42 with fluorogenic dyes. We demonstrated that a series of arylmethane dyes showed increased fluorescence with DAP-10-42, with auramine O in particular demonstrating at least a 12,500-fold increase in quantum yield upon aptamer binding, even though the aptamer exhibited decreased affinity to the dye as compared to dapoxyl sulfonyl dyes. We utilized auramine O with DAP-10-42 to detect the isoniazid-resistance-conferring mutation in *katG* and paired this sensor with the isothermal amplification technique NASBA. The sensors utilizing AO maintained a limit of detection (LOD) of 2.6 nM, compared the 5.4 nM limit of the same system with DSF. The linear dynamic range for the detection of the same nucleic acid analyte using the split aptamer probe was overlapping for the systems with either DSF or AO. Hence, no correlation between the LOD and the K_d for the dye-aptamer complex was observed. We utilized AO with a series of molecular logic gates based on the split dapoxyl aptamer to generate

an alphanumeric display with fluorescent readout for the *katG* fragment that can be seen through simple UV illumination.

In Chapter 5, we sought to further optimize the sequence of DAP-10-42 in terms of its ability to enhance fluorescence of specific fluorogens through targeted single-nucleotide substitutions and/or deletions in the aptamer sequence. In our testing, we discovered sequence alterations that provide an increase in fluorescence of DSF (34C, 141% of DAP-10-58 fluorescence), CV (35TTT, 120%), NMM (9A, 368%), AO (34X, 189%), and MG (34X, 119%). While this provides an immediate avenue to lowered limits of detection with any system utilizing a monolith aptamer, further testing is still required to determine if such a modification is beneficial for a split sensor.

The innovations and improvements to the G4:hemin peroxidase-like deoxyribozyme and DAP-10-42 outlined in this work expand upon the capacity to utilize these systems for signal transduction. Our work towards generating an algorithm to streamline split-G4 sensor design simplifies the design task towards new targets. We have discovered improvements towards the activities of monomer G4 and DAP-10-42-based output systems. The discovery of promiscuously binding dyes with DAP-10-42 offers opportunities for multiwavelength analyte detection and can contribute to the development of multiplex systems with the aptamer due to a broader choice in spectral signatures to pair with another aptamer. The capacity to provide a fluorescent output with simple UV excitation through the DAP-10-42:AO system expands upon the capacity of split-probes for utilization in point-of-care settings where expensive and complicated equipment may be unavailable. Further, the ability to utilize molecular logic gates inside the detection system can provide increased redundancy-based reliability of the test without relying on the use of computer algorithms

for output analysis. We believe that these advancements should provide a greater push towards the utilization of split-probes for single-nucleotide substitution analysis for infectious disease.

APPENDIX A: ARTICLE PERMISSIONS

Alphanumeric Visual Display Made of DNA Logic Gates for Drug Susceptibility Testing of Pathogens

Below is the full text of the agreement for reproduction of this article:

JOHN WILEY AND SONS LICENSE

TERMS AND CONDITIONS

Mar 10, 2022

This Agreement between Ryan Connelly ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number

5143801418017

License date

Sep 07, 2021

Licensed Content Publisher

John Wiley and Sons

Licensed Content Publication

ChemBioChem

Licensed Content Title

Alphanumerical Visual Display Made of DNA Logic Gates for Drug Susceptibility Testing of Pathogens

Licensed Content Author

Yulia V. Gerasimova, Evgeny S. Morozkin, Ryan P. Connelly

Licensed Content Date

Jan 4, 2018

Licensed Content Volume

19

Licensed Content Issue

3

Licensed Content Pages

4

Type of use

Dissertation/Thesis

Requestor type

Author of this Wiley article

Format

Electronic

Portion

Full article

Will you be translating?

No

Title

G-quadruplex reporters: structural studies and application for visual and fluorescent detection of point mutations in nucleic acid sequences

Institution name

University of Central Florida

Expected presentation date

May 2022

Requestor Location

Ryan Connelly

4111 Libra Drive

RM 255

ORLANDO, FL 32816

United States

Attn: Ryan Connelly

Publisher Tax ID

EU826007151

Total

0.00 USD

Terms and Conditions

Signatory Publishers clearing permission under the terms of the STM Permissions Guidelines only, the terms of the license are extended to include subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and does not involve the separate exploitation of the permitted figures or extracts, You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Wiley Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Wiley Materials on a stand-alone basis, or any of the rights granted to you hereunder to any other person.

The Wiley Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc, the Wiley Companies, or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Wiley Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Wiley Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto

NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.

WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

Any fee required for this permission shall be non-refundable after thirty (30) days from receipt by the CCC.

These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY

concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.

WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.

This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

WILEY OPEN ACCESS TERMS AND CONDITIONS

Wiley Publishes Open Access Articles in fully Open Access Journals and in Subscription journals offering Online Open. Although most of the fully Open Access journals publish open access articles under the terms of the Creative Commons Attribution (CC BY) License only, the subscription journals and a few of the Open Access Journals offer a choice of Creative Commons Licenses. The license type is clearly identified on the article.

The Creative Commons Attribution License

The Creative Commons Attribution License (CC-BY) allows users to copy, distribute and transmit an article, adapt the article and make commercial use of the article. The CC-BY license permits commercial and non-

Creative Commons Attribution Non-Commercial License

The Creative Commons Attribution Non-Commercial (CC-BY-NC)License permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.(see below)

Creative Commons Attribution-Non-Commercial-NoDerivs License

The Creative Commons Attribution Non-Commercial-NoDerivs License (CC-BY-NC-ND) permits use, distribution and reproduction in any medium, provided the original work is properly cited, is not used for commercial purposes and no modifications or adaptations are made. (see below)

Use by commercial "for-profit" organizations

Use of Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee.

Further details can be found on Wiley Online Library http://olabout.wiley.com/WileyCDA/Section/id-410895.html

Other Terms and Conditions:

98

v1.10 Last updated September 2015

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

Towards a Rational Approach to Design Split-G-Quadruplex Probes

The American Chemical Society provides innate permission for the reproduction of any works towards one's dissertation or thesis completion, as outlined below:

Toward a Rational Approach to Design Split G-Quadruplex Probes
Author: Ryan P. Connelly, Charles Verduzco, Serena Farnell, et al
Publication: ACS Chemical Biology
Publisher: American Chemical Society
Date: Dec 1, 2019

Copyright © 2019, American Chemical Society

Quick Price Estimate

This service provides permission for reuse only. If you do not have a copy of the portion you are using, you may copy and paste the content and reuse according to the terms of your agreement. Please be advised that obtaining the content you license is a separate transaction not involving RightsLink.

⚑ Permission for this particular request is granted for print and electronic formats, and translations, at no charge. Figures and tables may be modified. Appropriate credit should be given. Please print this page for your records and provide a copy to your publisher. Requests for up to 4 figures require only this record. Five or more figures will generate a printout of additional terms and conditions. Appropriate credit should read: "Reprinted with permission from {COMPLETE REFERENCE CITATION}. Copyright {YEAR} American Chemical Society." Insert appropriate information in place of the capitalized words.

I would like to... ⑦	reuse in a Thesis/Dissertation ▾	Format ⑦	Print ▾
Requestor Type ⑦	Author (original work) ▾	Select your currency	USD - $ ▾
Portion ⑦	make a selection ▾	Quick Price	Click Quick Price

QUICK PRICE   **CONTINUE**

To request permission for a type of use not listed, please contact the publisher directly.

Figure 21: Permission statement from ACS Publications.

Promiscuous Dye Binding by a Light-Up Aptamer: Application for Label-Free
Multiwavelength Biosensing

The Royal Society of Chemistry provides innage permission for the reproduction of any works towards one's dissertation or thesis completion, as outlined below:

If you are **an author contributing to an RSC publication, you do not need to request permission** provided correct acknowledgement is given.

Figure 22: RSC permission statement.

APPENDIX B: SUPPORTING INFORMATION FOR CHAPTER TWO

Experimental Section

Colorimetric assay. The strands of an individual DNA logic gate (1 μM for YES and OR, and 0.5 μM for NOT) were mixed in a running buffer containing 50 mM HEPES, pH 7.4, 50 mM $MgCl_2$, 20 mM KCl, 120 mM NaCl, 0.03% Triton X-100, 1% DMSO. For the NOT^R logic, the mixture contained 0.75 μM YES gate strands and 0.4 μM NOT gate strands. The mixtures were then split into three tubes, 30 μL each, followed by addition of either RIF^R or RIF^S input strand (1 μM) in two of the tubes. Then, hemin (375 nM), ABTS (1 mM) and H_2O_2 (1 mM) were sequentially added to the samples, followed by incubation at room temperature (22 °C) for 10-15 min. The absorbance of the samples at 420 nm was measured using a Nano Drop One^C UV-Vis spectrophotometer (Thermo Fisher Scientific). Alternatively, a picture of the tubes was taken using a smartphone camera. The data from three independent experiments were averaged to calculate the standard deviation from the average. Ten standard deviations were used as a threshold to distinguish high output from low output. The data was processed using Excel.

Time-drive experiments. A sample containing an individual DNA logic gate (1 μM for YES and OR, 0.5 μM for NOT) in 60 μL of the running buffer in the absence of in the presence of either RIF^R or RIF^S input strand (1 μM) was mixed with hemin (375 nM) and ABTS (1 mM). Immediately after the addition of H_2O_2 (1 mM), the sample was placed in a quartz micro cuvette, and absorbance kinetics over 60 min was measured at 420 nm using a Nano Drop One^C UV-Vis spectrophotometer (Thermo Fisher Scientific). The data was processed using Excel.

Visual Display. To assemble the display, four solutions were prepared in 1x running buffer: PASS1 (300 nM PW17), OR (1 µM OR_A, 1 µM OR_B), NOR (500 nM NOR_A, 500 nM NOR_B), and NOT^R (1 µM YES_A, 1 µM YES_B, 500 nM NOR_A, 500 nM NOR_B). Each solution (200 µL) was placed into an appropriate well of a 96-well plate as outlined in Figure 3B. One of the inputs (1 µM), hemin (375 nM), ABTS (1 mM), and H_2O_2 (1 mM) were added to each well, and the samples were mixed by pipetting up and down several times. The images we were captured upon a 10-min incubation at room temperature using a smartphone camera.

Gel analysis. The complexes corresponding to high and low output states were analyzed using 15% PAAG under non-denaturing conditions. The gel was supplemented with KCl (10 mM) and $MgCl_2$ (10 mM) to ensure proper formation of the G4-containing complexes. The samples for the analysis were prepared as described for the colorimetric assay. As controls, samples containing individual gate strands or input strands in the running buffer were used. The gel was run at 80 V for 2 h.

Table 5: Oligonucleotides used in this study.*

Name	Sequence
RIF^S	GGA CCA GAA CAA CCC GCT GTC GGG GTT GAC C<u>CA C</u>AA GCG CCG
RIF^R	GGA CCA GAA CAA CCC GCT GTC GGG GTT GAC C<u>TA C</u>AA GCG CCG
RIF^R-G	GGA CCA GAA CAA CCC GCT GTC GGG GTT GAC C<u>GA C</u>AA GCG CCG
OR-A	*GGG TT GGG* /isp9/ GGT TGT TCT GGT CCa a
OR-B	ccc tac GGT CAA CCC CGA CAG CG /isp9/ *GGG TA GGG*
YES^S-A	*GGG TT GGG* /isp9/ AAC CCC AAC AGC GGG TTG T
YES^S-B	ccc ta TGT GGG TC /isp9/ *GGG TA GGG*
NOR-A	*GGG TT GGG* TTG TTC TGG TCC
NOR-B	CAG AAC AAA CGT AGT ATA TCT ACG T *GGG TA GGG*
PW17	*GGG TA GGG C GGG TT GGG*

     *isp9 – triethylene glycol linker; nucleotides of codon 526 are underlined; single-nucleotide substitutions in RIF^R inputs are shown in cyan; nucleotides forming G4 structures are in italics; nucleotides complementary to portions of G4-forming fragments of strands A and B are in low case.
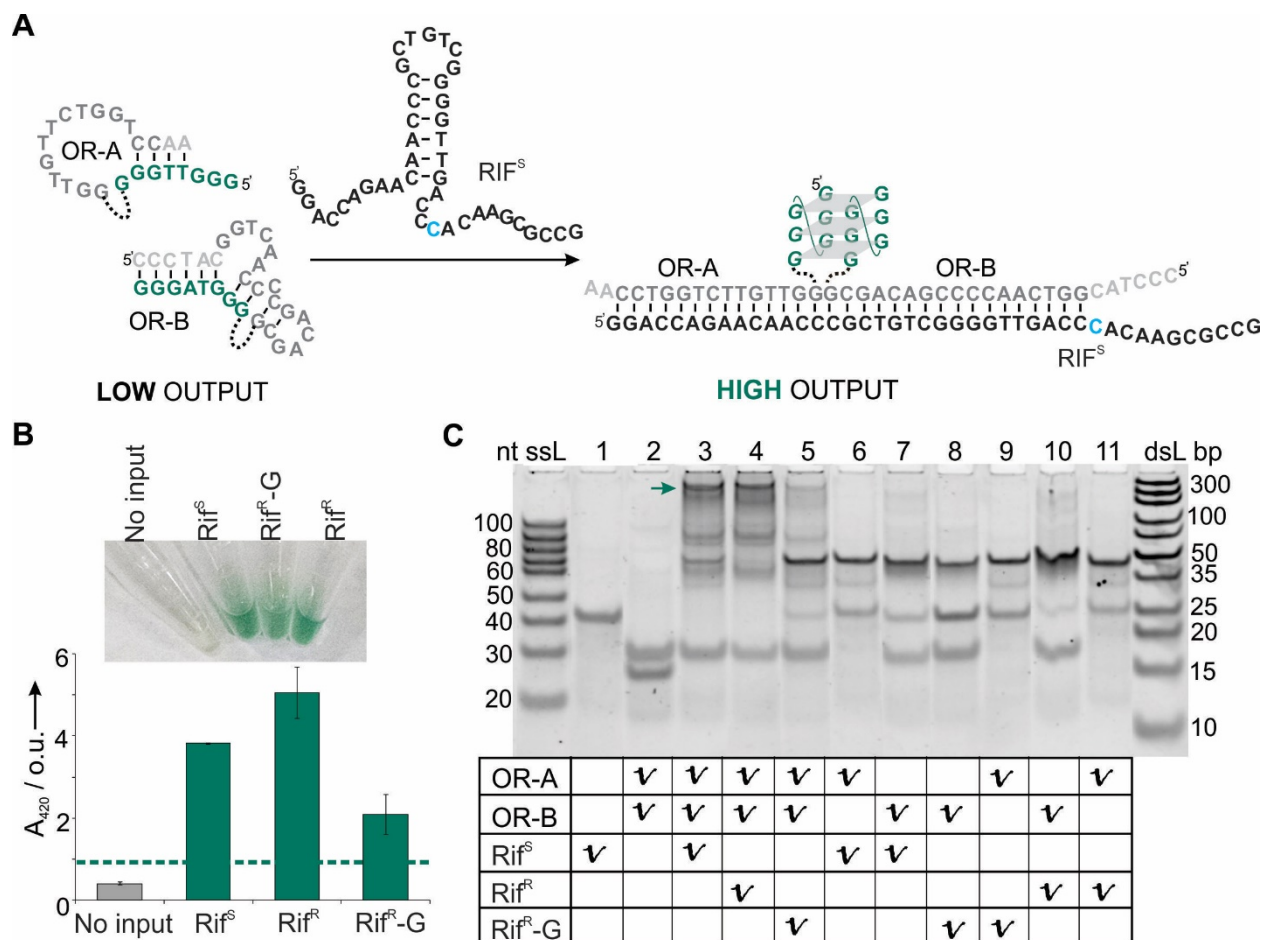
Figure 23: G4-OR gate. A) In the absence of an input, OR-A and OR-B strands acquire stem-loop conformations preventing their re-association into a G4 structure and, therefore, ensuring low background of the "LOW output" (left). Addition of an input (e.g. RIF^S) triggers association of OR-A and OR-B by hybridizing to the input-recognition fragments of the strands (shown in dark grey) and formation of the catalytically active G4 peroxidase responsible for green color of "HIGH output" state (right). The structures are shown as predicted by Mfold[213] or Nupack[111] software. Nucleotides of the gate strands that are not complementary to the input sequence are shown in light grey. Nucleotides involved in the formation of G4 structure are in green. Position of a point mutation in the input is in cyan. Triethylene glycol linkers between the output-generating and input-recognizing modules of the gate strands are shown as black dotted curves. B) Absorbance at 420 nm of the samples containing OR-A and OR-B in the absence ("No input") or presence of either RIF^S, RIF^R or RIF^R-G inputs. Threshold between low output and high output absorbance is represented by green dashed line. The data are average values for three independent measurements with standard deviations as error bars. Inset: Image of the tubes containing the same samples. C) Analysis of "LOW input" and "HIGH input" states of G4-OR gate in non-denaturing 15% PAAG. The lane contents are shown under the gel image. "ssL" and "dsL" designate single-stranded and double-

stranded DNA ladder, respectively. The sizes of the ladder fragments are indicated to the sides of the gel. The band corresponding to OR-A/OR-B complex is indicated by a green arrow.

G4-OR gate consisted of OR-A and OR-B strands, which contained output-generating G-rich sequences at their 5'- and 3'-ends, respectively. The input-recognition sequences at the opposite ends of the strands were complementary to the abutting positions of RIF^S or RIF^R input (nt 1-31) away from the site of a resistance-causing point-mutation (Fig. 23A). Therefore, OR-A and OR-B strands associated in response to the presence of either RIFS or RIFR input (Fig. 23A, right). To prevent self-association of the G-rich regions under low output conditions (no color), OR-A and OR-B were designed to acquire stem-loop conformations upon dissociation (Fig. 23A, left). For this purpose, additional nucleotides, not complementary to the inputs, were added to the input-recognition fragments of the strands (shown in light grey in Fig. 23A). To minimize the interfering effect of a double-stranded complex between the input and input-recognition site of the gate on G4 formation and, thus, catalytic activity, triethylene glycol linkers were introduced between the output-generating and communicating modules (black dotted curves in Fig. 23A).

The logical behavior of the OR gate was tested by measuring the absorbance at 420 nm of the samples containing the gate strands in the absence of inputs, or in the presence of RIF^S, RIF^R or RIF^R-G input (Fig. 23B). A threshold between the high and low output states of the OR gate was placed at $A_{420} = 1$ o.u. The input-triggered absorbance was 2-5 times higher than the threshold, which enabled reliable differentiation between the absence and presence of the inputs. Alternatively, the color development in the presence of an input could be seen by the naked eye (Fig. 23B, inset). As expected, intense blue-green color (high output) was observed only when the input corresponding to either rifampin-sensitive or rifampin-resistant bacterial phenotype was added to the sample.

The samples corresponding to low output and high output were analyzed in native 15% PAAG (Fig. 23C). High output complexes (lanes 3-5, indicated by a green arrow) containing 97 nt overall migrated at ~200 bp position due to the presence of the G4 structure. RIF^R-G input associated with the two strands of OR gate not as efficiently as the two other inputs. This was due to the formation of a 9-nt stem in the structure of RIF^R-G input (Fig. 27B), which rendered the input secondary structure too stable to open, thereby decreasing the amount of G4-forming OR-A/OR-B/RIF^R-G complex. This observation is in accordance with lower absorbance observed for the OR gate in the presence of RIF^R-G, as compared to the one for RIF^S and RIF^R (Fig. 23B, Inset). Either OR-A or OR-B alone could hybridize to the inputs (lanes 6-11), but these binary complexes did not trigger the color change because of the lack of the second portion of the G4 structure and thus inability to form catalytically active G4 peroxidase (data not shown).

G4-YESS gate consisted of YES-A and YES-B strands, which contained output-generating G-rich sequences at their 5'- and 3'-ends, respectively. The input-recognition fragment of YES-B was made of only 9 nt to ensure its binding to a fragment (nt 28-36) of only fully complementary RIF^S, but not to mismatched RIF^R or RIF^R-G (Fig. 24A, right). The input-recognition sequence of YES-A targeted a fragment (nt 9-27) of the input without the point-mutation and thus could bind either RIF^S or RIF^R (Fig. 24A and B). In the dissociated state YES-A and YES-B acquired stem-loop conformations to prevent self-association of the G-rich regions of the strands into the catalytically active G4 structure under low output conditions (Fig. 24A, left). For this purpose, the 5'-terminal part of YES-B was elongated with extra nucleotides, not complementary to the inputs (shown in light grey in Fig. 24A). For YES-A strand, a G>A substitution was introduced in its input-recognizing fragment to ensure stable binding of a portion of the input-recognition fragment to the output-generating

fragment. Without this substitution, different secondary structure was more stable (according to Mfold[213] software), where a stem-loop structure was formed within the input-recognition region of YES-A (Fig. 24C). To minimize the interfering effect of the double-stranded complex between the input and input-recognition site of the gate on G4 formation and, thus, catalytic activity, triethylene glycol linkers were introduced between the output-generating and communicating modules (black dotted curves in Fig. 24A).

The logical behavior of the YES gate was tested by measuring the absorbance at 420 nm of the samples containing the gate strands in the absence of inputs, or in the presence of RIF^S, RIF^R or RIF^R-G input (Fig. 24D). A threshold between the high and low output states was placed at $A_{420} = 1.1$ o.u, which corresponded to the average absorbance of the gate triggered by RIF^R input plus ten standard deviations. Alternatively, the difference in the color of the samples could be observed by the naked eye (Fig. 24D, inset). As expected, intense blue-green color (high output) was generated only in the presence of RIF^S input. The sample containing the gate strands and RIF^R input had slightly higher absorbance at 420 nm than in the absence of inputs or in the presence of RIF^R-G. The color attributed to the formation of a G-T wobble pair in YES-A/YES-B/RIF^R complex, which stabilized the high input complex. At the same time, RIF^S and RIF^R-G inputs could be still unambiguously differentiated by the YES^S gate (Fig. 24D), with about 3-fold difference in the color intensity (as calculated from the absorbance values).

Figure 24: G4-YES gate. A) In the absence of RIF[S] input, YES-A and YES-B strands acquire stem-loop conformations preventing their re-association into a G4 structure and, therefore, ensuring low background of the "LOW output" (left). Addition of the input triggers association of YES-A and YES-B by hybridizing to the input-recognition fragments of the strands (shown in dark grey) and formation of the catalytically active G4 peroxidase responsible for green color of "HIGH output" state (right). The structures are shown as predicted by Mfold[213] or Nupack[111] software. Nucleotides of the gate strands that are not complementary to the input sequence are shown in light grey. Nucleotides involved in the formation of G4 structure are in green. Position of a point mutation in the input is in cyan. Triethylene glycol linkers between the output-generating and input-recognizing modules of the gate strands are shown as black dotted curves. B) Structure of YES-A/RIF[R] complex predicted by Nupack.[111] C) Secondary structure of an alternative YES-A sequence without a G>A substitution, as predicted by Mfold.[213] The substitution position is underlined. D) Absorbance at 420 nm of the samples containing YES-A and YES-B in the absence ("No input") or presence of either RIF[S], RIF[R] or RIF[R]-G inputs. A threshold between low output and high output absorbance is represented by green dashed line. The data are average values for three independent measurements with standard deviations as error bars. Inset: Image of the tubes containing the same samples. E) Analysis of "LOW input" and "HIGH input" states of G4-YES gate in non-denaturing 12% PAAG.

The lane contents are shown under the gel image. "ssL" and "dsL" designate single-stranded and double-stranded DNA ladder, respectively. The sizes of the ladder fragments are indicated to the sides of the gel. The band corresponding to YES-A/YES-B/input complex is shown by a green arrow.

To prove the formation of the complexes corresponding to low input and high input states, gel electrophoresis analysis was performed (Fig. 24E). As expected, the low mobility band of the YES-A/YES-B/input complex was observed only in the presence of RIF^S input (lane 5, indicated by a green arrow; compare lanes 5-7). Either of the three inputs could bind to YES-A strand due to the long input-recognition fragment of the strand (lanes 8-10), but even RIF^S, which had fully complementary sequence to the input-recognition module of YES-B, did not complex with YES-B in the absence of YES-A (lane 11), thus ensuring high selectivity of the input recognition by YES^S gate.

Figure 25: G4-NOR gate. A) In the absence of an input, NOR-A and NOR-B strands associate into a catalytically active structure of G4 peroxidase, thus producing "HIGH output" (left). 3'-Terminal overhang of NOR-A serves as a toehold for an input (e.g. RIF[S]) to displace NOR-A from the NOR-A/NOR-B complex and switch the gate into the "LOW output" state (right). The structure are shown as predicted by Mfold[213] or Nupack[111] software. Nucleotides of the gate strands that are not complementary to the input sequence are shown in light grey. Nucleotides involved in the formation of the G4 structure are in green. Position of a point mutation in the input is in cyan. B) Absorbance at 420 nm of the samples containing NOR-A and NOR-B in the absence ("No input") or presence of either RIF[S], RIF[R] or RIF[R]-G inputs. A threshold between low output and high output absorbance is represented by green dashed line. The data are average values for three independent measurements with standard deviations as error bars. Inset: Image of the tubes containing the same samples. C) Gel analysis of "LOW input" and "HIGH input" states of G4-NOR gate. The lane contents are shown under the gel image. "ssL" and "dsL" designate single-stranded and double-stranded DNA ladder, respectively. The sizes of the ladder fragments are indicated to the sides of the gel. The bands corresponding to NOR-A/NOR-B and NOR-A/input complexes are indicated with green and cyan arrows, respectively.

G4-NOR gate was made of NOR-A and NOR-B strands, which associated into a G4-forming

complex in the absence of either of the inputs due to the complementarity between the 3'-terminal

fragment of NOR-B and part of the input-recognition fragment of NOR-A (Fig. 25A, left). The last 4 nt at the 3'-end of NOR-A served as a "toehold" for either of the three inputs, which displaced NOR-A from NOR-A/NOR-B associate, thereby destroying the catalytically active G4 structure (Fig. 25A, right). For the G4-NOR gate design, the triethylene glycol linkers were omitted due to two reasons: 1) to slightly destabilize the G4 structure in the absence of the inputs for easier turning-off of the gate; 2) to stabilize the NOR-B/input complex, since in the absence of the linkers three additional nucleotides of the input were available for hybridization with NOR-B (see secondary structure of NOR-B in Fig. 25A, right). NOR gate could unambiguously differentiate between the absence (high output) and presence (low output) of the inputs, as could be observed by the naked eye and the difference in the absorption at 420 nm (Fig. 25B). A threshold between the low and high output states of the NOR gate was placed at $A_{420} = 0.7$, which corresponded to the average response of the NOR gate to RIF^R input plus ten standard deviations. The intensity of the high output signal was 4.5-5-fold higher than one for the low output state. Formation of NOR-A/NOR-B complex could be seen in PAAG (Fig. 25C, lane 4, indicated by a green arrow). Addition of either of the three inputs resulted in disappearance of the correspondent band accompanied with the appearance of the band corresponding to NOR-A/input complex (Fig. 25C, lanes 5-7, indicated by a cyan arrow, compare with lane 8).
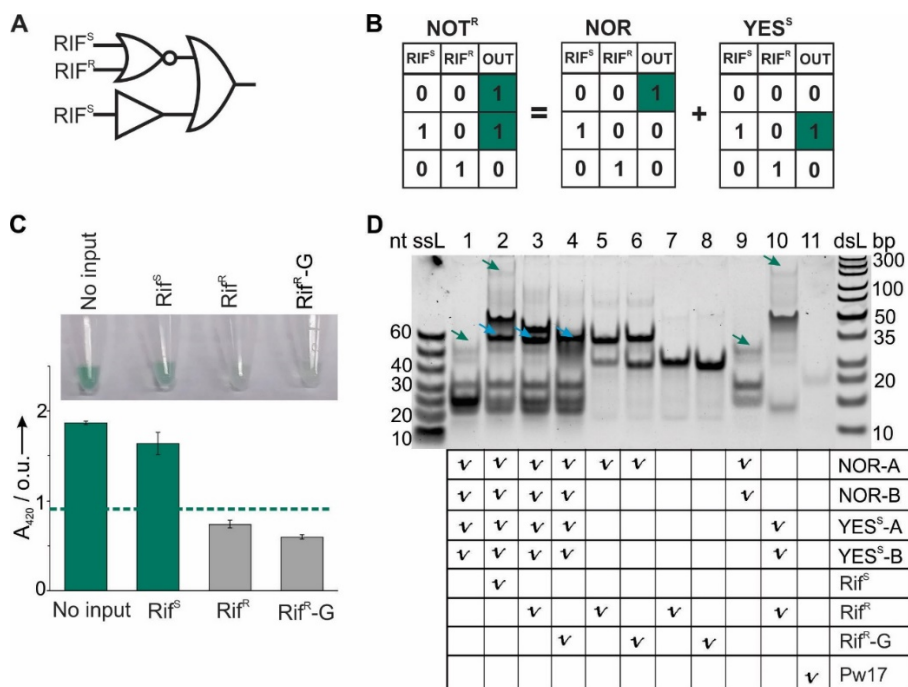
Figure 26: Logical negation of RIF^R (NOT^R function). A) Symbol for NOT^R logic function obtained by mixing NOR and YES^S gates. The NOR/YES^S combination produces an output as if the outputs of NOR and YES^S were recognized as inputs of an OR gate. B) Truth table for NOT^R logic function as a sum of NOR and YES^S logic functions. High output is highlighted in green. C) Absorbance at 420 nm of the samples containing a mixture of NOR-A, NOR-B, YES-A and YES-B in the absence ("No input") or presence of either RIF^S, RIF^R or RIF^R-G inputs. A threshold between low output and high output absorbance is represented by a green dashed line. The data are average values for three independent measurements with standard deviations as error bars. Inset: Image of the tubes containing the same samples. C) Analysis of the samples in non-denaturing 12% PAAG. The lane contents are shown under the gel image. "ssL" and "dsL" designate single-stranded and double-stranded DNA ladder, respectively. The sizes of the ladder fragments are indicated to the sides of the gel. The bands corresponding to either NOR-A/NOR-B or YES-A/YES-B/input complex are indicated by a green arrow. The bands corresponding to NOR-A/input complexes are indicated using a cyan arrow.

NOT^R logic function was obtained by mixing NOR and YES^S gates. The gates computed independently from each other, so their mixture could be perceived as if the outputs of NOR and YES^S were recognized as inputs of an OR gate (Fig. 26A). As a result, high output was expected in the absence of the inputs (due to NOR function) and the presence of RIF^S input (due to YES^S function), while the presence of either RIF^R or RIF^R-G resulted in low output, since the inputs turning

off the NOR were not recognized by YES^S gate (Fig. 26B). The observed response of NOR/YES^S mixture corresponded to the NOT^R logic (Fig. 26C). Formation of high output complexes was also confirmed using gel analysis (Fig. 26D, lanes 1 and 2, indicated by green arrows; compare with lanes 9 and 10). NOR-A/input complexes were formed in the presence of any input (Fig. 26D, lanes 2-4, indicated by cyan arrows).
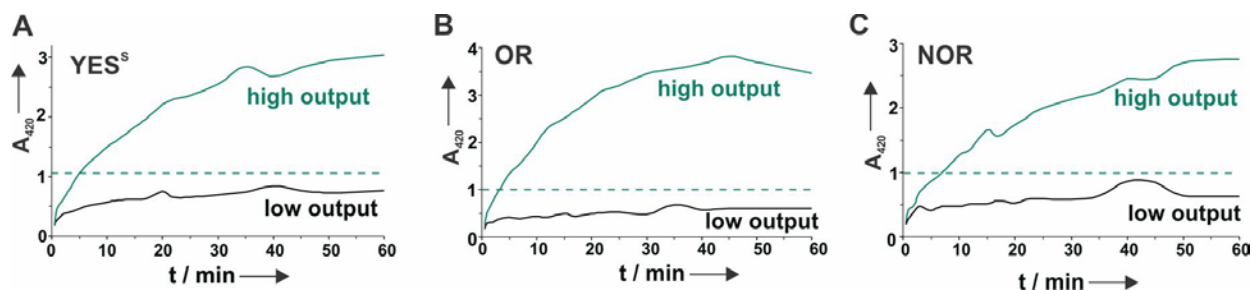
Figure 27: Time dependence of the color generation due to ABTS peroxidation reaction for A) YES^S gate. B) OR gate. C) NOR gate. High output sample (green curve) contained the gate strands (1 µM) in the presence of RIF^S input (1 µM) for the YES^S and OR gates, but only the gate strands (0.5 µM) for the NOR gate. Low output sample (black curve) contained just the gate strands (1 µM) for the YES^S and OR gates. For the NOR gate, it contained the gate strands (0.5 µM) in the presence of RIF^S input (1 µM). A threshold placed at A$_{420}$ ~ 1 o.u. is indicated by a green dashed line.
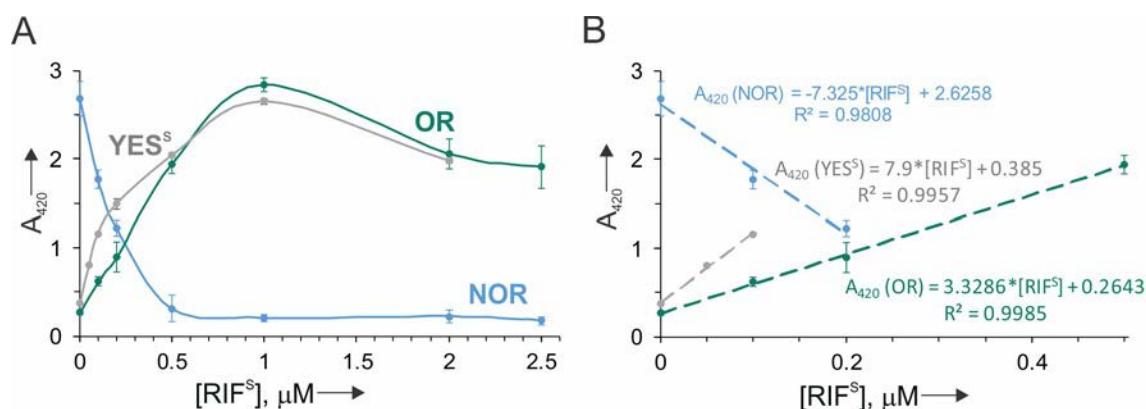
Figure 28: Dependence of the color generation due to ABTS peroxidation reaction on RIF^S input concentration for the YES^S, OR and NOR gates. A) RIF^S concentration in a range of 0-2.5 µM. B) Linear dependence of the absorbance on RIF^S concentration in a range of 0-0.5 µM.

The individual logic gates (YES^S, OR, and NOR) were further characterized in terms of dependence of the signal intensity on the time of peroxidation reaction (Fig. 27) and input concentration (Fig. 28). It can be seen that the high output signal was generated within the first 10 min after the addition of H$_2$O$_2$ and was stable over the course of at least 60 min. The smallest input concentration sufficient to trigger the gate response to statistically differentiate the low and high output signal (detection limit) was found to be in the range of 30-70 nM.
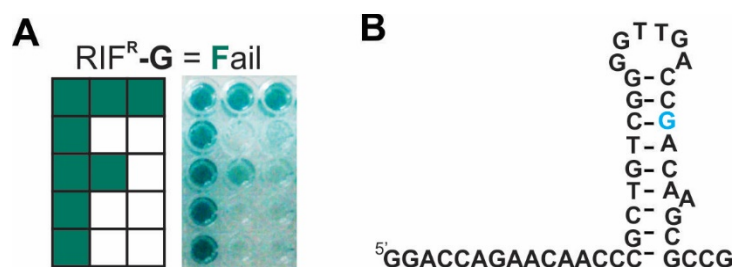
Figure 29: Alphanumerical display response to RIF[R]-G input. (A) Predicted (left) and observed (right) read-out of the logic gate display in response to RIF[R]-G input. (B) Secondary structure RIF[R]-G input, which is more stable and more difficult to open than that of RIF[S] and RIF[R].

When RIF[R]-G input containing a C>G substitution was used instead of RIF[R] input with C>T substitution, the logic gate array displayed letter "F", indicating that RIF-susceptibility test failed (Fig. 26A). The middle well in the 3rd row of the display was not colored as bright as other high output-producing wells. This led to a weaker response of the G4-OR gate to the presence of RIF[R]-G input as compared with RIF[S] and RIF[R] inputs (Fig. 23B). Such imperfect performance of OR gate was due to more stable stem-loop structure of RIF[R]-G (Fig. 23B) than, for example, of RIF[S] (Fig. 23A). At the same time, this structural hurdle did not prevent the formation of the correct alphanumerical read-out of the logic gate display and unequivocal classification of the input as a sequence causing bacterial RIF-resistance.

Further Works

Table 6: Oligonucleotides used in this study.*

Name	Sequence
RIF^S	GGG ATA AGG AGC CAG CTG AGC CAA TTC ATG GAC CAG AAC AAC CCG CTG TCG GGG TTG ACC CAC AAG CGC CGA CTG ACG GCG CT
RIF^R-516G	GGG ATA AGG AGC CAG CTG AGC CAA TTC ATG GGC CAG AAC AAC CCG CTG TCG GGG TTG ACC CAC AAG CGC CGA CTG ACG GCG CT
RIF^R-526T	GGG ATA AGG AGC CAG CTG AGC CAA TTC ATG GAC CAG AAC AAC CCG CTG TCG GGG TTG ACC TAC AAG CGC CGA CTG ACG GCG CT
rB-N	GGCTCCTTCTCCC *TGGGTTGGGTTGGG*
rB-N2	*TTGGGT* GGCATCTTTTTGATGCC GGGAGA
rB516-U	*TTGGGT* GAATTGGCTCAGCT
rB516-S	CTGGACCAT /isp9/ *TGGGTTGGGTTGGG*
rB526-U	*GGGTTGGGT* /isp9/ AACCCCAACAGCGGGGTTGTT
rB526-S	ccaa TTGTAGGTC /isp9/ *TGGGTTGGG*
rB-O	ccctac GGTCAACCCCGACAGCG *TGGGTTGGG*
rB-O2	*GGGTTGGGT* GGTTGTTCTGGTCCA accc
PW17	*GGG TA GGG C GGG TT GGG*

*isp9 – triethylene glycol linker; single-nucleotide substitutions in RIF^R inputs are shown in cyan; nucleotides forming G4 structures are in italics; nucleotides complimentary to the G4 sequence that are not part of the target-recognizing sequence are in lower case.

APPENDIX C: SUPPORTING INFORMATION FOR CHAPTER THREE
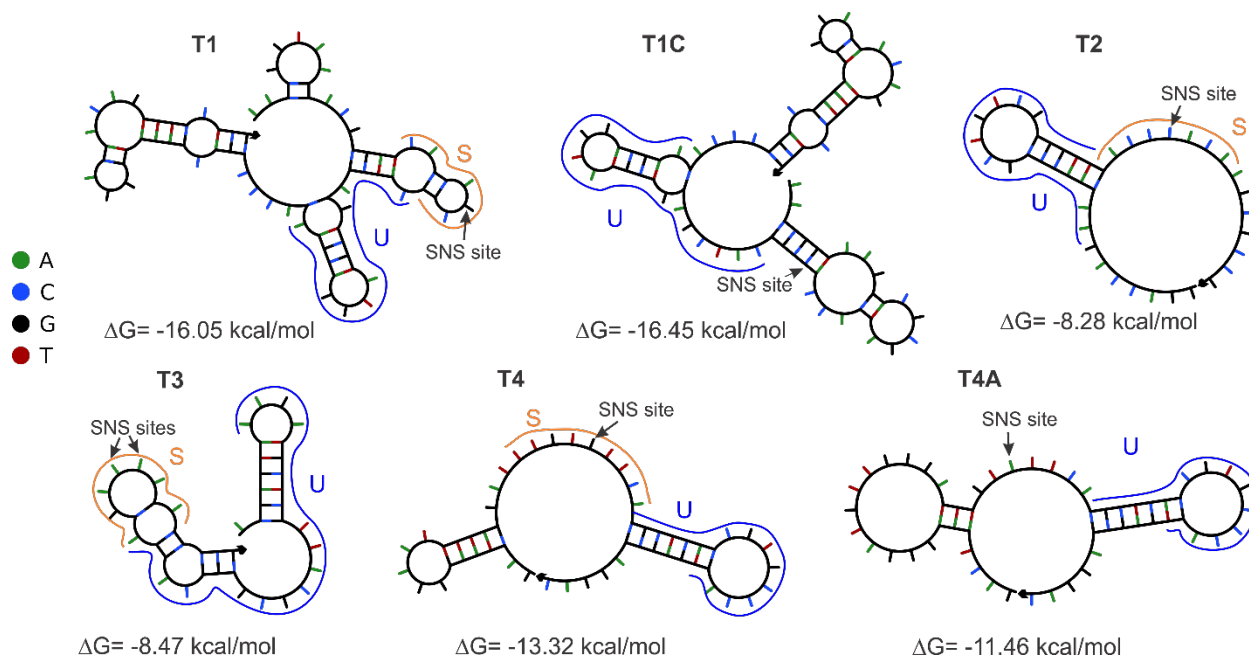
Figure 30: Minimum energy secondary structures for the targets T1, T1C, T2, T3, T3, T4 and T4A calculated by NUPACK (http://www.nupack.org). In the structures, nucleotides are color coded according to their nature (A, C, T or G, see color-coding on the left). The target fragments complementary to the target-recognition elements of strands *U* and *S* of the probes are shown by blue and orange curve, respectively. The predicted values for the secondary structure free energy are indicated. The secondary structures of the targets with SNSs – T2T, T3GA, T3AG, T3TA, T3AT, T3CA, T3AC, and T4T – are the same as for their "wild type" (specific for the sPDz probes) counterparts (T2, T3, and T4).

Table 7: Full sequences for the targets used in the study.[a]

Name		Sequence (5'-3')
T1	**T1**	AACCGGTAAGGACGCGATC*ACCAGCGGCATCGAGGTCGTATGGACGAA*CACCCCGACGAAATGGGACAACAGTTTCCTCG
	T1C	AACCGGTAAGGACGCGATC*ACCACCGGCATCGAGGTCGTATGGA*CGAACACCCCGACGAAATGGGACAACAGTTTCCTCG
T2	**T2**	GGACCAGA*ACAACCCGCTGTCGGGGTTGACCCACA*AGCGCCG
	T2T	GGACCAGA*ACAACCCGCTGTCGGGGTTGACCTACA*AGCGCCG
T3	**T3**	ACGAGTA*AAGATGCTCGTTACGCGCGGCAGGACGAAAAGA*CCCCG
	T3GA	ACGAGTA*AAGATGCTCGTTACGCGCGGCAGGACGAGAAGA*CCCCG
	T3TA	ACGAGTA*AAGATGCTCGTTACGCGCGGCAGGACGATAAGA*CCCCG
	T3CA	ACGAGTA*AAGATGCTCGTTACGCGCGGCAGGACGAGAAGA*CCCCG
	T3AG	ACGAGTA*AAGATGCTCGTTACGCGCGGCAGGACGAAGAGA*CCCCG
	T3AT	ACGAGTA*AAGATGCTCGTTACGCGCGGCAGGACGAATAGA*CCCCG
	T3AC	ACGAGTA*AAGATGCTCGTTACGCGCGGCAGGACGAACAGA*CCCCG
T4	**T4**	GACTAAGGGAATTTTAGGAT*TTGTGTTCACGCTCACCGTGCCCAGT*GAGCGAGGAC
	T4T	GACTAAGGG<u>G</u>ATTTTAGG<u>G</u>T*TTGTTTTCACGCTCACCGTGCCCAGT*GAGCGAGGAC
	T4A	GACTAA<u>A</u>GG<u>G</u>GAT<u>G</u>TT<u>G</u>GGAT*TTGTATTCACGCTCACCGTGCCCAGT*GAGCGAGGAC

[a] Nucleotides at the probed SNS positions are in red. Other SNS positions, if present, are underlined. Target fragments interrogated by the probes are in italics.

Table 8: The strand sequences for the P1 probes targeting a fragment of the *MTB katG* gene.

Name		Sequence (5'-3')[a]
P1[6:6]	[6]U1	TTCGTCCATACGACCTCGATGTGGGGTTGGG
	[6]S1	GGGTTGGGTCCGCTGGTG
P1[3:9]	[3]U1	TTCGTCCATACGACCTCGATGTGGGGTT
	[9]S1	TGGGATGGGTTGGGTCCGCTGGTG
P1[9:3]	[9]U1	TTCGTCCATACGACCTCGATGTGGGGTTGGGTAGGGT
	[3]S1	TTGGGTCCGCTGGTG
P1[6:6]-*teg*	[6]U1-*teg*	TTCGTCCATACGACCTCGATG/iSp9/TGGGTTGGG
	[6]S1-*teg*	GGGTTGGGT/iSp9/CCGCTGGTG
P1[3:9]-*teg*	[3]U1-*teg*	TTCGTCCATACGACCTCGATG/iSp9/TGGGTT
	[9]S1-*teg*	TGGGATGGGTTGGGT/iSp9/CCGCTGGTG
P1[9:3]-*teg*	[9]U1-*teg*	TTCGTCCATACGACCTCGATG/iSp9/TGGGTTGGGTAGGGT
	[3]S1-*teg*	TTGGGT/iSp9/CCGCTGGTG
sl-P1[6:6]	*sl*-[6]U1	CCAACTTCGTCCATACGACCTCGATGTGGGTTGGG
	sl-[6]S1	GGGTTGGGTCCGCTGGTGCAAC
sl-P1[3:9]	*sl*-[3]U1	AACCTTCGTCCATACGACCTCGATGTGGGTT
	sl-[9]S1	TGGGATGGGTTGGGTCCGCTGGTGCATCC
sl-P1[9:3]	*sl*-[9]U1	CCTACTTCGTCCATACGACCTCGATGTGGGTTGGGTAGGGT
	sl-[3]S1	TTGGGTCCGCTGGTGCCAA
sl-P1[6:6]-*teg*	*sl*-[6]U1-*teg*	CCAACTTCGTCCATACGACCTCGATG/iSp9/TGGGTTGGG
	sl-[6]S1-*teg*	GGGTTGGGT/iSp9/CCGCTGGTGCAAC
sl-P1[3:9]-*teg*	*sl*-[3]U1-*teg*	AACCTTCGTCCATACGACCTCGATG/iSp9/TGGGTT
	sl-[9]S1-*teg*	TGGGATGGGTTGGGT/iSp9/CCGCTGGTGCATCC

sl-P1[9:3]-teg	sl-[9]U1-teg	CCTACTTCGTCCATACGACCTCGATG/iSp9/TGGGTTGGGTAGGGT
	sl-[13]S1-teg	TTGGGT/iSp9/CCGCTGGTGCCAA

[a]Nucleotides corresponding to the G4 signal-transducing elements are in green. Nucleotides of the target-recognizing elements complementary to the target SNS sites are in red. Self-complementary regions of the strands are underlined. Nucleotides of the strands that do not belong to the target-recognizing or signal-transducing domains are in grey. Triethylene glycol (*teg*) linkers are indicated by "/iSp9/" (notation used by IDT, Inc.).

Table 9: The strand sequences for the P2 probes targeting a fragment of the *MTB rpoB* gene.

Name[a]		Sequence (5'-3')[b]
P2[6:6]	*sl*-[6]**U2**	GGGTTGGGTAACCCCAACAGCGGGTTGT
	[6]**S2**	TTGTGGGTCTGGGTTGGG
P2[3:9]	*sl*-[3]**U2**	TTGGGTAACCCCAACAGCGGGTTGT
	[9]**S2**	TTGTGGGTCTGGGTTGGGTAGGGT
P2[9:3]	*sl*-[9]**U2**	TGGGTTGGGTTGGGTAACCCCAACAGCGGGTTGT
	[3]**S2**	TTGTGGGTCTGGGTT
P2[6:6]-*teg*	*sl*-[6]**U2**-*teg*	GGGTTGGG/iSp9/AACCCCAACAGCGGGTTGT
	[6]**S2**-*teg*	TTGTGGGTC/iSp9/GGGTAGGG
P2[9:3]-*teg*	*sl*-[9]**U2**-*teg*	TGGGTTGGGTTGGGT/iSp9/AACCCCAACAGCGGGTTGT
	[3]**S2**-*teg*	TTGTGGGTC/iSp9/GGGTT
sl-**P2**[6:6]	*sl*-[6]**U2**	GGGTTGGGTAACCCCAACAGCGGGTTGT
	sl-[6]**S2**	CCCTATTGTGGGTCTGGGTAGGG
sl-**P2**[3:9]	*sl*-[3]**U2**	TTGGGTAACCCCAACAGCGGGTTGT
	sl-[9]**S2**	CCAACCTTGTGGGTCTGGGTTGGGTTGGGT
sl-**P2**[9:3]	*sl*-[9]**U2**	TGGGTTGGGTTGGGTAACCCCAACAGCGGGTTGT
	sl-[3]**S2**	AACCTTGTGGGTCTGGGTT
sl-**P2**[6:6]-*teg*	*sl*-[6]**U2**-*teg*	GGGTTGGG/iSp9/AACCCCAACAGCGGGTTGT
	sl-[6]**S2**-*teg*	CCCTATTGTGGGTC/iSp9/GGGTAGGG
sl-**P2**[3:9]-*teg*	*sl*-[3]**U2**-*teg*	TTGGGT/iSp9/AACCCCAACAGCGGGTTGT
	sl-[9]**S2**-*teg*	CCAACCTTGTGGGTC/iSp9/TGGGTTGGGTTGGGT
sl-**P2**[9:3]-*teg*	*sl*-[9]**U2**-*teg*	TGGGTTGGGTTGGGT/iSp9/AACCCCAACAGCGGGTTGT
	sl-[3]**S2**-*teg*	AACCTTGTGGGTC/iSp9/GGGTT
	sl-[6]**U2**-*teg*	GGGTTGGG/iSp9/AACCCCAACAGCGGGTTGT

sl-**P2**[6:6]-*teg*-**T**	*sl*-[6]**S2**-*teg*-**T**	CCCTATTGTTGGTC/iSp9/GGGTAGGG

[a]Strands *U* contained an intrinsic stem-loop motif in all the sPDz probe designs.

[b]Nucleotides corresponding to the G4 signal-transducing elements are in green. Nucleotides of the target-recognizing elements complementary to the target SNS sites are in red. Self-complementary regions of the strands are underlined. Nucleotides of the strands that do not belong to the target-recognizing or signal-transducing domains are in grey. Triethylene glycol (*teg*) linkers are indicated by "/iSp9/" (notation used by IDT, Inc.). In strand *U* sequences, a SNS G>A was introduced (shown in blue).

Table 10: The strand sequences for the P3 probes targeting a fragment of the *Mabs rrl* gene.

Name		Sequence (5'-3')[a]
P3[6:6]	[6]**U3**	GGGTTGGGTTCCTGCCGCGCGTAACGAGCATCTT
	[6]**S3**	GTCTTTTCGTGGGTTGGG
P3[3:9]	[3]**U3**	TTGGGTTCCTGCCGCGCGTAACGAGCATCTT
	[9]**S3**	GTCTTTTCGTGGGTTGGGTTGGGT
P3[9:3]	[9]**U3**	TGGGTTGGGTTGGGTTCCTGCCGCGCGTAACGAGCATCTT
	[3]**S3**	GTCTTTTCGTGGGTT
P3[6:6]-*teg*	[6]**U3**-*teg*	GGGTTGGGT/iSp9/TCCTGCCGCGCGTAACGAGCATCTT
	[6]**S3**-*teg*	GTCTTTTCG/iSp9/TGGGTTGGG
P3[3:9]-*teg*	[3]**U3**-*teg*	TTGGGT/iSp9/TCCTGCCGCGCGTAACGAGCATCTT
	[9]**S3**-*teg*	GTCTTTTCG/iSp9/TGGGTTGGGTTGGGT
P3[9:3]-*teg*	[9]**U3**-*teg*	TGGGTTGGGTTGGGT/iSp9/TCCTGCCGCGCGTAACGAGCATCTT
	[3]**S3**-*teg*	GTCTTTTCG/iSp9/TGGGTT
sl-**P3**[6:6]	[6]**U3**	GGGTTGGGTTCCTGCCGCGCGTAACGAGCATCTT
	sl-[6]**S3**	CCTAAGTCTTTTCGTGGGTTGGG
sl-**P3**[3:9]	[3]**U3**	TTGGGTTCCTGCCGCGCGTAACGAGCATCTT
	sl-[9]**S3**	CCCAAGTCTTTTCGTGGGTTGGGTTGGGT
sl-**P3**[9:3]	[9]**U3**	TGGGTTGGGTTGGGTTCCTGCCGCGCGTAACGAGCATCTT
	sl-[3]**S3**	AACCGTCTTTTCGTGGGTT
sl-**P3**[6:6]-*teg*	[6]**U3**-*teg*	GGGTTGGGT/iSp9/TCCTGCCGCGCGTAACGAGCATCTT
	sl-[6]**S3**-*teg*	CCTAAGTCTTTTCG/iSp9/TGGGTTGGG
	[3]**U3**-*teg*	TTGGGT/iSp9/TCCTGCCGCGCGTAACGAGCATCTT

sl-**P3**[3:9]-*teg*	*sl*-[9]**S3**-*teg*	CCTAAGTCTTTTCG/iSp9/TGGGTTGGGTTGGGT
sl-**P3**[9:3]-*teg*	[9]**U3**-*teg*	TGGGTTGGGTTGGGT/iSp9/TCCTGCCGCGCGTAACGAGCATCTT
	sl-[3]**S3**-*teg*	AACCGTCTTTTCG/iSp9/TGGGTT

[a]Nucleotides corresponding to the G4 signal-transducing elements are in green. Nucleotides of the target-recognizing elements complementary to the target SNS sites are in red. Self-complementary regions of the strands are underlined. Nucleotides of the strands that do not belong to the target-recognizing or signal-transducing domains are in grey. Triethylene glycol (*teg*) linkers are indicated by "/iSp9/" (notation used by IDT, Inc.).

Table 11: The strand sequences for the P4 probes containing a fragment of the IAV segment 7.

Name[a]		Sequence (5'-3')[b]
P4⁶:⁶	⁶U4	ACTGGGCACGGTGAGCGTTGGGTTGGG
	sl-⁶S4	GGGTTGGGTTGAACACAA
P4³:⁹	³U4	ACTGGGCACGGTGAGCGTTGGGTT
	sl-⁹S4	TGGGTTGGGTTGGGTTGAACACAA
P4³:⁹-T	³U4	ACTGGGCACGGTGAGCGTTGGGTT
	⁹S4-T	TGGGTTGGGTTGGGTTGAAAACAA
P4³:⁹-A	³U4	ACTGGGCACGGTGAGCGTTGGGTT
	⁹S4-A	TGGGTTGGGTTGGGTTGAATACAA
sl-P4³:⁹-A	³U4	ACTGGGCACGGTGAGCGTTGGGTT
	sl-⁹S4-A	TGGGTTGGGTTGGGTTGAATACAAAACC
P4⁹:³	⁹U4	ACTGGGCACGGTGAGCGTTGGGTTGGGTTGGGT
	sl-³S4	TTGGGTTGAACACAA

[a] Strand S of the probes P4³:⁹ targeting T4 contained an intrinsic conformational constraint that was absent in case of strands S targeting T4T or T4A.

[b] Nucleotides corresponding to the G4 signal-transducing elements are in green. Nucleotides of the target-recognizing elements complementary to the target SNS sites are in red. Self-complementary regions of the strands are underlined. Nucleotides of the strands that do not belong to the target-recognizing or signal-transducing domains are in grey. In strand ⁹U4, there is a G-triplet (shown in blue) in the target-recognizing element of strand U, which is brought in the proximity with the signal-transducing element of the same strand by a stem-loop (underlined nucleotides).
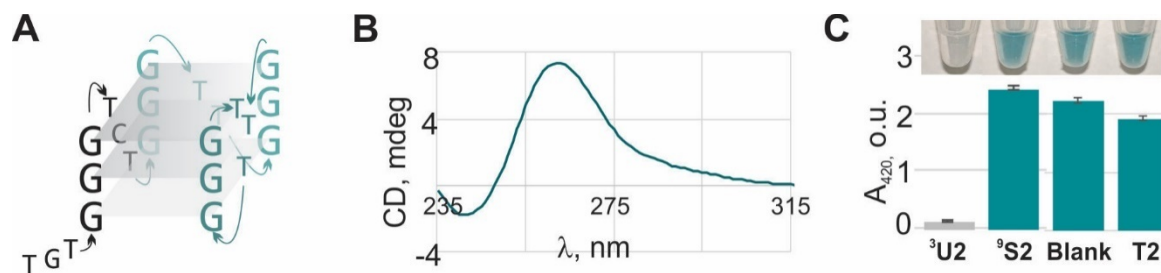
Figure 31: Folding of strand ⁹S2 into a monomolecular parallel G4 structure with peroxidase-like activity. A. Schematic illustration of a parallel G4 folding. B. CD spectrum C. Peroxidase activity of strands ³U2 or ⁹S2 alone, or both strands in the absence (blank) or presence of the fully complementary T2 target.
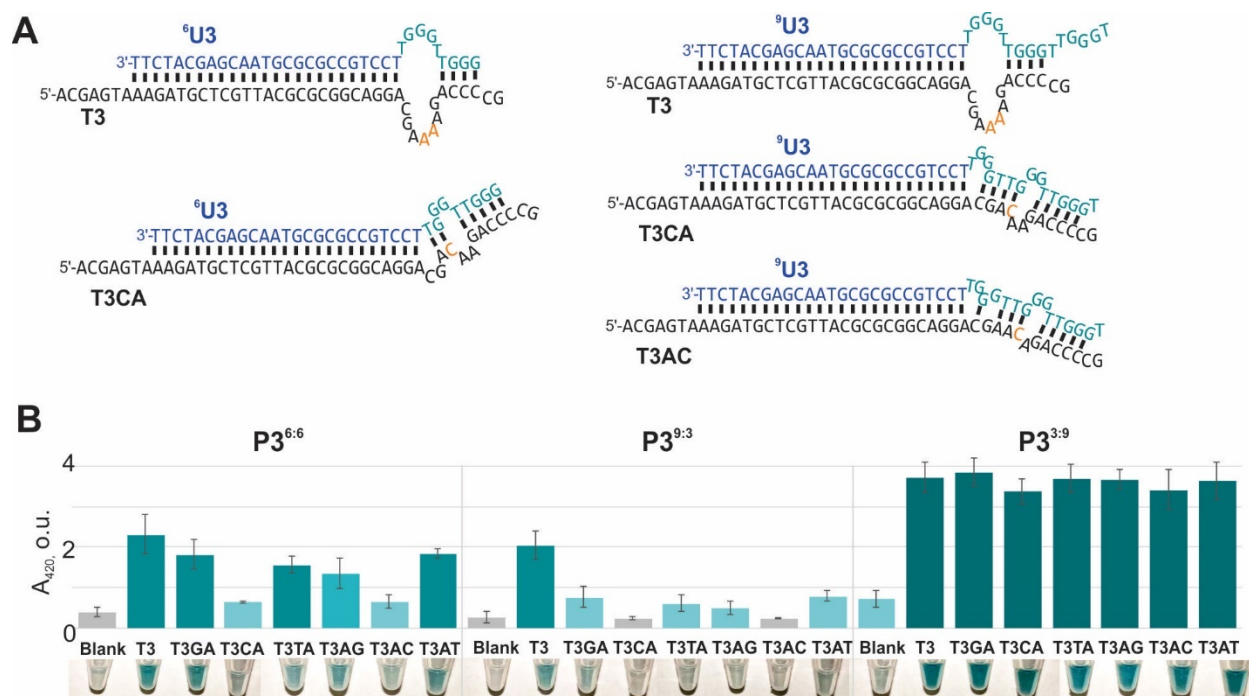
Figure 32: Performance and selectivity of the sPDz probes interrogating a fragment of *Mabs rrl* gene. A. Proposed interactions between T3 or its SNS-counterparts T3CA and T3AC, and either ⁶U3 or ⁹U3 strands. The structures are drawn as predicted by NUPACK (https://www.nupack.org). The signal-transducing region is shown in green, and the target-recognizing element of strand *U* is in blue; the SNS position is highlighted in orange. B. Probes P3⁶:⁶, P3⁹:³ and P3³:⁹ in the absence of the targets ("Blank"), or in the presence of the specific target T3 or SNS-containing targeted (as indicated).
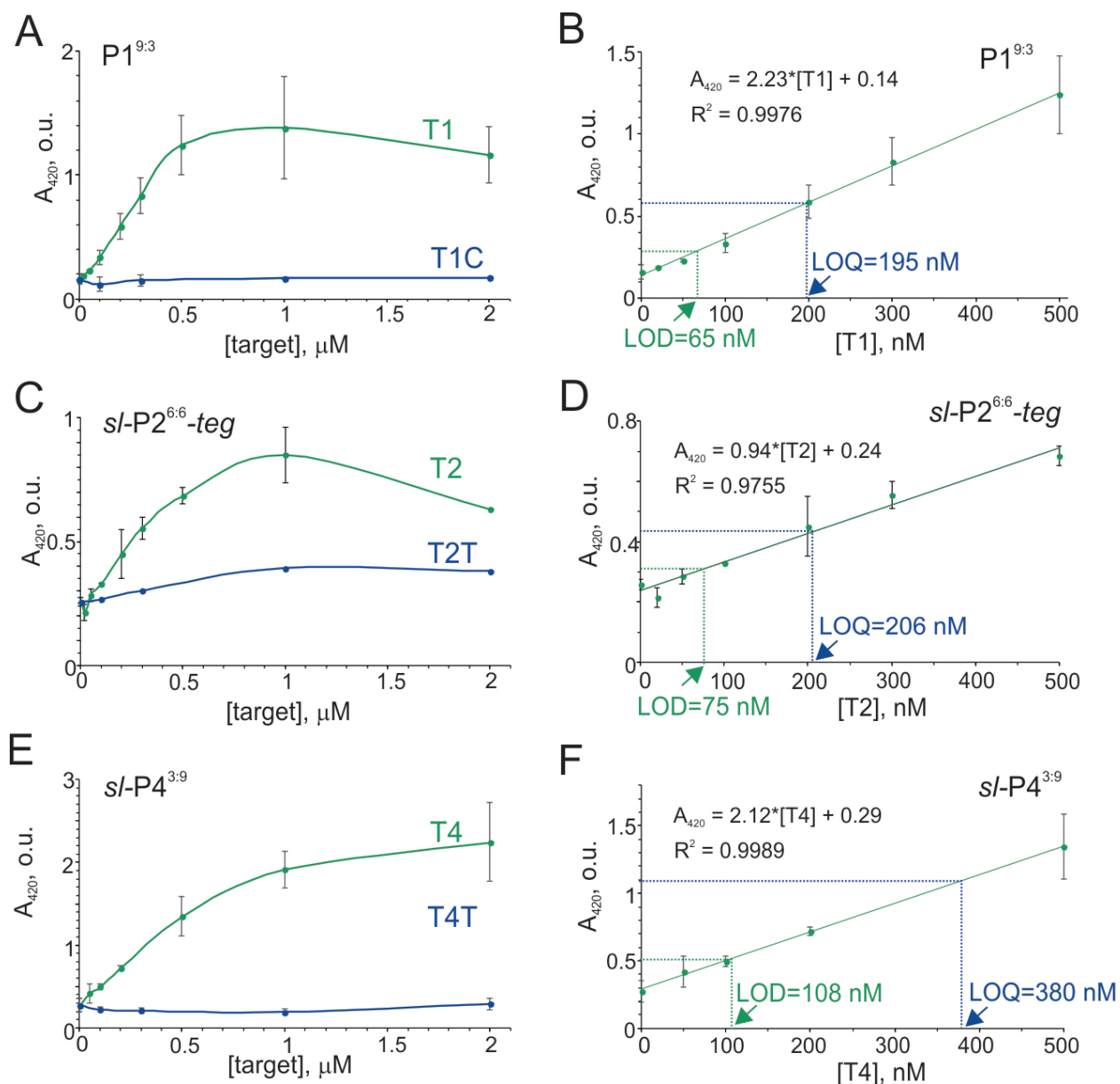
Figure 33: Dependence of the signal of the sPDz probes of the P1, P2, and P4 sets on the target concentration. A,C,E. Absorbance at 420 nm for samples containing 1 μM P19:3 (A), *sl*-P2$^{6:6}$-*teg* (C), or *sl*-P4$^{3:9}$ (E) with varying concentrations of T1 and T1C (A), T2 and T2T (C), or T4 and T4T (E). B,D,F. Linear dynamic range of T1 with P1$^{9:3}$ (B), T2 with *sl*-P2$^{6:6}$-*teg* (D), and T4 with *sl*-P4$^{3:9}$ (F) indicating the limits of detection (LOD) and quantitation (LOQ). LOD and LOQ were calculated as the target concentrations triggering the signals equal to the average background signal plus three and ten standard deviations from the blank, respectively.

Table 12: Target sequences and probe designs reported in the literature.

Target sequence[a]	C-clusters in the target's fragment interacting with strand S	SNS in the C-cluster	G-clusters in the target's fragment interacting with strand S	C-clusters in the target's fragment interacting with strand U	Recommended G4 splitting[b]	Used G4 splitting	Additional elements used	Reference
GGGTGCACC AGGACACGG TTTT	no	no	yes	no	$P^{3:9*}$	$P^{6:6}$	teg-linkers	103
ACCCAGCACA ACATTAGAG	yes	yes	no	no	$P^{6:6}$	$P^{6:6}$	teg-linkers	141
GGCAGCAATT TCACCAGTAC TACAGTTAAG GCCGCCTGT	yes	no	no	yes	$P^{6:6}$	$P^{9:3}$	Competiti on probe was used for	106
CTGGCGGCG CTGGGGCCC GGCGGTCTG TCACGTGAG	no	no	yes	no	$P^{3:9}$	$P^{9:3}$	improved selectivity	108
AGAGTCCACC AACCGCTTCA GATTAG	no	no	no	yes	$P^{3:9*}$	$P^{3:9}$	TTTT-linkers to increase signal	146
CACCAGCGGC ATCGAGGTCG TATGGACGAA	yes	no	yes	no	$P^{9:3*}$	$P^{9:3}$	teg-linkers improved the signal intensity	This paper
ACAACCCGCT GTCGGGGTT GACCCACA	yes	yes	no	yes	$P^{6:6}$	$P^{6:6}$	teg-linkers were required to increase	This paper

							the signal intensity; while "closing" of strand S helped to discriminate C>T SNS	
AAGATGCTCG TTACGCGCGG CAGGACGAA AAGACCCC	yes	no	no	no	$P^{9:3*}$	$P^{9:3}$	"closing" of strand S into a stem-loop helped to discriminate all SNSs in two adjacent positions	This paper
TTGTGTTCAC GCTCACCGTG CCCAGT	no	no	no	yes	$P^{3:9*}$	$P^{3:9}$		This paper

[a] The position of SNSs is in red. An underlined nucleotide was not complementary to strands S and U (gap nucleotide)

[b] Sub-optimal, but acceptable (depending on the type of SNS) probe designs are indicated with an asterisk (*).
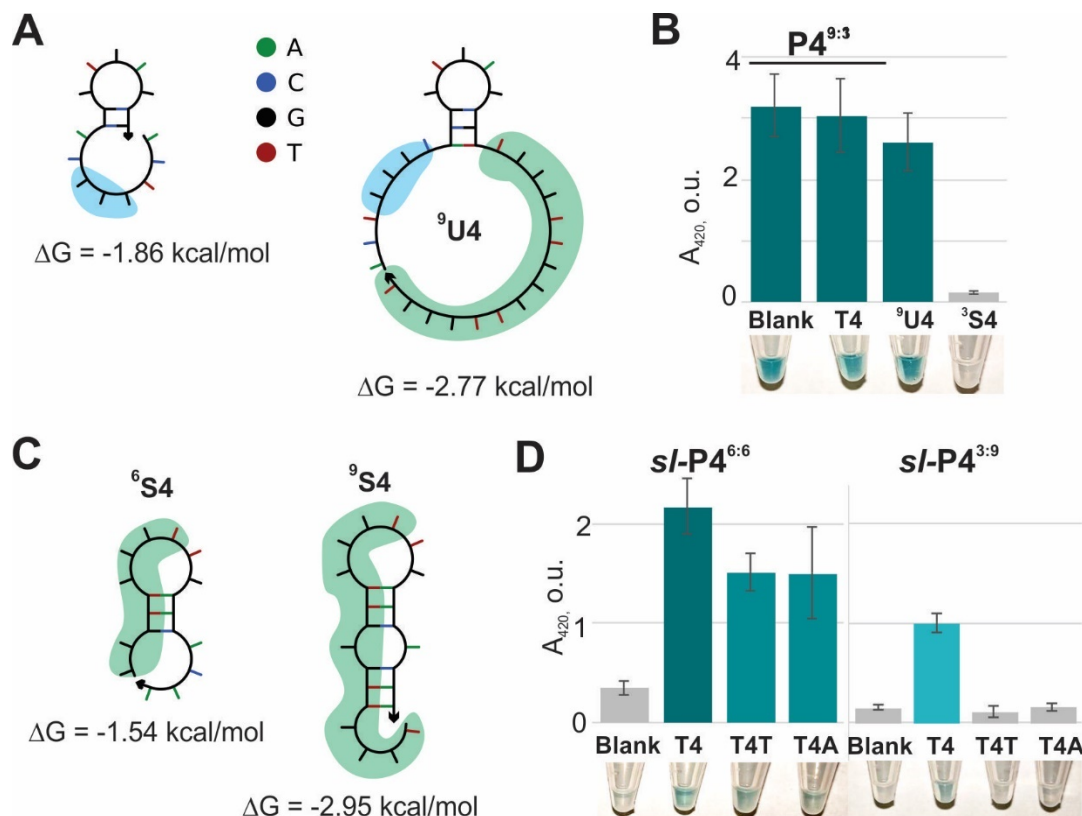
Figure 34: Performance and selectivity of the sPDz probes interrogating a fragment of segment 7 of IAV. A. Minimum energy secondary structures for the target-recognition element (left) or whole sequence of strand ^{9}U4 calculated by NUPACK (http://www.nupack.org). In the structures, nucleotides are color coded according to their nature (A, C, T or G, color-coding is indicated). The nucleotides of the signal-transducing element of the strand are highlighted by a green area. Light-blue area highlights a G-triplet in the target-recognizing element of the strand. B. Absorbance at 420 nm and tube images for the samples containing probe P4$^{9:3}$ in the absence of the targets ("Blank"), or in the presence of the specific target T4; or strands ^{9}U4 or ^{3}S4 of the probe alone. C. Minimum energy secondary structures for the target-recognition element (left) or whole sequence of strand ^{9}U4 calculated by NUPACK. D. Absorbance at 420 nm and tube images for the samples containing either probe sl-P4$^{6:6}$ or probe sl-P4$^{3:9}$ in the absence of the targets ("Blank"), in the presence of the specific target T4, or one of the non-specific targets T4T and T4A.

Table 13: Performance and selectivity of the sPDz probes of the P4 set.[a]

Probe	P4 set					
	P4$^{6:6}$	P4$^{3:9}$	P4$^{3:9}$-T	P4$^{3:9}$-A	sl-P4$^{3:9}$-A	P4$^{9:3}$
S/B	6.3±0.1	**6.5±0.1**	**7.2±0.1**	6.0±0.1	4.7±0.2	1.0±0.5
SF^{b}	36	**99**	**96**	58	76	0

[a] The values for S/B and/or SF indicating poor performance of the probe are in grey. Acceptable probes are in bold.

[b] In case of a mathematical value for SF>100%, the SF was made equal to 100%. When the signal triggered by the mismatched target was higher when the signal in the presence of the fully complementary target, the SF was 0.

Further Works

Table 14: The strand sequences for the G4 loop-modifications study.

Sequence (5'-3')[b]
GGG T GGG T GGG T GGG T
GGG TT GGG TT GGG TT GGG T
GGG TTT GGG TTT GGG TTT GGG T
GGG TTTT GGG TTTT GGG TTTT GGG T
GGG TTTTT GGG TTTTT GGG TTTTT GGG T
GGG A GGG A GGG A GGG A
GGG AT GGG AT GGG AT GGG A
GGG ATT GGG ATT GGG ATT GGG A
GGG ATTT GGG ATTT GGG ATTT GGG A
GGG ATTTT GGG ATTTT GGG ATTTT GGG A
GGG ATT GGG TTT GGG TTT GGG T
GGG TTT GGG ATT GGG TTT GGG T
GGG TTT GGG TTT GGG ATT GGG T
GGG TTT GGG TTT GGG TTT GGG A
GGG ATT GGG ATT GGG ATT GGG AA
GGG ATT GGG ATT GGG ATT GGG AC
GGG ATT GGG ATT GGG ATT GGG AAC

APPENDIX D: SUPPORTING INFORMATION FOR CHAPTER FOUR

Experimental Section

Materials. Oligonucleotides were custom-synthesized by IDT, Inc. (Coralville, IA, USA). Auramine O and brilliant green were purchased from MP Biomedicals (Solon, OH, USA). Basic fuchsin and Michler's ketone were purchased from Alfa Aesar (Ward Hill, MA, USA). Crystal violet and malachite green were purchased from Sigma-Aldrich (St. Louis, MO, USA). Fluorescein, methyl violet 2B, pararosaniline, thioflavin T, and victoria blue B were purchased from ACROS Organics (Morris, NJ, USA). N¬-methyl mesoporphyrin IX was purchased from Frontier Scientific (Logan, UT, USA). Rhodamine 6G was from Armor Forensics (Jacksonville, FL, USA). Quartz cuvettes for measuring absorbance, fluorescence, or circular dichroism were purchased from Starna Cells (Atascadero, CA, USA). NASBA liquid kit was purchased from Life Sciences Advanced Technologies Inc. (Saint-Petersburg, FL). RiboRuler Low Range RNA Ladder was from ThermoFisher Scientific (Waltham, MA). Agarose was from Lonza (Basel, Switzerland).

Characterization of dapoxyl sulfonyl fluoride (DSF). Dapoxyl sulfonyl fluoride (DSF) was synthesized as previously described[214] and kindly provided by Dr. Kikuchi. The dye sample was characterized prior to use in this work. ESI-MS: $[M+H]^+$ calculated=347.09, $[M+H]^+$ observed=347.08. 1H-NMR (CDCl$_3$): 8.30 (2H, d); 8.09 (2H, d); 7.61 (2H, d); 7.34 (1H, s); 6.77 (2H, d); 3.05 (6H, s). 19F-NMR (CDCl$_3$): 66.26 (s).

Affinity constant determination. Samples contained 0-50 μM DAP-10-42 and 2 μM of the respective dye. DSF, AO, and MK were suspended in buffer A (20 mM Tris-HCl, pH 7.4, 20 mM

KCl, and 25 mM MgCl$_2$); all other dyes were measured in buffer B (50 mM HEPES, pH 7.4, 20 mM KCl, 50 mM MgCl$_2$, 120 mM NaCl, 1% DMSO, and 0.03% Triton X-100). Fluorescence data was blank subtracted then normalized to the largest value in range as 100% Turn-On (TO%). Excitation/emission wavelengths were as follows: DSF, 390/507 nm; TFT, 440/500 nm; CV and MV, 590/630 nm; MG and BG, 617/660 nm; PR, 557/600 nm; BF, 560/600 nm; VB, 635/690 nm; AO, 475/540 nm; MK, 390/450 nm. Fluorescence intensity was plotted as a function of the aptamer concentration ([DAP]) and analyzed using Origin software. The K$_d$ values were calculated using Equation 2:

$$S/B = \frac{(A_T + D_T + K_d) - \sqrt{(A_T + D_T + K_d)^2 - 4A_T D_T}}{2A_T} \qquad (2)$$

where S/B is the signal-to-background ratio calculated by dividing the fluorescence intensity of the aptamer-containing sample by the average intrinsic dye's fluorescence, AT is the total concentration of DAP-10-42 in a given sample and DT=2 μM is the total dye concentration.

Quantum yield analysis. Samples containing 0-7.5 μM DAP-10-58 and 7.5 μM AO, or 7.5 μM DAP-10-58 and 0-7.5 μM AO, were prepared in the buffer A. Standards contained fluorescein in 0.1 M sodium hydroxide or rhodamine 6G in ethanol. Absorbance spectra were collected with a NanoDrop One^C (ThermoScientific, Waltham, MA, USA) at ambient temperature using a 1 cm quartz cuvette. Fluorescent spectra were recorded upon excitation at 475 nm using a Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies, Santa Clara, CA, USA) at ambient temperature using a 3 mm quartz cuvette.

Job plot analysis. Samples were prepared in the assay buffer (20 mM Tris-HCl, pH 7.4, 20 mM KCl, 25 mM MgCl$_2$) at 4 μM total molar concentration of both DAP-10-42 and either DSF or

auramine O (AO). Fluorescence emission was measured using a Cary Eclipse Spectrofluorimeter at 507 nm upon excitation at 399 nm (DSF), or at 540 nm upon excitation at 475 nm (AO) in a 1 cm quartz cuvette.

SDA probe assay. Samples (50 µl) containing SDA-S and SDA-U (each at 0.5 µM) in buffer A supplemented with 1% DMSO and 2 µM dye (DSF, AO, CV, or TFT) in the presence of a synthetic target (*katG* or *katG*-G>C) at the indicated concentration or 2% (v/v) NASBA sample were incubated at 22 °C for 15 min prior to measuring their fluorescence using a Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies, Santa Clara, CA, USA) at ambient temperature using a 1 cm quartz cuvette. Alternatively, fluorescence measurements were performed in a time-drive mode using an Infinite® 200 PRO Plate Reader (Tecan, Switzerland). The measurements in this case began within 5-10 min of the target addition. As a negative control, a sample in the absence of the target was used.

Calculation of LOD and LOQ was performed using a 3σ/S and 10σ/S rule,[215] respectively, where σ is the standard deviation of the y-axis intercept, and S is the slope of the linear trendline for the signal concentration dependence in the linear dynamic range. The data processing was done using OriginLab© 2018b graphing software.

NASBA reaction was performed according to the manufacturer's instruction, with the sample volume decreased to 12 µl. Specifically, samples were prepared by mixing 4 µl of 3×NASBA Reaction Buffer, 2 µl of 6×Nucleotide Mix, 2 µl of 1.5 µM primer mix, and 1 µl of either RNase-free water (for no-target control, NTC) or total bacterial RNA (10 pg/µl). The samples were incubated at 65 °C for 2 min followed by cooling to 41 °C for 10 min. Then, 3 µl of 4×NASBA enzyme cocktail was added, and the samples were incubated at 41 °C for 90 min. The obtained samples were analyzed by 2%

agarose gel electrophoresis (1 µl per lane) with RNA in the gel stained by GelRed. Amplicon concentration in the NASBA samples was determined by comparing the intensity of the corresponding band with the intensity of the band containing a 300-nt marker of the RiboRuler Low Range RNA ladder using a BioRad GelDoc XR+ Molecular Imager coupled with ImageLab software.

Table 15: Sequences of oligonucleotides used in this study.

Name	Sequence, 5'→3'[a]
ssDNA	GGACATTTTAGAACATTACATACCCCG
slDNA	GGACCAGAACAACC<u>CGCTGTCGG</u>GGGTTGA<u>CCGACAA</u><u>GCG</u>CCG
TBA	GGTTGGTGTGGTTGG
PW17	GGGTAGGGCGGGTTGGG
CV30S	AACGACCACCGGTGCGCCGTACAGGTAACTAGCGTCGTCGTT
Tel23a	AGGGTTAGGGTTAGGGTTAGGGT
MG1-3	CTCAGATCTAACCTTGTTAAATTGAG
DIR2-1	GACGACGACGCTAGGAAGGCGTTGGTGGGCACGCCGGTCGTC
AT11	TGGTGGTGGTTGTTGTGGTGGTGGTGGT
EAD2	CTGGGAGGGAGGGAGGGA
Agro100	GGTGGTGGTGGTTGTGGTGGTGGTGG
HT	AGGGTTAGGGTTAGGGTTAGGG
RHT	GGGTTAGGGTTAGGGTTAGGGA
DAP-10-42	<u>CAATT</u><u>ACG</u>*GGGGAGGGTGTGTGGTCTTGCTTGGTT*<u>CGTATTG</u>
SDA-U	TTCGTCCATACGACCTCGATGTT<u>CTACG</u>*GGGGAGGGTGTGTGGTTTTG* <u>GTCAT</u>
SDA-S	<u>ATGACC</u>*TTGGTT*<u>CGTAG</u>TTCCG**C**TGGTG
katG	AACCGGTAAGGACGCGATCACCA**G**CGGCATCGAGGTCGTATGGACGA ACACCCCGACGAAATGGGACAACAGTTTCCTCG
katG-G>C	AACCGGTAAGGACGCGATCACCA**C**CGGCATCGAGGTCGTATGGACGA ACACCCCGACGAAATGGGACAACAGTTTCCTCG
NASBA-*katG*	gggagaagggcuugggcuggaagagcucguauggcaccggaaccgguaaggacgcgaucacca**g**cggcaucgag gucguauggacgaacaccccgacgaaaugggacaacaguuuccucgagauccuguacggcuacgagugg

a Stem-forming nucleotides are underlined; nucleotides of the dye-binding aptamer core are in italics "linker" nucleotides in the sensor strands are in grey; single-nucleotide substitution positions in *katG* and *katG*-Mut are in bold; complementary nucleotides of the sensor strands and the targets are color-coded; ribonucleotides in the NASBA amplicon of the *katG* gene are in lowercase, with nucleotides absent in the mRNA indicated in grey lowercase.

Table 16: Characteristics of DAP-10-42 interaction with different fluorogenic dyes.

Dye	K_d (µM)	S/B$_{MAX}$[a]	LOD[b] (nM)	Dye	K_d (µM)	S/B$_{MAX}$[a]	LOD[b] (nM)
DSF	0.011 ± 0.022	132	5.4	PR	24 ± 1.6	18	N/D
CV	0.21 ±0.07	131	4.7	BF	5.9 ± 1.2	20	N/D
MV	0.42 ± 0.05	104	N/D	VB	3.4 ± 2.2	2	N/D
MG	6.9 ± 0.82	258	N/D	AO	0.74 ± 0.12	2070	2.6
BG	4.8 ± 0.87	59	N/D	MK	6.6 ± 0.89	6	N/D
TFT	10 ± 1.1	1870	2.8	NMM	130 ± 79	121	N/D

a The signal-to-background ratio (S/B$_{MAX}$) was determined from the fitting of the data in Figure 1B using the Origin software. It serves as a measure of the dyes' fluorescence turn-on.

b The LOD values are determined for the SDA probe in the presence of the selected dyes. N/D – not determined.
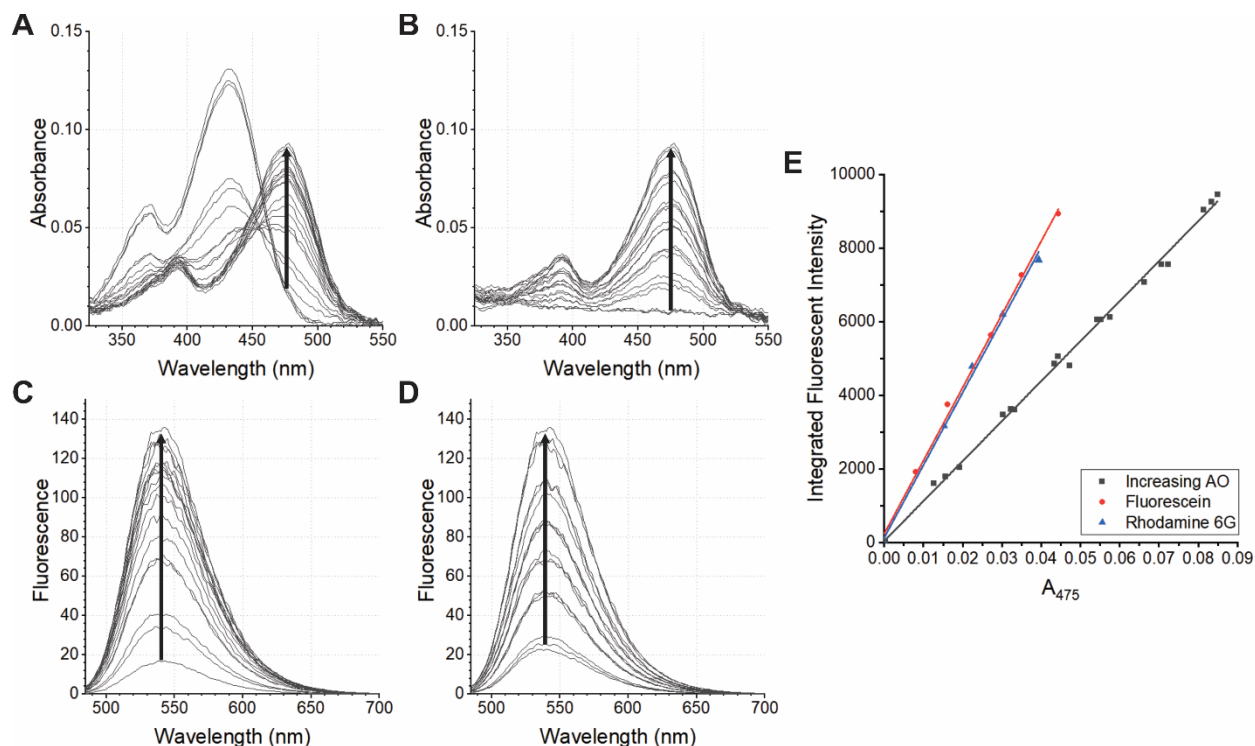
Figure 35: Quantum yield (φf) determination for the DAP-10-42:AO complex. Absorbance spectra for increasing concentrations (0-7.5 μM) of DAP-10-42 at 7.5 μM AO (A) and increasing concentrations (0-7.5 μM) of AO at 7.5 μM DAP-10-58 (B), as well as their respective fluorescence emission spectra upon excitation at 475 nm (C and D, respectively) were recorded in the assay buffer. Fluorescent spectra are shown after 7-point Savitzky-Golay smoothing. The slopes of integrated fluorescent intensity versus A_{475} for each set (E) were compared to two standards (fluorescein, rhodamine 6G) with literature φf values of 0.95 and experimental φf values of 0.934 and 0.966, respectively. The average φf compared to the two standards is 0.511.

143
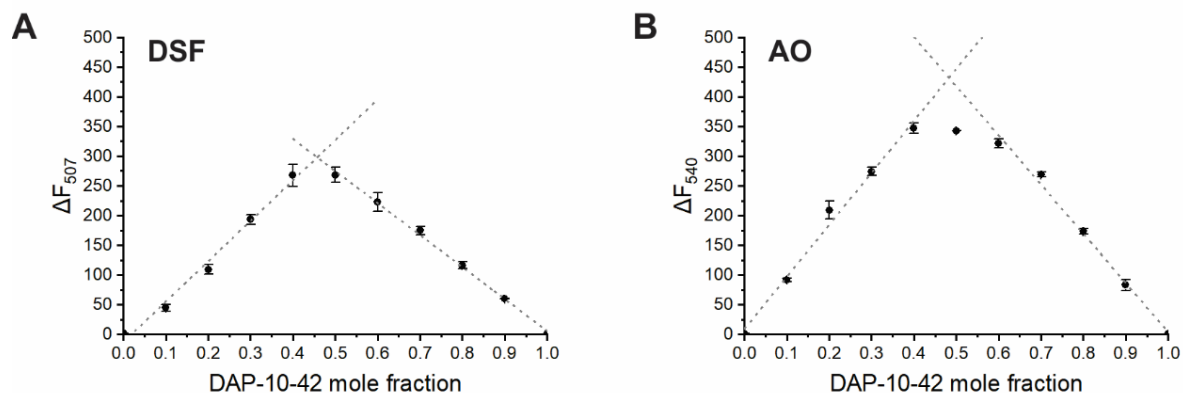
Figure 36: Continuous variations analysis (Job plot) for the complex of DAP-10-42 with DSA (A) or AO (B). The total concentration of the dye and aptamer was 4 µM. The DAP-10-42 molar fraction was calculated as the ratio between the aptamer concentration and the total (aptamer and dye) concentration. Average values from three independent experiments are shown with the error bars as standard deviations.
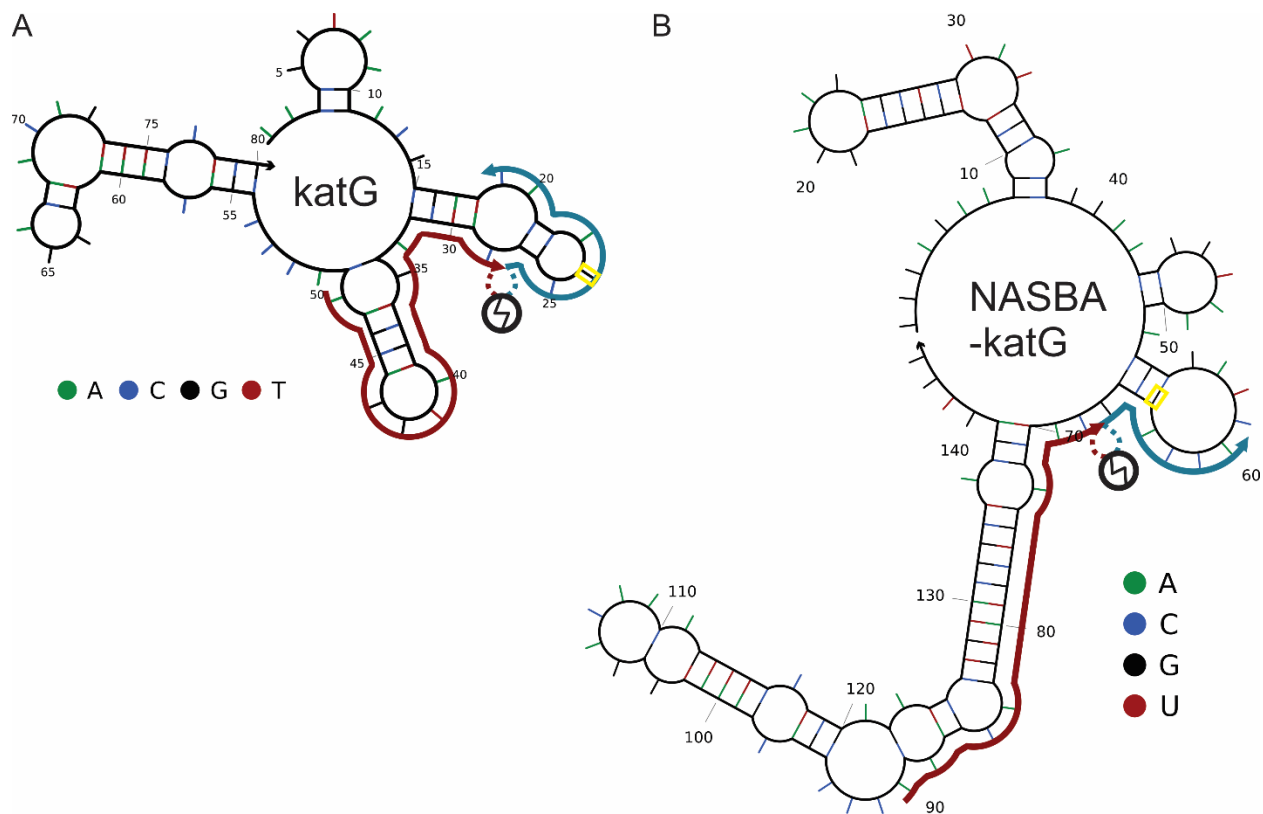
Figure 37: Minimum energy secondary structures of the synthetic target *katG* (A) and NASBA amplicon NASBA-*katG* (B) as predicted by NUPACK software (http://www.nupack.org/). The G315C mutation site is outlined with a yellow box. The nucleotide identity is color-coded. The target fragments interacting with SDA-S and SDA-U are indicated with cyan and red lines, respectively.
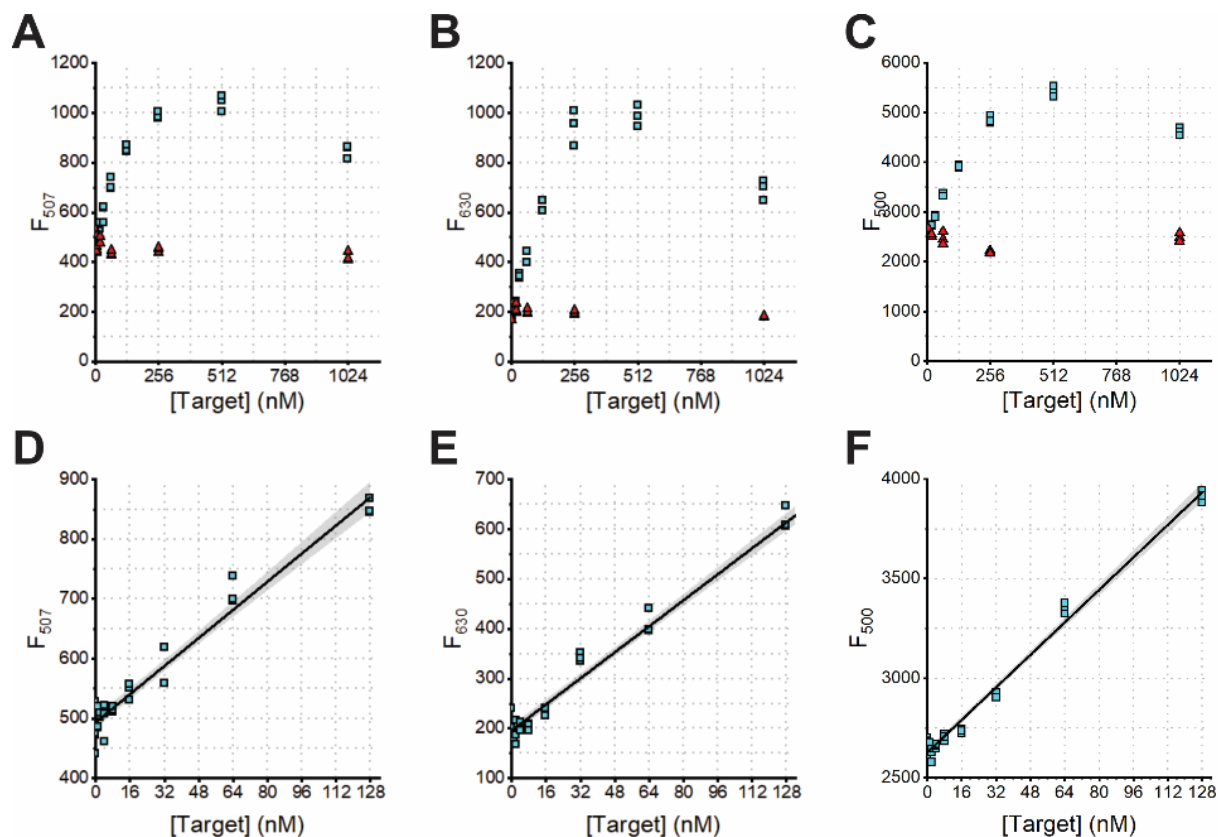
Figure 38: Response of the SDA probe utilizing DSF (A and D), CV (B and E), or TFT (C and F) to increasing concentrations (0-1024 nM) of *katG* (cyan squares) or *katG*-G>C (red triangles). Response to *katG* within the linear dynamic range (D-F) was used to calculate the limits of detection and quantification using the 3σ and 10σ rules, respectively. The calculated values are listed in Table 4. The line of best fit and 95% confidence interval are shown (black line, gray shading). For each graph shown, the fluorescence intensity values from three independent experiments were combined in one plot.
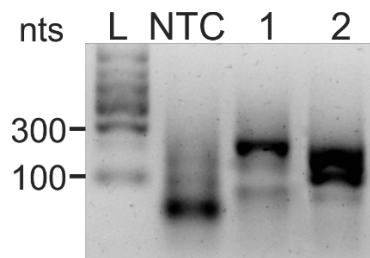
Figure 39: Analysis of NASBA samples in 2% agarose gel electrophoresis. NASBA no-target control (NTC) and two amplicon-containing samples (1 and 2) obtained in parallel by two different experimentalists using total RNA from *M. tuberculosis* (10 pg per reaction) were loaded on the gel in aliquots of 1 µl. L – RiboRuler Low Range RNA ladder (70 ng/µl per band, 0.5 µl per lane).

A Logical Array with Alphanumeric Fluorescent Output

Table 17: Oligonucleotides used in this study.*

Name	Sequence
INH^s	AAC CGG TAA GGA CGC GAT CAC CAG CGG CAT CGA GGT CGT ATG GAC GAA CAC CCC GAC GAA ATG GGA CAA CAG TTT CCT CG
INH^R	AAC CGG TAA GGA CGC GAT CAC CAC CGG CAT CGA GGT CGT ATG GAC GAA CAC CCC GAC GAA ATG GGA CAA CAG TTT CCT CG
SDA-U	<u>TGTTCGTCCATACGACCTCGATG</u> TT CTACG *GGGGAGGGTGTGTGGTTTT* GGTCAT
SDA-S	ATGACC *TTGGTT* CGTAG TT <u>CCGCTGGTG</u>
SDA-O	ATGACC *TTGGTT* CGTAG TT <u>CCGCTGGTGATCGCG</u>
SDA-N	ATGACC *TTGGTT* CGTAG TT GGCATCTTTTTGATGCC CATCGAGG

* single-nucleotide substitution in INH^R input is shown in cyan; nucleotides forming aptameric core fragments are in italics.

Experimental Section

Samples were prepared in the running buffer (20 mM Tris-HCl pH 7.4, 20 mM KCl, 25 mM MgCl$_2$, 0.1% Triton X-100) in the presence of 2 µM of either DSF or AO and 1 µM of SDA-U and SDA-S. Measurements were taken using an Agilent Cary Eclipse Fluorescent Spectrophotometer after 15 minutes of incubation at room temperature. Fluorometer settings for AO included excitation at 475 nm, emission measurements at 540 nm, and 5/5 Ex/Em slitwidths. DSF settings included excitation at 390 nm, emission at 507 nm, and 10/10 Ex/Em slitwidths.

APPENDIX E: SUPPORTING INFORMATION FOR CHAPTER FIVE

Experimental Section

Materials

Oligonucleotides were custom-made by IDT, Inc. (Coralville, IA, USA). Auramine O was purchased from MP Biomedicals (Solon, OH, USA). Thioflavin T was purchased from ACROS Organics (Morris, NJ, USA). N-methyl mesoporphyrin IX was purchased from Frontier Scientific (Logan, UT, USA). Quartz cuvettes for measuring absorbance, fluorescence, or circular dichroism were purchased from Starna Cells (Atascadero, CA, USA).

Methods

Fluorescent assays. Samples containing a dye and the aptamer or one of its mutants at the indicated concentrations were prepared in the assay buffer (20 mM Tris-HCl pH 7.4, 20 mM KCl, 25 mM $MgCl_2$) and incubated at 25 °C for 24h to reach equilibrium. Fluorescence measurements were performed using a Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies, Santa Clara, CA, USA) at ambient temperature using a 3 mm quartz cuvette (Starna Cells, Atascadero, CA, USA).

CD analysis. Samples (2 μM) were prepared in the assay buffer (20 mM Tris-HCl pH 7.4, 20 mM KCl, 25 mM $MgCl_2$) in the absence or presence of 4 μM of the respective dye without annealing and allowed to reach equilibrium at 22 °C over 48 hours in a dark environment before measurement. CD spectra were recorded in a 4 mm quartz cuvette using a JASCO J-810 spectropolarimeter (JASCO Inc., Easton, MD, USA), and data was obtained with a 2 nm slit width from 350 to 235 nm at 1 nm intervals. CD spectra were averaged over five scans. Spectra were normalized by subtracting the spectra of the blank and normalizing the baseline by subtracting the average reading for each sample in the 320-350 nm range. Difference G4 spectra were obtained by subtracting the spectra of a sample

containing only the equivalent stem-loop portion of the aptamer from the spectrum of the aptamer

(or one of its mutants).

Table 18: Oligonucleotides used in this study.*

Name	Sequence
DAP-10-42	CAATTACGGGGGAGGGTGTGTGGTCTTGCTTGGTTCGTATTG
DAP-10-58	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTCTTGCTTGGTTCGTATTG ATATTCGC
Core	GGGGAGGGTGTGTGGTCTTGCTTGGTT
StemA	GCGAATAT CAATTACG
StemB	CGTATTG ATATTCGC
8A,36T	GCGAATAT CAATTACAGGGGAGGGTGTGTGGTCTTGCTTGGTTTGTATTG ATATTCGC
8C,36G	GCGAATAT CAATTACCGGGGAGGGTGTGTGGTCTTGCTTGGTTGGTATTG ATATTCGC
9A	GCGAATAT CAATTACGAGGGAGGGTGTGTGGTCTTGCTTGGTTCGTATTG ATATTCGC
9I	GCGAATAT CAATTACGIGGGAGGGTGTGTGGTCTTGCTTGGTTCGTATTG ATATTCGC
10A	GCGAATAT CAATTACGGAGGAGGGTGTGTGGTCTTGCTTGGTTCGTATTG ATATTCGC
10I	GCGAATAT CAATTACGGIGGAGGGTGTGTGGTCTTGCTTGGTTCGTATTG ATATTCGC
11A	GCGAATAT CAATTACGGGAGAGGGTGTGTGGTCTTGCTTGGTTCGTATTG ATATTCGC
11I	GCGAATAT CAATTACGGGIGAGGGTGTGTGGTCTTGCTTGGTTCGTATTG ATATTCGC
12I	GCGAATAT CAATTACGGGGIAGGGTGTGTGGTCTTGCTTGGTTCGTATTG ATATTCGC
12T	GCGAATAT CAATTACGGGGTAGGGTGTGTGGTCTTGCTTGGTTCGTATTG ATATTCGC
14A	GCGAATAT CAATTACGGGGGAAGGTGTGTGGTCTTGCTTGGTTCGTATTG ATATTCGC
14I	GCGAATAT CAATTACGGGGGAIGGTGTGTGGTCTTGCTTGGTTCGTATTG ATATTCGC
15A	GCGAATAT CAATTACGGGGGAGAGTGTGTGGTCTTGCTTGGTTCGTATTG ATATTCGC
15I	GCGAATAT CAATTACGGGGGAGIGTGTGTGGTCTTGCTTGGTTCGTATTG ATATTCGC
16A	GCGAATAT CAATTACGGGGGAGGATGTGTGGTCTTGCTTGGTTCGTATTG ATATTCGC
16I	GCGAATAT CAATTACGGGGGAGGITGTGTGGTCTTGCTTGGTTCGTATTG ATATTCGC
18A	GCGAATAT CAATTACGGGGGAGGGTATGTGGTCTTGCTTGGTTCGTATTG ATATTCGC
18I	GCGAATAT CAATTACGGGGGAGGGTITGTGGTCTTGCTTGGTTCGTATTG ATATTCGC
18T	GCGAATAT CAATTACGGGGGAGGGTTTGTGGTCTTGCTTGGTTCGTATTG ATATTCGC
20A	GCGAATAT CAATTACGGGGGAGGGTGTATGGTCTTGCTTGGTTCGTATTG ATATTCGC
20I	GCGAATAT CAATTACGGGGGAGGGTGTITGGTCTTGCTTGGTTCGTATTG ATATTCGC
20T	GCGAATAT CAATTACGGGGGAGGGTGTTTGGTCTTGCTTGGTTCGTATTG ATATTCGC

22A	GCGAATAT CAATTACGGGGGAGGGTGTGTAGTCTTGCTTGGTTCGTATTG ATATTCGC
22I	GCGAATAT CAATTACGGGGGAGGGTGTGTIGTCTTGCTTGGTTCGTATTG ATATTCGC
23A	GCGAATAT CAATTACGGGGGAGGGTGTGTGATCTTGCTTGGTTCGTATTG ATATTCGC
23I	GCGAATAT CAATTACGGGGGAGGGTGTGTGITCTTGCTTGGTTCGTATTG ATATTCGC
24-31T10	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTTTTTTTTTTGGTTCGTATTG ATATTCGC
24-31T8	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTTTTTTTTGGTTCGTATTG ATATTCGC
24-31T6	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTTTTTTGGTTCGTATTG ATATTCGC
24-31T4	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTTTTGGTTCGTATTG ATATTCGC
28T,29T	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTCTTTTTTGGTTCGTATTG ATATTCGC
32A	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTCTTGCTTAGTTCGTATTG ATATTCGC
32I	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTCTTGCTTIGTTCGTATTG ATATTCGC
33A	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTCTTGCTTGATTCGTATTG ATATTCGC
33I	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTCTTGCTTGITTCGTATTG ATATTCGC
34A	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTCTTGCTTGGATCGTATTG ATATTCGC
34C	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTCTTGCTTGGCTCGTATTG ATATTCGC
34X	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTCTTGCTTGG_TCGTATTG ATATTCGC
35A	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTCTTGCTTGGTACGTATTG ATATTCGC
35C	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTCTTGCTTGGTCCGTATTG ATATTCGC
35TTT	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTCTTGCTTGGTTTTCGTATTG ATATTCGC
34X,35X	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTCTTGCTTGG__CGTATTG ATATTCGC
34X,35A	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTCTTGCTTGG_ACGTATTG ATATTCGC
34X,35C	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTCTTGCTTGG_CCGTATTG ATATTCGC
34A,35A	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTCTTGCTTGGAACGTATTG ATATTCGC
34C,35C	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTCTTGCTTGGCCCGTATTG ATATTCGC
34X,35TEG	GCGAATAT CAATTACGGGGGAGGGTGTGTGGTCTTGCTTGG/isp9/CGTATTG ATATTCGC

* Nucleotide substitutions / deletions are shown in cyan. Additional nucleotides added to stem are shown in red. /isp9/:

triethylene glycol linker

LIST OF REFERENCES

1. Organization, W. H. The Top 10 Causes of Death. https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death (accessed March 6 2022).

2. Aminov, R. I., A brief history of the antibiotic era: lessons learned and challenges for the future. *Front Microbiol* **2010,** *1*, 134.

3. Blair, J. M.; Webber, M. A.; Baylay, A. J.; Ogbolu, D. O.; Piddock, L. J., Molecular mechanisms of antibiotic resistance. *Nat Rev Microbiol* **2015,** *13* (1), 42-51.

4. Peterson, E.; Kaur, P., Antibiotic Resistance Mechanisms in Bacteria: Relationships Between Resistance Determinants of Antibiotic Producers, Environmental Bacteria, and Clinical Pathogens. *Front Microbiol* **2018,** *9*, 2928.

5. Zaman, S. B.; Hussain, M. A.; Nye, R.; Mehta, V.; Mamun, K. T.; Hossain, N., A Review on Antibiotic Resistance: Alarm Bells are Ringing. *Cureus* **2017,** *9* (6), e1403.

6. Prevention, C. f. D. C. a. 2019 AR Threats Report. https://www.cdc.gov/drugresistance/biggest-threats.html (accessed March 6 2022).

7. Gradmann, C., Sensitive Matters: The World Health Organisation and Antibiotic Resistance Testing, 1945–1975. *Social History of Medicine* **2013,** *26* (3), 555-574.

8. Khan, Z. A. S., M. F. Park, S., Current and Emerging Methods of Antibiotic Susceptibility Testing. *Diagnostics* **2019,** *9* (2).

9. Soini, H. M., J. M., Molecular Diagnosis of Mycobacteria. *Clin Chem* **2001,** *47* (5), 809-814.

10. Barken, K. B.; Haagensen, J. A.; Tolker-Nielsen, T., Advances in nucleic acid-based diagnostics of bacterial infections. *Clin Chim Acta* **2007,** *384* (1-2), 1-11.

11. Niemz, A.; Ferguson, T. M.; Boyle, D. S., Point-of-care nucleic acid testing for infectious diseases. *Trends Biotechnol* **2011,** *29* (5), 240-50.

12. Craw, P.; Balachandran, W., Isothermal nucleic acid amplification technologies for point-of-care diagnostics: a critical review. *Lab Chip* **2012,** *12* (14), 2469-86.

13. Scheler, O.; Glynn, B.; Kurg, A., Nucleic acid detection technologies and marker molecules in bacterial diagnostics. *Expert Rev Mol Diagn* **2014,** *14* (4), 489-500.

14. Kauppinen, S.; Vester, B.; Wengel, J., Locked nucleic acid (LNA): High affinity targeting of RNA for diagnostics and therapeutics. *Drug Discov Today Technol* **2005,** *2* (3), 287-90.

15. Wang, Q.; Chen, L.; Long, Y.; Tian, H.; Wu, J., Molecular Beacons of Xeno-Nucleic Acid for Detecting Nucleic Acid. *Theranostics* **2013,** *3* (6), 395-408.

16. Tyagi, S.; Kramer, F. R., Molecular Beacons: Probes that Fluoresce upon Hybridization. *Nature Biotechnology* **1996,** *14* (3), 303-308.

17. Goel, G.; Kumar, A.; Puniya, A. K.; Chen, W.; Singh, K., Molecular beacon: a multitask probe. *J Appl Microbiol* **2005,** *99* (3), 435-42.

18. Bidar, N.; Amini, M.; Oroojalian, F.; Baradaran, B.; Hosseini, S. S.; Shahbazi, M. A.; Hashemzaei, M.; Mokhtarzadeh, A.; Hamblin, M. R.; de la Guardia, M., Molecular beacon strategies for sensing purpose. *Trac-Trend Anal Chem* **2021,** *134*.

19. Holland, P. M.; Abramson, R. D.; Watson, R.; Gelfand, D. H., Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of Thermus aquaticus DNA polymerase. *Proc Natl Acad Sci U S A* **1991,** *88* (16), 7276-80.

20. Zanoli, L. M.; Spoto, G., Isothermal Amplification Methods for the Detection of Nucleic Acids in Microfluidic Devices. *Biosensors* **2013,** *3* (1), 18-43.

21. Demidov, V. V.; Frank-Kamenetskii, M. D., Two sides of the coin: affinity and specificity of nucleic acid interactions. *Trends Biochem Sci* **2004,** *29* (2), 62-71.

22. Kolpashchikov, D. M., Binary probes for nucleic acid analysis. *Chem Rev* **2010,** *110* (8), 4709-23.

23. Bezerra, A. B.; Kurian, A. S. N.; Easley, C. J., Nucleic-Acid Driven Cooperative Bioassays Using Probe Proximity or Split-Probe Techniques. *Anal Chem* **2021,** *93* (1), 198-214.

24. Ebata, K. M., M. Ohtani, H. Kashiwasake-Jibu, M., NUCLEIC ACID HYBRIDIZATION ACCOMPANIED WITH EXCIMER FORMATION FROM TWO PYRENE-LABELED PROBES. *Photochem Photobiol Sci* **1995,** *62* (5), 836-839.

25. Yusuke, K.; Toshihiro, I.; Yusuke, T.; Masato, T.; Akinori, J., DNA-templated Cooperative Formation of the Luminous Lanthanide Complex and Its Analytical Application to Gene Detection. *Chemistry Letters* **2005,** *34* (12), 1606-1607.

26. Gerasimova, Y. V.; Hayson, A.; Ballantyne, J.; Kolpashchikov, D. M., A single molecular beacon probe is sufficient for the analysis of multiple nucleic acid sequences. *Chembiochem* **2010,** *11* (12), 1762-8.

27. Park, J. A.; Kim, J.; Kim, S. M.; Sohn, H.; Park, C.; Kim, T. H.; Lee, J. H.; Lee, M. H.; Lee, T., Fabrication of Electrochemical Influenza Virus (H1N1) Biosensor Composed of Multifunctional DNA Four-Way Junction and Molybdenum Disulfide Hybrid Material. *Materials (Basel)* **2021,** *14* (2).

28. Wang, T.; Peng, Q.; Guo, B.; Zhang, D.; Zhao, M.; Que, H.; Wu, H.; Yan, Y., An integrated electrochemical biosensor based on target-triggered strand displacement amplification and "four-way" DNA junction towards ultrasensitive detection of PIK3CA gene mutation. *Biosens Bioelectron* **2020,** *150*, 111954.

29. Huang, S. Y.; Chang, J. R.; Liao, Y. C.; Dou, H. Y.; Chuang, M. C., Simultaneous detections of genetic fragment and single nucleotide mutation with a three-tiered output for tuberculosis diagnosis. *Anal Chim Acta* **2018,** *1007*, 1-9.

30. Megalathan, A.; Wijesinghe, K. M.; Dhakal, S., Single-Molecule FRET-Based Dynamic DNA Sensor. *ACS Sens* **2021,** *6* (3), 1367-1374.

31. Lynch, C. A., 3rd; Foguel, M. V.; Reed, A. J.; Balcarcel, A. M.; Calvo-Marzal, P.; Gerasimova, Y. V.; Chumbimuni-Torres, K. Y., Selective Determination of Isothermally Amplified Zika Virus RNA Using a Universal DNA-Hairpin Probe in Less than 1 Hour. *Anal Chem* **2019,** *91* (21), 13458-13464.

32. Foguel, M. V. B., A. M. Reed, M. A. Calvo-Marzal, P. Gerasimova, Y. V. Kolpashchikov, D. M. Chumbimuni-Torres, K. Y., MVF Sensor Enables Analysis of Nucleic Acids with Stable Secondary Structures. *Electroanalysis* **2020,** *32* (4), 835-841.

33. Mills, D. M. F., M. V. Martin, C. P. Trieu, T. T. Kamar, O. Calvo-Marzal, P. Kolpashchikov, D. M. Chumbimuni-Torres, K. Y., Rapid detection of different DNA analytes using a single electrochemical sensor. *Sens Actuators B Chem* **2019,** *293*, 11-15.

34. Song, S. W., L. Lia, J. Fana, C. Zhaob, J., Aptamer-based biosensors. *TrAC Trends in Analytical Chemistry* **2008,** *27* (2), 108-117.

35. Ellington, A. D.; Szostak, J. W., In vitro selection of RNA molecules that bind specific ligands. *Nature* **1990,** *346* (6287), 818-22.

36. Tuerk, C.; Gold, L., Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **1990,** *249* (4968), 505-10.

37. Wilson, D. S.; Szostak, J. W., In vitro selection of functional nucleic acids. *Annu Rev Biochem* **1999,** *68*, 611-47.

38. Darmostuk, M.; Rimpelova, S.; Gbelcova, H.; Ruml, T., Current approaches in SELEX: An update to aptamer selection technology. *Biotechnol Adv* **2015,** *33* (6 Pt 2), 1141-61.

39. Iliuk, A. B. H., L. Tao, W. A., Aptamer in Bioanalytical Applications. *Anal Chem* **2011,** *83* (12), 4440-4452.

40. Gao, T.; Luo, Y.; Li, W.; Cao, Y.; Pei, R., Progress in the isolation of aptamers to light-up the dyes and the applications. *Analyst* **2020,** *145* (3), 701-718.

41. Klymchenko, A. S., Solvatochromic and Fluorogenic Dyes as Environment-Sensitive Probes: Design and Biological Applications. *Acc Chem Res* **2017,** *50* (2), 366-375.

42. Chen, Y.; Wang, J.; Zhang, Y.; Xu, L.; Gao, T.; Wang, B.; Pei, R., Selection and characterization of a DNA aptamer to crystal violet. *Photochem Photobiol Sci* **2018,** *17* (6), 800-806.

43. Sun, C. Y.; Su, R. F.; Bie, J. X.; Sun, H. J.; Qiao, S. N.; Ma, X. Y.; Sun, R.; Zhang, T. H., Label-free fluorescent sensor based on aptamer and thiazole orange for the detection of tetracycline. *Dyes and Pigments* **2018,** *149*, 867-875.

44. Wang , H.; Wang, J.; Sun, N.; Cheng , H.; Chen, H.; Pei, R., Selection and Characterization of Malachite Green Aptamers for the Development of Light-up Probes. *ChemistrySelect* **2016,** *1* (8), 1571-1574.

45. Wang, J. E.; Zhang, Y. J.; Wang, H. Y.; Chen, Y.; Xu, L. J.; Gao, T.; Pei, R. J., Selection and analysis of DNA aptamers to berberine to develop a label-free light-up fluorescent probe. *New Journal of Chemistry* **2016,** *40* (11), 9768-9773.

46. Wang, H.; Wang, J.; Wang, Q.; Chen, X.; Liu, M.; Chen, H.; Pei, R., Selection and characterization of dimethylindole red DNA aptamers for the development of light-up fluorescent probes. *Talanta* **2017,** *168*, 217-221.

47. Kato, T.; Shimada, I.; Kimura, R.; Hyuga, M., Light-up fluorophore--DNA aptamer pair for label-free turn-on aptamer sensors. *Chem Commun (Camb)* **2016,** *52* (21), 4041-4.

48. Sando, S.; Narita, A.; Aoyama, Y., Light-up Hoechst-DNA aptamer pair: generation of an aptamer-selective fluorophore from a conventional DNA-staining dye. *Chembiochem* **2007,** *8* (15), 1795-803.

49. Paige, J. S.; Wu, K. Y.; Jaffrey, S. R., RNA mimics of green fluorescent protein. *Science* **2011,** *333* (6042), 642-6.

50. Dolgosheina, E. V.; Jeng, S. C.; Panchapakesan, S. S.; Cojocaru, R.; Chen, P. S.; Wilson, P. D.; Hawkins, N.; Wiggins, P. A.; Unrau, P. J., RNA mango aptamer-fluorophore: a bright, high-affinity complex for RNA labeling and tracking. *ACS Chem Biol* **2014,** *9* (10), 2412-20.

51. Constantin, T. P.; Silva, G. L.; Robertson, K. L.; Hamilton, T. P.; Fague, K.; Waggoner, A. S.; Armitage, B. A., Synthesis of new fluorogenic cyanine dyes and incorporation into RNA fluoromodules. *Org Lett* **2008,** *10* (8), 1561-4.

52. Strack, R. L.; Disney, M. D.; Jaffrey, S. R., A superfolding Spinach2 reveals the dynamic nature of trinucleotide repeat-containing RNA. *Nat Methods* **2013,** *10* (12), 1219-24.

53. Song, W.; Strack, R. L.; Svensen, N.; Jaffrey, S. R., Plug-and-play fluorophores extend the spectral properties of Spinach. *J Am Chem Soc* **2014,** *136* (4), 1198-201.

54. Filonov, G. S.; Moon, J. D.; Svensen, N.; Jaffrey, S. R., Broccoli: rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution. *J Am Chem Soc* **2014,** *136* (46), 16299-308.

55. Gotrik, M.; Sekhon, G.; Saurabh, S.; Nakamoto, M.; Eisenstein, M.; Soh, H. T., Direct Selection of Fluorescence-Enhancing RNA Aptamers. *J Am Chem Soc* **2018,** *140* (10), 3583-3591.

56. Yang, J.; Bowser, M. T., Capillary electrophoresis-SELEX selection of catalytic DNA aptamers for a small-molecule porphyrin target. *Anal Chem* **2013,** *85* (3), 1525-30.

57. Tucker, W. O.; Shum, K. T.; Tanner, J. A., G-quadruplex DNA aptamers and their ligands: structure, function and application. *Curr Pharm Des* **2012,** *18* (14), 2014-26.

58. Roxo, C.; Kotkowiak, W.; Pasternak, A., G-Quadruplex-Forming Aptamers-Characteristics, Applications, and Perspectives. *Molecules* **2019,** *24* (20).

59. Platella, C.; Riccardi, C.; Montesarchio, D.; Roviello, G. N.; Musumeci, D., G-quadruplex-based aptamers against protein targets in therapy and diagnostics. *Biochim Biophys Acta Gen Subj* **2017,** *1861* (5 Pt B), 1429-1447.

60. Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S., Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res* **2006,** *34* (19), 5402-15.

61. Balagurumoorthy, P.; Brahmachari, S. K., Structure and stability of human telomeric sequence. *J Biol Chem* **1994,** *269* (34), 21858-69.

62. Bhattacharyya, D.; Mirihana Arachchilage, G.; Basu, S., Metal Cations in G-Quadruplex Folding and Stability. *Front Chem* **2016,** *4*, 38.

63. Umar, M. I.; Ji, D.; Chan, C. Y.; Kwok, C. K., G-Quadruplex-Based Fluorescent Turn-On Ligands and Aptamers: From Development to Applications. *Molecules* **2019,** *24* (13).

64. Nicoludis, J. M.; Barrett, S. P.; Mergny, J. L.; Yatsunyk, L. A., Interaction of human telomeric DNA with N-methyl mesoporphyrin IX. *Nucleic Acids Res* **2012,** *40* (12), 5432-47.

65. Tippana, R.; Xiao, W.; Myong, S., G-quadruplex conformation and dynamics are determined by loop length and sequence. *Nucleic Acids Res* **2014,** *42* (12), 8106-14.

66. Kong, D. M.; Ma, Y. E.; Guo, J. H.; Yang, W.; Shen, H. X., Fluorescent sensor for monitoring structural changes of G-quadruplexes and detection of potassium ion. *Anal Chem* **2009,** *81* (7), 2678-84.

67. Ren, J.; Wang, J.; Wang, J.; Luedtke, N. W.; Wang, E., Contribution of potassium ion and split modes of G-quadruplex to the sensitivity and selectivity of label-free sensor toward DNA detection using fluorescence. *Biosens Bioelectron* **2012,** *31* (1), 316-22.

68. Zhu, J.; Zhang, L.; Wang, E., Measurement of the base number of DNA using a special calliper made of a split G-quadruplex. *Chem Commun (Camb)* **2012,** *48* (98), 11990-2.

69. Yett, A.; Lin, L. Y.; Beseiso, D.; Miao, J.; Yatsunyk, L. A., N-methyl mesoporphyrin IX as a highly selective light-up probe for G-quadruplex DNA. *J Porphyr Phthalocyanines* **2019,** *23* (11n12), 1195-1215.

70. Zhang, X. F.; Li, N.; Ling, Y.; Tang, L.; Li, N. B.; Luo, H. Q., Linked bridge hybridizing-induced split G-quadruplex DNA machine and its application to uracil-DNA glycosylase detection. *Sensors and Actuators B-Chemical* **2018,** *255*, 2589-2594.

71. Kolpashchikov, D. M., Binary malachite green aptamer for fluorescent detection of nucleic acids. *J Am Chem Soc* **2005,** *127* (36), 12442-3.

72. Kikuchi, N.; Kolpashchikov, D. M., Split Spinach Aptamer for Highly Selective Recognition of DNA and RNA at Ambient Temperatures. *Chembiochem* **2016,** *17* (17), 1589-92.

73. Kikuchi, N.; Kolpashchikov, D. M., A universal split spinach aptamer (USSA) for nucleic acid analysis and DNA computation. *Chem Commun (Camb)* **2017,** *53* (36), 4977-4980.

74. Kikuchi, N.; Reed, A.; Gerasimova, Y. V.; Kolpashchikov, D. M., Split Dapoxyl Aptamer for Sequence-Selective Analysis of Nucleic Acid Sequence Based Amplification Amplicons. *Anal Chem* **2019,** *91* (4), 2667-2671.

75. Cheng, X.; Liu, X.; Bing, T.; Cao, Z.; Shangguan, D., General peroxidase activity of G-quadruplex-hemin complexes and its application in ligand screening. *Biochemistry* **2009,** *48* (33), 7817-23.

76. Deng, M.; Zhang, D.; Zhou, Y.; Zhou, X., Highly effective colorimetric and visual detection of nucleic acids using an asymmetrically split peroxidase DNAzyme. *J Am Chem Soc* **2008,** *130* (39), 13095-102.

77. Kolpashchikov, D. M., Split DNA enzyme for visual single nucleotide polymorphism typing. *J Am Chem Soc* **2008,** *130* (10), 2934-5.

78. Roembke, B. T.; Nakayama, S.; Sintim, H. O., Nucleic acid detection using G-quadruplex amplification methodologies. *Methods* **2013,** *64* (3), 185-98.

79. Li, T. L., B. Wang, E. Dong, S., G-quadruplex-based DNAzyme for sensitive mercury detection with the naked eye. *Chem Commun* **2009**, 3551-3553.

80. Thirstrup, D. B., G. S., Histochemical Application Of A Peroxidase DNAzyme with a Covalently Attached Hemin Cofactor. *Anal Chem* **2010,** *82* (6), 2498-2504.

81. Liang, G. M., Y. Li, A. Jin, X. Pan, L. Liu, X., Chemiluminescence assay for detection of 2-hydroxyfluorene using the G-quadruplex DNAzyme-H2O2-luminol system. *Microchimica Acta* **2018,** *185*.

82. de Silva, P. A.; Gunaratne, N. H. Q.; McCoy, C. P., A molecular photoionic AND gate based on fluorescent signalling. *Nature* **1993,** *364* (6432), 42-44.

83. Cornett, E. M.; Campbell, E. A.; Gulenay, G.; Peterson, E.; Bhaskar, N.; Kolpashchikov, D. M., Molecular logic gates for DNA analysis: detection of rifampin resistance in M. tuberculosis DNA. *Angew Chem Int Ed Engl* **2012,** *51* (36), 9075-7.

84. Poje, J. E.; Kastratovic, T.; Macdonald, A. R.; Guillermo, A. C.; Troetti, S. E.; Jabado, O. J.; Fanning, M. L.; Stefanovic, D.; Macdonald, J., Visual Displays that Directly Interface and Provide Read-Outs of Molecular States via Molecular Graphics Processing Units. *Angew Chem Int Edit* **2014,** *53* (35), 9222-9225.

85. Lai, Y. H.; Sun, S. C.; Chuang, M. C., Biosensors with built-in biomolecular logic gates for practical applications. *Biosensors (Basel)* **2014,** *4* (3), 273-300.

86. Gerasimova, Y. V.; Kolpashchikov, D. M., Divide and control: split design of multi-input DNA logic gates. *Chem Commun (Camb)* **2015,** *51* (5), 870-2.

87. Chen, Y.; Song, Y.; Wu, F.; Liu, W.; Fu, B.; Feng, B.; Zhou, X., A DNA logic gate based on strand displacement reaction and rolling circle amplification, responding to multiple low-abundance DNA fragment input signals, and its application in detecting miRNAs. *Chem Commun (Camb)* **2015,** *51* (32), 6980-3.

88. Bi, S.; Ye, J.; Dong, Y.; Li, H.; Cao, W., Target-triggered cascade recycling amplification for label-free detection of microRNA and molecular logic operations. *Chem Commun (Camb)* **2016,** *52* (2), 402-5.

89. Ge, L.; Wang, W.; Sun, X.; Hou, T.; Li, F., Versatile and Programmable DNA Logic Gates on Universal and Label-Free Homogeneous Electrochemical Platform. *Anal Chem* **2016,** *88* (19), 9691-9698.

90. Dobrin, A.; Saxena, P.; Fussenegger, M., Synthetic biology: applying biological circuits beyond novel therapies. *Integr Biol (Camb)* **2016,** *8* (4), 409-30.

91. Vijayakumar, P.; Macdonald, J., A DNA Logic Gate Automaton for Detection of Rabies and Other Lyssaviruses. *Chemphyschem* **2017,** *18* (13), 1735-1741.

92. E. E. May, J. C. H., S. M. Brozik *Modeling, Methodologies and Tools for Molecular and Nano-scale Communications.* Springer: 2017.

93. Ahmad, S.; Mokaddas, E., Recent advances in the diagnosis and treatment of multidrug-resistant tuberculosis. *Respiratory Medicine CME* **2010,** *3* (2), 51-61.

94. Chiang, C. Y.; Centis, R.; Migliori, G. B., Drug-resistant tuberculosis: past, present, future. *Respirology* **2010,** *15* (3), 413-32.

95. Musser, J. M., Antimicrobial agent resistance in mycobacteria: molecular genetic insights. *Clin Microbiol Rev* **1995,** *8* (4), 496-514.

96. Almeida Da Silva, P. E.; Palomino, J. C., Molecular basis and mechanisms of drug resistance in Mycobacterium tuberculosis: classical and new drugs. *J Antimicrob Chemother* **2011,** *66* (7), 1417-30.

97. Palomino, J. C.; Martin, A., Drug Resistance Mechanisms in Mycobacterium tuberculosis. *Antibiotics (Basel)* **2014,** *3* (3), 317-40.

98. Pavlov, V.; Xiao, Y.; Gill, R.; Dishon, A.; Kotler, M.; Willner, I., Amplified chemiluminescence surface detection of DNA and telomerase activity using catalytic nucleic acid labels. *Anal Chem* **2004,** *76* (7), 2152-6.

99. Li, T.; Wang, E.; Dong, S., Potassium-lead-switched G-quadruplexes: a new class of DNA logic gates. *J Am Chem Soc* **2009,** *131* (42), 15082-3.

100. Kong, D. M.; Cai, L. L.; Guo, J. H.; Wu, J.; Shen, H. X., Characterization of the G-quadruplex structure of a catalytic DNA with peroxidase activity. *Biopolymers* **2009,** *91* (5), 331-9.

101. Nakayama, S.; Wang, J.; Sintim, H. O., DNA-based peroxidation catalyst--what is the exact role of topology on catalysis and is there a special binding site for catalysis? *Chemistry* **2011,** *17* (20), 5691-8.

102. Darius, A. K.; Ling, N. J.; Mahesh, U., Visual detection of DNA from salmonella and mycobacterium using split DNAzymes. *Mol Biosyst* **2010,** *6* (5), 792-4.

103. Neo, J. L.; Aw, K. D.; Uttamchandani, M., Visual SNP genotyping using asymmetric PCR and split DNA enzymes. *Analyst* **2011,** *136* (8), 1569-72.

104. Hooyberghs, J.; Carlon, E., Hybridisation thermodynamic parameters allow accurate detection of point mutations with DNA microarrays. *Biosens Bioelectron* **2010,** *26* (4), 1692-5.

105. Grimes, J.; Gerasimova, Y. V.; Kolpashchikov, D. M., Real-Time SNP Analysis in Secondary-Structure-Folded Nucleic Acids. *Angewandte Chemie* **2010,** *122* (47), 9134-9137.

106. Deng, M.; Feng, S.; Luo, F.; Wang, S.; Sun, X.; Zhou, X.; Zhang, X. L., Visual detection of rpoB mutations in rifampin-resistant Mycobacterium tuberculosis strains by use of an asymmetrically split peroxidase DNAzyme. *J Clin Microbiol* **2012,** *50* (11), 3443-50.

107. Zhu, J.; Zhang, L.; Dong, S.; Wang, E., How to split a G-quadruplex for DNA detection: new insight into the formation of DNA split G-quadruplex. *Chem Sci* **2015,** *6* (8), 4822-4827.

108. Lu, X.; Shi, X.; Wu, G.; Wu, T.; Qin, R.; Wang, Y., Visual detection and differentiation of Classic Swine Fever Virus strains using nucleic acid sequence-based amplification (NASBA) and G-quadruplex DNAzyme assay. *Sci Rep* **2017,** *7*, 44211.

109. Ramaswamy, S.; Musser, J. M., Molecular genetic basis of antimicrobial agent resistance in Mycobacterium tuberculosis: 1998 update. *Tuber Lung Dis* **1998,** *79* (1), 3-29.

110. Caws, M.; Duy, P. M.; Tho, D. Q.; Lan, N. T.; Hoa, D. V.; Farrar, J., Mutations prevalent among rifampin- and isoniazid-resistant Mycobacterium tuberculosis isolates from a hospital in Vietnam. *J Clin Microbiol* **2006,** *44* (7), 2333-7.

111. Zadeh, J. N.; Steenberg, C. D.; Bois, J. S.; Wolfe, B. R.; Pierce, M. B.; Khan, A. R.; Dirks, R. M.; Pierce, N. A., NUPACK: Analysis and design of nucleic acid systems. *J Comput Chem* **2011,** *32* (1), 170-3.

112. Lv, M. G., Y. Ren, J. Wang, E., Exploration of intramolecular split G-quadruplex and its analytical applications. *Nucleic Acids Res* **2019,** *47* (18), 9502-9510.

113. SantaLucia, J., Jr., A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc Natl Acad Sci U S A* **1998,** *95* (4), 1460-5.

114. Urakawa, H.; El Fantroussi, S.; Smidt, H.; Smoot, J. C.; Tribou, E. H.; Kelly, J. J.; Noble, P. A.; Stahl, D. A., Optimization of single-base-pair mismatch discrimination in oligonucleotide microarrays. *Appl Environ Microbiol* **2003,** *69* (5), 2848-56.

115. Jaksik, R.; Iwanaszko, M.; Rzeszowska-Wolny, J.; Kimmel, M., Microarray experiments and factors which affect their reliability. *Biol Direct* **2015,** *10*, 46.

116. Tyagi, S.; Kramer, F. R., Molecular beacons: probes that fluoresce upon hybridization. *Nat Biotechnol* **1996,** *14* (3), 303-8.

117. Tsourkas, A.; Behlke, M. A.; Bao, G., Structure-function relationships of shared-stem and conventional molecular beacons. *Nucleic Acids Res* **2002,** *30* (19), 4208-15.

118. Tsourkas, A.; Behlke, M. A.; Rose, S. D.; Bao, G., Hybridization kinetics and thermodynamics of molecular beacons. *Nucleic Acids Res* **2003,** *31* (4), 1319-30.

119. Bonnet, G.; Tyagi, S.; Libchaber, A.; Kramer, F. R., Thermodynamic basis of the enhanced specificity of structured DNA probes. *Proc Natl Acad Sci U S A* **1999,** *96* (11), 6171-6.

120. Nguyen, C.; Grimes, J.; Gerasimova, Y. V.; Kolpashchikov, D. M., Molecular-beacon-based tricomponent probe for SNP analysis in folded nucleic acids. *Chemistry* **2011,** *17* (46), 13052-8.

121. Gerasimova, Y. V.; Kolpashchikov, D. M., Detection of bacterial 16S rRNA using a molecular beacon-based X sensor. *Biosens Bioelectron* **2013,** *41*, 386-90.

122. Cornett, E. M.; O'Steen, M. R.; Kolpashchikov, D. M., Operating Cooperatively (OC) sensor for highly specific recognition of nucleic acids. *PLoS One* **2013,** *8* (2), e55919.

123. Stancescu, M.; Fedotova, T. A.; Hooyberghs, J.; Balaeff, A.; Kolpashchikov, D. M., Nonequilibrium Hybridization Enables Discrimination of a Point Mutation within 5-40 degrees C. *J Am Chem Soc* **2016,** *138* (41), 13465-13468.

124. Gerasimova, Y. V.; Cornett, E.; Kolpashchikov, D. M., RNA-cleaving deoxyribozyme sensor for nucleic acid analysis: the limit of detection. *Chembiochem* **2010,** *11* (6), 811-7, 729.

125. Schaaf, C. P.; Wiszniewska, J.; Beaudet, A. L., Copy number and SNP arrays in clinical diagnostics. *Annu Rev Genomics Hum Genet* **2011,** *12*, 25-51.

126. Lieben, L., Disease genetics: SNP location helps predict disease aetiology. *Nat Rev Genet* **2016,** *17* (1), 4.

127. Laing, R. E.; Hess, P.; Shen, Y.; Wang, J.; Hu, S. X., The role and impact of SNPs in pharmacogenomics and personalized medicine. *Curr Drug Metab* **2011,** *12* (5), 460-86.

128. Nielsen, R., Population genetic analysis of ascertained SNP data. *Hum Genomics* **2004,** *1* (3), 218-24.

129. Oldoni, F.; Podini, D., Forensic molecular biomarkers for mixture analysis. *Forensic Sci Int Genet* **2019,** *41*, 107-119.

130. Black, W. C. t.; Vontas, J. G., Affordable assays for genotyping single nucleotide polymorphisms in insects. *Insect Mol Biol* **2007,** *16* (4), 377-87.

131. Cirillo, D. M.; Miotto, P.; Tortoli, E., Evolution of Phenotypic and Molecular Drug Susceptibility Testing. In *Strain Variation in the Mycobacterium tuberculosis Complex: Its Role in Biology, Epidemiology and Control*, Gagneux, S., Ed. Springer International Publishing: Cham, 2017; pp 221-246.

132. Yoo, S. M.; Lee, S. Y., Optical Biosensors for the Detection of Pathogenic Microorganisms. *Trends Biotechnol* **2016,** *34* (1), 7-25.

133. Kosman, J.; Juskowiak, B., Peroxidase-mimicking DNAzymes for biosensing applications: a review. *Anal Chim Acta* **2011,** *707* (1-2), 7-17.

134. Neo, J. L.; Kamaladasan, K.; Uttamchandani, M., G-quadruplex based probes for visual detection and sensing. *Curr Pharm Des* **2012,** *18* (14), 2048-57.

135. Lv, L.; Guo, Z.; Wang, J.; Wang, E., G-quadruplex as signal transducer for biorecognition events. *Curr Pharm Des* **2012,** *18* (14), 2076-95.

136. Ruttkay-Nedecky, B.; Kudr, J.; Nejdl, L.; Maskova, D.; Kizek, R.; Adam, V., G-quadruplexes as sensing probes. *Molecules* **2013,** *18* (12), 14760-79.

137. Ren, J.; Wang, T.; Wang, E.; Wang, J., Versatile G-quadruplex-mediated strategies in label-free biosensors and logic systems. *Analyst* **2015,** *140* (8), 2556-72.

138. Tian, T.; Xiao, H.; Zhou, X., A Review: G-Quadruplex's Applications in Biological Target Detection and Drug Delivery. *Current Topics in Medicinal Chemistry* **2015,** *15* (19), 1988-2001.

139. Batule, B. S.; Kim, S. U.; Mun, H.; Choi, C.; Shim, W. B.; Kim, M. G., Colorimetric Detection of Norovirus in Oyster Samples through DNAzyme as a Signaling Probe. *J Agric Food Chem* **2018,** *66* (11), 3003-3008.

140. Li, T.; Dong, S.; Wang, E., Enhanced catalytic DNAzyme for label-free colorimetric detection of DNA. *Chem Commun (Camb)* **2007,** (41), 4209-11.

141. Jiang, X.; Zhang, H.; Wu, J.; Yang, X.; Shao, J.; Lu, Y.; Qiu, B.; Lin, Z.; Chen, G., G-quadruplex DNA biosensor for sensitive visible detection of genetically modified food. *Talanta* **2014,** *128*, 445-9.

142. Connelly, R. P.; Morozkin, E. S.; Gerasimova, Y. V., Alphanumerical Visual Display Made of DNA Logic Gates for Drug Susceptibility Testing of Pathogens. *Chembiochem* **2018,** *19* (3), 203-206.

143. Travascio, P.; Li, Y.; Sen, D., DNA-enhanced peroxidase activity of a DNA aptamer-hemin complex. *Chemistry & Biology* **1998,** *5* (9), 505-517.

144. Travascio, P.; Bennet, A. J.; Wang, D. Y.; Sen, D., A ribozyme and a catalytic DNA with peroxidase activity: active sites versus cofactor-binding sites. *Chemistry & Biology* **1999,** *6* (11), 779-787.

145. Witting, P. K.; Travascio, P.; Sen, D.; Mauk, A. G., A DNA oligonucleotide-hemin complex cleaves t-butyl hydroperoxide through a homolytic mechanism. *Inorg Chem* **2001,** *40* (19), 5017-23.

146. Nakayama, S.; Sintim, H. O., Colorimetric split G-quadruplex probes for nucleic acid sensing: improving reconstituted DNAzyme's catalytic efficiency via probe remodeling. *J Am Chem Soc* **2009,** *131* (29), 10320-33.

147. Bohlin, J.; Eldholm, V.; Pettersson, J. H.; Brynildsrud, O.; Snipen, L., The nucleotide composition of microbial genomes indicates differential patterns of selection on core and accessory genomes. *BMC Genomics* **2017,** *18* (1), 151.

148. Lee, M. R.; Sheng, W. H.; Hung, C. C.; Yu, C. J.; Lee, L. N.; Hsueh, P. R., Mycobacterium abscessus Complex Infections in Humans. *Emerg Infect Dis* **2015,** *21* (9), 1638-46.

149. Stefan, M. A.; Ugur, F. S.; Garcia, G. A., Source of the Fitness Defect in Rifamycin-Resistant Mycobacterium tuberculosis RNA Polymerase and the Mechanism of Compensation by Mutations in the beta' Subunit. *Antimicrob Agents Chemother* **2018,** *62* (6).

150. Tseng, S. T.; Tai, C. H.; Li, C. R.; Lin, C. F.; Shi, Z. Y., The mutations of katG and inhA genes of isoniazid-resistant Mycobacterium tuberculosis isolates in Taiwan. *J Microbiol Immunol Infect* **2015,** *48* (3), 249-55.

151. Suwannakarn, K.; Payungporn, S.; Chieochansin, T.; Samransamruajkit, R.; Amonsin, A.; Songserm, T.; Chaisingh, A.; Chamnanpood, P.; Chutinimitkul, S.; Theamboonlers, A.; Poovorawan, Y., Typing (A/B) and subtyping (H1/H3/H5) of influenza A viruses by multiplex real-time RT-PCR assays. *J Virol Methods* **2008,** *152* (1-2), 25-31.

152. Influenza Type A Viruses. https://www.cdc.gov/flu/avianflu/influenza-a-virus-subtypes.htm (accessed August 2).

153. Wolfenden, B. S.; Willson, R. L., Radical-cations as reference chromogens in kinetic studies of ono-electron transfer reactions: pulse radiolysis studies of 2,2'-azinobis-(3-ethylbenzthiazoline-6-sulphonate). *J. Chem. Soc., Perkin Trans. 2* **1982,** (7), 805-812.

154. Kolpashchikov, D. M., A binary DNA probe for highly specific nucleic Acid recognition. *J Am Chem Soc* **2006,** *128* (32), 10625-8.

155. Kolpashchikov, D. M., A binary deoxyribozyme for nucleic acid analysis. *Chembiochem* **2007,** *8* (17), 2039-42.

156. Xiao, Y.; Pavlov, V.; Niazov, T.; Dishon, A.; Kotler, M.; Willner, I., Catalytic beacons for the detection of DNA and telomerase activity. *J Am Chem Soc* **2004,** *126* (24), 7430-1.

157. Luo, Y.; Yu, H.; Alkhamis, O.; Liu, Y.; Lou, X.; Yu, B.; Xiao, Y., Label-Free, Visual Detection of Small Molecules Using Highly Target-Responsive Multimodule Split Aptamer Constructs. *Anal Chem* **2019,** *91* (11), 7199-7207.

158. Hou, T.; Li, C.; Wang, X.; Zhao, C.; Li, F., Label-free colorimetric detection of coralyne utilizing peroxidase-like split G-quadruplex DNAzyme. *Analytical Methods* **2013,** *5* (18).

159. Cheng, M.; Zhou, J.; Jia, G.; Ai, X.; Mergny, J. L.; Li, C., Relations between the loop transposition of DNA G-quadruplex and the catalytic function of DNAzyme. *Biochim Biophys Acta Gen Subj* **2017,** *1861* (8), 1913-1920.

160. Karsisiotis, A. I.; Hessari, N. M.; Novellino, E.; Spada, G. P.; Randazzo, A.; Webba da Silva, M., Topological characterization of nucleic acid G-quadruplexes by UV absorption and circular dichroism. *Angew Chem Int Ed Engl* **2011,** *50* (45), 10645-8.

161. Li, W.; Li, Y.; Liu, Z.; Lin, B.; Yi, H.; Xu, F.; Nie, Z.; Yao, S., Insight into G-quadruplex-hemin DNAzyme/RNAzyme: adjacent adenine as the intramolecular species for remarkable enhancement of enzymatic activity. *Nucleic Acids Res* **2016,** *44* (15), 7373-84.

162. Chen, J.; Guo, Y.; Zhou, J.; Ju, H., The Effect of Adenine Repeats on G-quadruplex/hemin Peroxidase Mimicking DNAzyme Activity. *Chemistry* **2017,** *23* (17), 4210-4215.

163. Chen, J.; Zhang, Y.; Cheng, M.; Guo, Y.; Šponer, J.; Monchaud, D.; Mergny, J.-L.; Ju, H.; Zhou, J., How Proximal Nucleobases Regulate the Catalytic Activity of G-Quadruplex/Hemin DNAzymes. *ACS Catalysis* **2018,** *8* (12), 11352-11361.

164. Kejnovská, I. K., J. Vorlíčková, M., Circular dichroism spectroscopy of conformers of (guanine + adenine) repeat strands of DNA. *Chirality* **2003,** *15* (7), 584-592.

165. Qiu, D.; Mo, J.; Liu, Y.; Zhang, J.; Cheng, Y.; Zhang, X., Effect of Distance from Catalytic Synergy Group to Iron Porphyrin Center on Activity of G-Quadruplex/Hemin DNAzyme. *Molecules* **2020,** *25* (15).

166. Gerasimova, Y. V.; Cornett, E. M.; Edwards, E.; Su, X.; Rohde, K. H.; Kolpashchikov, D. M., Deoxyribozyme cascade for visual detection of bacterial RNA. *Chembiochem* **2013,** *14* (16), 2087-90.

167. Reed, A. J.; Connelly, R. P.; Williams, A.; Tran, M.; Shim, B. S.; Choe, H.; Gerasimova, Y. V., Label-Free Pathogen Detection by a Deoxyribozyme Cascade with Visual Signal Readout. *Sens Actuators B Chem* **2019,** *282*, 945-951.

168. Banerjee, J.; Nilsen-Hamilton, M., Aptamers: multifunctional molecules for biomedical research. *J Mol Med (Berl)* **2013,** *91* (12), 1333-42.

169. Bauer, M.; Strom, M.; Hammond, D. S.; Shigdar, S., Anything You Can Do, I Can Do Better: Can Aptamers Replace Antibodies in Clinical Diagnostic Applications? *Molecules* **2019,** *24* (23).

170. Ouellet, J., RNA Fluorescence with Light-Up Aptamers. *Front Chem* **2016,** *4*, 29.

171. Tan, X.; Constantin, T. P.; Sloane, K. L.; Waggoner, A. S.; Bruchez, M. P.; Armitage, B. A., Fluoromodules Consisting of a Promiscuous RNA Aptamer and Red or Blue Fluorogenic Cyanine Dyes: Selection, Characterization, and Bioimaging. *J Am Chem Soc* **2017,** *139* (26), 9001-9009.

172. Neubacher, S.; Hennig, S., RNA Structure and Cellular Applications of Fluorescent Light-Up Aptamers. *Angew Chem Int Ed Engl* **2019,** *58* (5), 1266-1279.

173. Debiais, M.; Lelievre, A.; Smietana, M.; Muller, S., Splitting aptamers and nucleic acid enzymes for the development of advanced biosensors. *Nucleic Acids Res* **2020,** *48* (7), 3400-3422.

174. Shelke, S. A.; Shao, Y.; Laski, A.; Koirala, D.; Weissman, B. P.; Fuller, J. R.; Tan, X.; Constantin, T. P.; Waggoner, A. S.; Bruchez, M. P.; Armitage, B. A.; Piccirilli, J. A., Structural basis for activation of fluorogenic dyes by an RNA aptamer lacking a G-quadruplex motif. *Nat Commun* **2018,** *9* (1), 4542.

175. Kolpashchikov, D. M.; Spelkov, A. A., Binary (Split) Light-up Aptameric Sensors. *Angewandte Chemie* **2020,** *133* (10), 5040-5051.

176. Babendure, J. R.; Adams, S. R.; Tsien, R. Y., Aptamers switch on fluorescence of triphenylmethane dyes. *J Am Chem Soc* **2003,** *125* (48), 14716-7.

177. Chen, R. F., Fluorescence of free and protein-bound Auramine O. *Archives of Biochemistry and Biophysics* **1977,** *179* (2), 672-681.

178. Silva, M. S.; Senna, S. G.; Ribeiro, M. O.; Valim, A. R.; Telles, M. A.; Kritski, A.; Morlock, G. P.; Cooksey, R. C.; Zaha, A.; Rossetti, M. L., Mutations in katG, inhA, and ahpC genes of Brazilian isoniazid-resistant isolates of Mycobacterium tuberculosis. *J Clin Microbiol* **2003,** *41* (9), 4471-4.

179. Kolpashchikov, D. M., An elegant biosensor molecular beacon probe: challenges and recent solutions. *Scientifica (Cairo)* **2012,** *2012*, 928783.

180. Compton, J., Nucleic acid sequence-based amplification. *Nature* **1991,** *350* (6313), 91-2.

181. Dhar, B. C.; Reed, A. J.; Mitra, S.; Rodriguez Sanchez, P.; Nedorezova, D. D.; Connelly, R. P.; Rohde, K. H.; Gerasimova, Y. V., Cascade of deoxyribozymes for the colorimetric analysis of drug resistance in Mycobacterium tuberculosis. *Biosens Bioelectron* **2020,** *165*, 112385.

182. Sato, Y.; Nishizawa, S.; Teramae, N., Label-free molecular beacon system based on DNAs containing abasic sites and fluorescent ligands that bind abasic sites. *Chemistry* **2011,** *17* (41), 11650-6.

183. Du, Y.; Li, B.; Wang, E., "Fitting" makes "sensing" simple: label-free detection strategies based on nucleic acid aptamers. *Acc Chem Res* **2013,** *46* (2), 203-13.

184. Guo, Y.; Xu, L.; Hong, S.; Sun, Q.; Yao, W.; Pei, R., Label-free DNA-based biosensors using structure-selective light-up dyes. *Analyst* **2016,** *141* (24), 6481-6489.

185. Ma, D. L.; He, H. Z.; Leung, K. H.; Zhong, H. J.; Chan, D. S.; Leung, C. H., Label-free luminescent oligonucleotide-based probes. *Chem Soc Rev* **2013,** *42* (8), 3427-40.

186. Shi, S.; Wang, X.; Sun, W.; Wang, X.; Yao, T.; Ji, L., Label-free fluorescent DNA biosensors based on metallointercalators and nanomaterials. *Methods* **2013,** *64* (3), 305-14.

187. Del Bonis-O'Donnell, J. T.; Vong, D.; Pennathur, S.; Fygenson, D. K., A universal design for a DNA probe providing ratiometric fluorescence detection by generation of silver nanoclusters. *Nanoscale* **2016,** *8* (30), 14489-96.

188. Guo, Y.; Pan, X.; Zhang, W.; Hu, Z.; Wong, K. W.; He, Z.; Li, H. W., Label-free probes using DNA-templated silver nanoclusters as versatile reporters. *Biosens Bioelectron* **2020,** *150*, 111926.

189. Connelly, R. P.; Madalozzo, P. F.; Mordeson, J. E.; Pratt, A. D.; Gerasimova, Y. V., Promiscuous dye binding by a light-up aptamer: application for label-free multi-wavelength biosensing. *Chem Commun (Camb)* **2021,** *57* (30), 3672-3675.

190. Nimjee, S. M.; White, R. R.; Becker, R. C.; Sullenger, B. A., Aptamers as Therapeutics. *Annu Rev Pharmacol Toxicol* **2017,** *57*, 61-79.

191. Wang, Z.; Luo, Y.; Xie, X.; Hu, X.; Song, H.; Zhao, Y.; Shi, J.; Wang, L.; Glinsky, G.; Chen, N.; Lal, R.; Fan, C., In Situ Spatial Complementation of Aptamer-Mediated Recognition Enables Live-Cell Imaging of Native RNA Transcripts in Real Time. *Angew Chem Int Ed Engl* **2018,** *57* (4), 972-976.

192. Paige, J. S.; Nguyen-Duc, T.; Song, W.; Jaffrey, S. R., Fluorescence imaging of cellular metabolites with RNA. *Science* **2012,** *335* (6073), 1194.

193. Jaffrey, S. R., RNA-Based Fluorescent Biosensors for Detecting Metabolites in vitro and in Living Cells. *Adv Pharmacol* **2018,** *82*, 187-203.

194. Wu, R.; Karunanayake Mudiyanselage, A.; Shafiei, F.; Zhao, B.; Bagheri, Y.; Yu, Q.; McAuliffe, K.; Ren, K.; You, M., Genetically Encoded Ratiometric RNA-Based Sensors for Quantitative Imaging of Small Molecules in Living Cells. *Angew Chem Int Ed Engl* **2019,** *58* (50), 18271-18275.

195. Yu, Q.; Shi, J.; Mudiyanselage, A.; Wu, R.; Zhao, B.; Zhou, M.; You, M., Genetically encoded RNA-based sensors for intracellular imaging of silver ions. *Chem Commun (Camb)* **2019,** *55* (5), 707-710.

196. Stojanovic, M. N.; Kolpashchikov, D. M., Modular aptameric sensors. *J Am Chem Soc* **2004,** *126* (30), 9266-70.

197. Baugh, C.; Grate, D.; Wilson, C., 2.8 A crystal structure of the malachite green aptamer. *J Mol Biol* **2000,** *301* (1), 117-28.

198. Warner, K. D.; Chen, M. C.; Song, W.; Strack, R. L.; Thorn, A.; Jaffrey, S. R.; Ferre-D'Amare, A. R., Structural basis for activity of highly efficient RNA mimics of green fluorescent protein. *Nat Struct Mol Biol* **2014,** *21* (8), 658-63.

199. Trachman, R. J., 3rd; Demeshkina, N. A.; Lau, M. W. L.; Panchapakesan, S. S. S.; Jeng, S. C. Y.; Unrau, P. J.; Ferre-D'Amare, A. R., Structural basis for high-affinity fluorophore binding and activation by RNA Mango. *Nat Chem Biol* **2017,** *13* (7), 807-813.

200. Warner, K. D.; Sjekloca, L.; Song, W.; Filonov, G. S.; Jaffrey, S. R.; Ferre-D'Amare, A. R., A homodimer interface without base pairs in an RNA mimic of red fluorescent protein. *Nat Chem Biol* **2017,** *13* (11), 1195-1201.

201. Dolinnaya, N. G.; Yuminova, A. V.; Spiridonova, V. A.; Arutyunyan, A. M.; Kopylov, A. M., Coexistence of G-quadruplex and duplex domains within the secondary structure of 31-mer DNA thrombin-binding aptamer. *J Biomol Struct Dyn* **2012,** *30* (5), 524-31.

202. Troisi, R.; Napolitano, V.; Spiridonova, V.; Russo Krauss, I.; Sica, F., Several structural motifs cooperate in determining the highly effective anti-thrombin activity of NU172 aptamer. *Nucleic Acids Res* **2018,** *46* (22), 12177-12185.

203. Russo Krauss, I.; Napolitano, V.; Petraccone, L.; Troisi, R.; Spiridonova, V.; Mattia, C. A.; Sica, F., Duplex/quadruplex oligonucleotides: Role of the duplex domain in the stabilization of a new generation of highly effective anti-thrombin aptamers. *Int J Biol Macromol* **2018,** *107* (Pt B), 1697-1705.

204. Keniry, M. A., Quadruplex structures in nucleic acids. *Biopolymers* **2001,** *56* (3), 123-146.

205. Miyahara, T.; Nakatsuji, H.; Sugiyama, H., Helical structure and circular dichroism spectra of DNA: a theoretical study. *J Phys Chem A* **2013,** *117* (1), 42-55.

206. del Villar-Guerra, R.; Trent, J. O.; Chaires, J. B., G-Quadruplex Secondary Structure Obtained from Circular Dichroism Spectroscopy. *Angewandte Chemie* **2018,** *130* (24), 7289-7293.

207. Ambrus, A.; Chen, D.; Dai, J.; Bialis, T.; Jones, R. A.; Yang, D., Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic Acids Res* **2006,** *34* (9), 2723-35.

208. Butovskaya, E.; Heddi, B.; Bakalar, B.; Richter, S. N.; Phan, A. T., Major G-Quadruplex Form of HIV-1 LTR Reveals a (3 + 1) Folding Topology Containing a Stem-Loop. *J Am Chem Soc* **2018,** *140* (42), 13654-13662.

209. Zaitseva, S. O.; Baleeva, N. S.; Zatsepin, T. S.; Myasnyanko, I. N.; Turaev, A. V.; Pozmogova, G. E.; Khrulev, A. A.; Varizhuk, A. M.; Baranov, M. S.; Aralov, A. V., Short Duplex Module Coupled to G-Quadruplexes Increases Fluorescence of Synthetic GFP Chromophore Analogues. *Sensors (Basel)* **2020,** *20* (3).

210. Fernandez-Millan, P.; Autour, A.; Ennifar, E.; Westhof, E.; Ryckelynck, M., Crystal structure and fluorescence properties of the iSpinach aptamer in complex with DFHBI. *RNA* **2017,** *23* (12), 1788-1795.

211. Zhao, D.; Dong, X.; Jiang, N.; Zhang, D.; Liu, C., Selective recognition of parallel and anti-parallel thrombin-binding aptamer G-quadruplexes by different fluorescent dyes. *Nucleic Acids Res* **2014,** *42* (18), 11612-21.

212. Nicoludis, J. M.; Miller, S. T.; Jeffrey, P. D.; Barrett, S. P.; Rablen, P. R.; Lawton, T. J.; Yatsunyk, L. A., Optimized end-stacking provides specificity of N-methyl mesoporphyrin IX for human telomeric G-quadruplex DNA. *J Am Chem Soc* **2012,** *134* (50), 20446-56.

213. Zuker, M., Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **2003,** *31* (13), 3406-15.

214. Diwu, Z.; Lu, Y.; Zhang, C.; Klaubert, D. H.; Haugland, R. P., Fluorescent Molecular Probes II. The Synthesis, Spectral Properties and Use of Fluorescent Solvatochromic Dapoxyl Dyes. *Photochemistry and Photobiology* **1997,** *66* (4), 424-431.

215. MacDougall, D.; Crummett, W. B., Guidelines for data acquisition and data quality evaluation in environmental chemistry. *Analytical Chemistry* **1980,** *52* (14), 2242-2249.