# Structural Recycling of Model Fluorescent Aptamers

By

Jonathan Carey Savage

A DISSERTATION

Presented to the Department of Chemical Physiology & Biochemistry

and the Oregon Health & Science University

School of Medicine

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

June 2021

School of Medicine

Oregon Health & Science University

CERTIFICATE OF APPROVAL

This is to certify that the PhD dissertation of

Jonathan Carey Savage

has been approved

---

Ujwal Shinde, PhD (Mentor)

---

Monika Davare, PhD (co-Mentor)

---

Matt Thayer, PhD (Committee Chair)

---

Dave Farrens, PhD

---

Larry David, PhD

---

Andrew Adey, PhD

# Table of Contents

# List of Figures

# Acknowledgments

First and foremost, I would like to thank my primary scientific mentor, Dr. Ujwal Shinde. This dissertation contains the published fraction of our six year dive into research field that was new to both of us. For years, it has been just myself and Ujwal in the lab, which coupled with our shared exploration of a completely new field, produced a friendship and closeness that far exceeds that of a common student-advisor mentorship. The personal and scientific respect I received from Ujwal have been invaluable to me in my growth as a scientist and a person. Perhaps most significantly, however, Ujwal has exemplified how to maintain spirit, optimism, and dignity in the face of day in, day out struggle and adversity. For all of this, I thank him.

Similarly, I wish to thank my co-mentor, Dr. Monika Davare. Monika has been a terrific source of scientific criticism, research strategy, and frank advice, and I have always appreciated her candor. I have witnessed Monika progress from becoming faculty, to building a lab from one to technician to a team of graduate students, clinician fellows and staff scientists, to receiving her first RO1 on the cusp of tenure. Her experience has been an education to be part of and observe.

I have to acknowledge the assistance I received from a summer student turned scientific peer, Pushkar Shinde. I cannot say enough about the quality and caliber of Pushkar as a student, a scientist, and ultimately, a person. His brilliance and raw talent are matched only by his work ethic and humility. It has been a pleasure to work with him and witness his personal development.

My Dissertation Advisory Committee provided me with substantial help. Drs. Matt Thayer, Dave Farrens, and Larry David have all been extremely generous with their time, experience, criticism, and scientific equipment throughout my graduate education. This work would not have been successful without their help, and to each of them I am thankful.

I have been inspired by the resolve of the remaining members of the Department of Biochemistry & Molecular Biology, who have continued to produce exemplary research, mentorship, and education in the face of difficult times. This department has known struggle in the last decade, and the personnel numbers dwindled so low during my time I can acknowledge nearly everyone individually, so I will: Thank you Ujwal Shinde, Dave Farrens, Matt Thayer, Larry David, Linda Musil, Show-Ling Shyng, Buddy Ullman, Dave Kabat, Michael Chapman, Peter Rotwein, Maureen Hoatlin, Hans-Peter Bächinger, Phil Yates, Steve Mansoor, Judy VanSlyke, Leslie Smith, Zhongying Yang, Emily Platt, Bruce Boswell, Ashok Reddi, Omar Davulcu, Jon Fay, Rob Meza-Romero, Yoshi Ishikawa, Raj Shah, Cindy McDermid, Johannes Elferich, Danielle Williamson, Jessica Martin, Nathan Montgomery, Chris Schafer, Amber Jones, Greg Martin, Chiara del Piccolo, Veronica Cochrane, Sam Burke, Anthony Shumate, Jonny Flores, Rich Posert, and Sigrid Noreng.

I wish to thank two personal role models, Cory Soto and Jim Pesout. Ebbs and flows of life have put significant time and distance between myself and both of these individuals, but their impacts on who I am and who I wish to become are immeasurable.

I owe so much to my close friends, my sister, and my parents for their support and enduring patience with my frustrations and obsessions and absences during my graduate education. I cannot possibly thank all of them enough.

Finally, this work would not have been possible without the endless love and support of Sami Friedrich. The blend of romance, friendship, and partnership I have experienced with her is more fulfilling than I used to think possible. Sami inspires me to think deeply, act with intention, and live vibrantly, and I look forward to our next adventure together.

# Abbreviations

**SELEX** ..... Systematic Evolution of Ligands by Exponential enrichment

**PCR** ..... Polymerase Chain Reaction

**2'OH** ..... Ribose 2' Hydroxyl

**2'F** ..... Ribose 2' Fluorine

**DNA** ..... Deoxyribonucleic Acid

**RNA** ..... Ribonucleic Acid

**FRET** ..... Förster Resonance Energy Transfer

**Py** ..... Pyrimidine

**Pu** ..... Purine

**$K^+$** ..... Potassium

**$Pb^{2+}$** ..... Lead

# Abstract

This dissertation explores the capacity of direct chemical modification of existing aptamer structures for the discovery of novel functional applications using a closely related family of fluorescent aptamers as a model system. Further, this dissertation explores the merits of reusing validated structural motifs for the accelerated discovery of novel aptamer utility. This body of work contains (i) the first instances of direct chemical modification of internal residues for the execution of a function an aptamer was not selected to perform, (ii) the first generation of an aptamer chimera of mixed nucleotide chemistries, engineering via a universal approach; (iii) the first RNA $K^+$ biosensor combined with the widest range of fluorescent $K^+$ detection of any nucleic acid cation sensor yet published, covering nearly three orders of magnitude. This work additionally raises broad questions about the influence of sequence identity and proximity of structure in G-quadruplex-cation interactions, and its relationship to G-Quadruplex behavior with implications for rational tuning, and the merits of structural recycling and functional reprogramming of aptamers as a conceptual alternative to *de novo* aptamer generation.

# 1

# Introduction

## 1.1 Nucleic Acids Aptamers and SELEX

### *1.1.1 Aptamers*

The advent of PCR and molecular cloning methods yielded an explosion of techniques and tools for the expression and purification of recombinant proteins, facilitating the biochemical characterization of reductionist *in vitro* reactions and manipulation of complex cellular systems. During this same period, our understanding of the biological significance of RNA shifted from a necessary and ubiquitous, though unstable and seemingly benign intermediary molecule for the transfer of genetic information, to a complex polymer with structure-dependent function capable of independent catalysis, then further to the theorized center of life preceding the current iteration requiring DNA and protein.[1] During this time, a wealth of basic biological knowledge of nucleic acid structure and function, and a rich library of techniques and tools for manipulating them were produced. One particular technique arose from the distillation of this knowledge, inspired by the concept of the "RNA World" while simultaneously providing evidence for it.[1–3] The **S**ystematic **E**volution of **L**igands by

**EX**ponential enrichment (SELEX) technique was developed concurrently and independently by three research groups in the late 1980s for the simple and rapid identification of functional nucleic acid structures (Figure 1-1).[4–6] The selection of these structures, named "Aptamers", has helped fundamentally shift our basic understanding of nucleic acid biophysics and molecular evolution, while simultaneously providing invaluable research and diagnostic tools. The execution of the technique SELEX has remained relatively niche, however its conceptual simplicity, broad applicability, and utility for garnering powerful evolutionary insights has ensured its place as a staple biochemical technique.

The word aptamer is a fusion of Latin "*aptus*" (to fit), and Greek "*meros*" (part), and is the name given to the synthetic nucleic acid sequences with structure-dependent functions that are identified by an *in vitro* evolution technique, SELEX. Under this definition, aptamers are distinct entities from naturally evolved functional RNAs such as ribozymes or riboswitches, though aptamers can theoretically be selected to perform the identical function of a natural RNA with identical sequence. However, given their *in vitro* genesis, aptamers are composed from a broader chemical repertoire, and can be composed of both DNA or RNA, or polymers with alternative bioorthogonal backbone chemistries collectively denoted "XNAs".[7,8] Broadly speaking, all aptamers identified can be binned into two functional classes, catalysts and affinity binders. Within these two classes thousands of unique aptamers have been identified which selectively bind diverse small-molecule and protein targets, and catalyze numerous and diverse biochemical reactions.

As affinity reagents—their most common use and by far most populous class—aptamers are attractive for many reasons. Aptamers are produced *in vitro* by organic synthesis or enzymatic polymerization, and thus display no batch-to-batch variability with highly scalable production. Being composed of short nucleic acid sequence, aptamers are inherently an alternative biochemical complement to more commonly used proteinaceous reagents, principally mammalian antibodies. Nucleic acids are generally less immunogenic than antibodies, and are more easily modified with dyes, fluorophores, lipids, bioorthogonal functional groups, and synthetic nucleotide analogs (a deeper discussion of nucleotide modifications can be found in section 1.3). The genotype-phenotype relation in nucleic acids (referring to a polynucleotide's capacity to both encode for itself and perform a structure-dependent function)—which is the foundation and basis of SELEX—facilitates aptamer selection in a biochemically simpler and more rapid manner than comparative *in vitro* evolution of protein by techniques such as phage display.[9,10] SELEX, by its very nature, produces the simplest solution to a given selection paradigm.[11] Aptamers therefore tend to be very small (on the order of 5-30kDa), which in turn makes them reversibly denaturable by heat and inexpensive to synthesize in large quantities. Similarly, aptamers are readily accessible to small biological compartments and less obtrusive when used to probe macromolecular complexes (Figure 1-2.A). Nucleic acid structure is highly flexible, with backbones containing six bonds with individual torsional rotation (compared to the peptide backbone which contains two) and a glycosidic X bond around which nucleobases can fully rotate (Figure 1-2.B). This flexibility

results in aptamer constructs with comparable affinities, often more stringent selectivities, and higher mean shape complementarity indices than antibodies.[12,13]

Perhaps the most attractive feature of aptamers is the ease of their dissemination. Identifying new aptamer *de novo* is conceptually simple but requires patience, precision, repetition, experience, and luck. However, acquiring and using an aptamer sequence that has been previously selected, sequenced, and characterized is very simple. Aptamers do not require specialized cell systems for expression, equipment for their purification, or molecular chaperones for folding, and are extremely soluble in aqueous solution. An aptamer' sequence is all the information needed for a laboratory to validate or utilize that aptamer to their own ends.

### *1.1.2 The Basics of SELEX*

As the name suggests SELEX is an *in vitro* evolution technique that follows the same principles of Darwinian selection observed in natural evolution.[3] The minimalist framework for a SELEX experiment is simple. Figure 1-1 provides a visual outline of a simplified SELEX framework for identifying protein-binding aptamers. A polynucleotide progenitor library of uniform length, flanked by constant primer sequences, is synthesized with some degree of heterogeneity ranging from a few specific positions to total randomness with equal representation of each nucleotide at every position. Theoretically, the heterogeneous population contains within it at least one species with a desired biophysical trait that is inherent to its sequence and structure. That desired species is isolated by subjecting the population to selective pressure favoring that desired trait, say, binding to a target protein. Favorable binders are physically separated from the

bulk population and then amplified by PCR, enriching their representation (reverse transcription is necessarily performed in RNA-based selections). This process is repeated iteratively, with increasingly stringent selective pressure, favoring fewer and fewer species with increasingly favorable behavior with each generation. When sufficient homogeneity is suspected the population is sequenced, after which major species are screened for function in isolation. Within this framework of heterogeneity → selective pressure → isolate → enrich → repeat, substantial freedom of customization exists for identifying aptamers with diverse functions. Numerous distinct SELEX derivations have been described, with selective pressure the most commonly explored variable. Selections based on size, shape, rate of association or dissociation, equilibrium affinity, catalytic activity, thermal stability, structural fold, and much else have been explored*. A proper summation of even a fraction of the unique SELEX canon is well beyond the scope of discussion here. However reviews are published annually, many of which are informative, all of which contain a unique combination of examples.[14–19]

The power of SELEX lies in the extreme heterogeneity of sequence and structure it theoretically samples. It is routine to initiate a SELEX experiment with a progenitor library containing 40 or more positions of randomness, yielding $>4^{40}$ ($10^{24}$) unique species, theoretically embodying as many unique biophysical properties. Within such a library, every sequence and structure that can possibly ever exist for that nucleotide length is

*It should be noted, "*in vitro* evolution" and "*in vitro* selection" are distinct approaches in the literature. The former introduces subtle variation through additional mutation throughout the process while the latter does not, acting instead to simply filter the sequence distribution of the original library. This subtle distinction is important, however, for the purposes of discussion in this dissertation the words "selection" and "evolution" will be used interchangeably.

simultaneously present in equal proportions (again, theoretically). The spontaneous generation of complex structures from a pool of monomeric nucleotides is so unlikely, particularly in primordial conditions, as to be effectively impossible. Natural evolution overcomes this obstacle through subtle change of existing structure, allowing for increased size and complexity over deep time. Natural evolution "discovers" new sequence-space as it *diverges* outward toward increasing heterogeneity from a common ancestral locus. SELEX, on the other hand, utilizes the reciprocal approach. SELEX begins with the chemical synthesis of vast random sequence pools that sample a vaster swath of disparately related sequence space. From this population SELEX is used to filter out unfavorable sequences, *converging* on increasing homogeneity. It is this contrast with natural evolution that is extremely valuable in identifying novel biophysical functions in experimentally relevant time scales. The reciprocation of initial heterogeneity and initial homogeneity provides another advantage to SELEX. Every individual selection can begin with a progenitor library that samples the same sequence space in totality, in effect removing any evolutionary memory of previous successful or unsuccessful selections. This provides each SELEX attempt unbiased access to sequence-space which may have been unfavorable in a previous selection for an alternate function.

## 1.2 Aptamer Applications

### *1.2.1 Unanticipated Functional Diversity*

Standard chemistry DNA and RNA polymers are each composed of just four base units, AT/UGC. The statistical likelihood that many favorable Watson-Crick (WC) pairs, G-

U wobble pairs, or G-Quadruplex motifs will exist in even short nucleic acid sequences is very high. This results in stable regular structures for the vast majority of all possible nucleotides sequence combinations (>99.9% of all possible RNA sequence combinations are predicted to fold stably, as opposed to <2% of peptide sequences).[20,21] Even large combinatorial libraries are therefore mostly structured. With such structural diversity simultaneously present, it is reasonable to presume every progenitor population contains at least one structure that is favorable for nearly any selection paradigm that can be biochemically executed. The diversity of progenitor populations can be gleaned from the broad range of aptamer functions that have been successfully selected and characterized, particularly so for catalytic aptamer (deoxy)ribozymes (also referred to as "artificial ribozymes" or "aptazymes").[22] Interesting catalytic activities include (but are not limited to) aptamers performing non-covalent biphenyl isomerization, $Pb^{2+}$-dependent phosphodiester hydrolysis, *trans*- and *cis*- aminoacylation, amide hydrolysis, acetyl transferase, porphyrin methylation, peptide bond condensation, Diels-Alder cycloaddition, and the aldol reaction.[23–32] One SELEX-derived RNA enzyme displays true *trans*-aminoacylation catalysis despite being only 5 nucleotides long, the smallest catalytic nucleic acid or protein so far known (strong evidence supporting the feasibility of an RNA World).[33–35] Most of these aptazymes were identified less than fifteen years after nucleic acid catalysis was discovered to even be possible let alone so potentially diverse.[36] Additional discussion of catalytic aptamer mechanism and chemistry, with comparison to natural ribozymes, can be found in several detailed reviews.[37,38]

Similarly impressive, and greater in number, is the ever-growing collection of aptamers selected to bind protein or small molecule ligands*. Aptamer affinities for protein ligands can be extremely high, regularly selected in the picomolar ($10^{-12}$ M) to mid nanomolar ($10^{-8}$ M) range prior to any additional optimization. While aptamer affinities for small-molecules tend to be lower than for protein (a result of the chemical homogeneity of aptamers combined with the reduced chemi-structural information provided by a small-molecule ligand), the selectivity of aptamers for both protein and small-molecule ligands is exquisite. An anti-p53 aptamer discriminates between WT and a single amino acid mutant $R_{175}H$ sufficiently to alter heterozygous cell survival, and an anti-theophylline aptamer distinguishes between theophylline and caffeine—different by just a single methyl group—with a 10,000-fold preference.[12,39]

At the time of this writing, thirty years of selection has produced thousands of unique aptamers against as many chemically distinct ligands. While certain biophysical restrictions must exist, it appears the limiting factor of successful aptamer selection is more aligned with the cleverness and skill (and luck) of the experimenter than an inherent property of nucleic acids (admittedly, however, it is difficult to find publications describing failed SELEX attempts).[40,41] The Ellington Lab established the Aptamer Database in 2004.[42] This repository, currently referred to as the "Apta-Index" and maintained by the company Aptagen, provides sequence information, nucleotide chemistry, ligand identity, buffer

*Despite "L" in SELEX standing for "ligand" and thus describing the aptamer and not its binding target, this dissertation uses the word "ligand" exclusively in reference to the molecule an aptamer binds.

conditions, affinity data, and publication links for voluntarily deposited aptamers. As of May 2021, the database contains 547 non-redundant nucleic acid affinity aptamers, which while non-exhaustive remains a terrific resource. This collection of affinity aptamers pales in comparison, however, to the private collection held by the company SOMAlogic, Inc. The company has reported the development of equipment and protocols for highly parallel and automated selection of modified aptamers for the development of lab-on-chip diagnostic devices (more on modified nucleotides and SOMAlogic in section 1.3.3). The company reports a proprietary collection of >10,000 unique protein-binding aptamers targeting >5,000 distinct human proteins.[40] The vast majority of this collection is proprietary, however the sheer size of their reported collection is indicative of the research, diagnostic, and therapeutic potential of aptamers.

Given their capacity for exquisite binding affinity and ligand selectivity, and physical traits discussed above (principally their small size, batch-to-batch consistency, and ease of modification or conjugation), aptamers have been increasingly looked to for in *in vitro* analytical applications as biosensors.[17] Any selective affinity interaction between two binding partners is a form of "detection". Transducing detection into a self-contained and functionally associated output signal engenders a "sensor", for which aptamers appear to be extremely conducive. Considering specialized instrumentation—all of which contain their own internal detectors and sensors—and precise computation is required to accurately quantify biosensor output signals, the line that delineates a true direct aptamer biosensor from an aptamer-based binding assay is subjective. However,

significant strides have been made across the spectrum, all of which are interesting and merit deeper exploration beyond the discussion below.[17,40,43,44]

### *1.2.2 Optical Biosensors*

Optical aptamer biosensors typically take advantage of the chemical conjugation of one or more fluorophores to residues proximal to ligand binding or to conformationally flexible domains, most commonly a 5' or 3' terminus. Strategic fluorophore placement produces a gain or loss of fluorescent signal that is specific to ligand binding, and proportional over some range of ligand concentration. In on example, two fluorescent dyes (acridine and fluorescein) were individually conjugated to residues adjacent to ligand binding sites in the anti-ATP aptamer, resulting in increased fluorescent signal upon ligand binding. Bulky aromatic conjugates in the ligand binding pocket reduced ATP affinity by ~10-fold, but produced a sensor that directly senses ATP over an order of magnitude in concentration in the high micromolar range.[45] Similarly, the fluorophore BODIPY (boron dipyrromethene) was successfully conjugated to ligand-binding proximal residues. Alterations to localized environments (e.g. ligand binding) produce detectable changes in BODIPY quantum yield, allowing the generation of multiple BODIPY-based fluorescent aptamer sensors against AMP, tyrosinamide, and argininamide.[46,47] To overcome issues of reduced affinity common with fluorophore conjugation, fluorescent nucleotide analogs can instead be directly incorporated (a deeper discussion of modified nucleotides can be found in section 1.3). Aptamers that bind the proteins thrombin, IgE, and PDGF (platelet derived growth factor) (among others), have been successfully engineered into sensors by the direct incorporation of fluorescent adenine or guanosine analogs *post*-SELEX.[48]

Incorporation sites were chosen by a combination of explicit structural determination and mutagenic screening to determine the optimal position of incorporation to yield fluorescent changes upon ligand binding. Strategies for the incorporation of fluorescent nucleotide analogs *during* SELEX have also been developed, bypassing the sequence-specific optimization that is required for extant aptamers.[49,50]

Dual-reporter systems that rely on fluorophore/quencher pairs, Förster resonance energy transfer (FRET) pairs, or excimer pairs have similarly found great application in optical aptamer sensors.[51] These approaches are more conducive to aptamer constructs that display some ligand-dependent structural conformation which conditionally alters the proximity of the reporter pair. In one example, a termini-linked fluorescein/dabcyl fluorophore/quencher pair were used to turn an anti-cocaine aptamer into a fluorescent sensor.[52] The sensor displays reduced fluorescence (by increased quenching) upon cocaine binding, and is effective over two orders of magnitude of cocaine concentration from ~10-1000 micromolar. This paradigm has been applied to protein-binding aptamers, using both fluorophore/quenchers and FRET pairs.[53,54] The high selectivity of aptamers shines in protein sensing applications such as this, wherein high-throughput analysis can be performed for the quantification of differential expression or solution presence of highly similar protein isoforms, typically in low or sub-nanomolar concentrations.[40,53,55]

All of the preceding approaches rely on labelling the same aptamer sequence in *cis*-, however *trans*- pairs of interacting nucleic acids sequence can be employed to similar ends. One subset of *trans*-acting sensors utilize a short quencher-linked antisense strand that competes with the ligand for aptamer binding. In one particular example, two short

DNA sequences complementary to the primary sequence of different regions of an anti-ATP aptamer were linked with the fluorescein/dabcyl fluorophore/quencher pair. In the absence of ATP, the aptamer structure is destabilized and amenable to complete denaturation by intermolecular Watson-Crick (WC) base pairing, which sequesters both antisense strands in a manner that quenches fluorescence. Addition of ATP stabilizes the aptamer fold, favoring the dissociation of antisense strands, un-quenching fluorescence.[56] This general approach has been used for numerous described aptamers, and has additionally been incorporated into SELEX schemes for the identification of *trans*-acting aptamer sensors *de novo*.[57,58] The other major subset of *trans*-acting aptamer sensors are known collectively as "split aptamers". With reasonable troubleshooting, aptamers can be split into two strands, which act independently in solution but associate and fold in the presence of their respective ligand. The same fluorophore/quencher or FRET pairs are conducive for termini conjugation in split aptamer pieces for the observed gain or loss of signal upon ligand association. The short length and generally simple architecture of aptamer structures appears to make them amenable to splitting, as numerous aptamers—including all the common models (anti- thrombin, ATP, cocaine, etc.) and several *in vitro* evolved aptazymes—have been split. Splitting aptamers has become a subfield in its own right. Approaches for generating a split aptamer and several informative reviews are worth further exploration.[43,59,60]

A third approach of optical aptamer sensors is the incorporation of nanoparticle reporters, principally Quantum Dots (QDs) and Gold nanoparticles (GNPs). QDs provide several fluorescent advantages over small-molecule organic fluorophores: QDs generally

have larger stokes shifts, longer fluorescent lifetimes, and can be easily multiplexed with the same excitation wavelength but vastly different emission wavelengths.[61,62] QD-aptamer sensors are generally constructed in quencher or FRET pairs, similar to small-molecule fluorophores (though much larger), and have been successfully implemented to sense small-molecules, proteins, and whole cells.[63–65] GNPs typically rely on changes in absorption or colorimetry upon changes in concentration or aggregation state. This has been achieved in several ways. One way employs the same logic of competitive *trans*-acting antisense strands discussed above, which promoted the diffusion of otherwise aggregated GNP-aptamer conjugates in the presence of ligand, detected by colorimetric shift.[66] Alternatively, GNPs conjugated to primers complementary to short sequences within the aptamer will aggregate in the absence of ligand. Ligand presence induces aptamer folding and dissociation from the GNP-primer, reducing the surface negative charge and allowing aggregation proportional to ligand concentration, similarly detectable by colorimetric shift.[67] One study has found a seemingly hybrid approach, wherein ligand-induced folding released an aptamer conjugated to TAMRA (tetramethyl-6-carboxyrhodamine) fluorophore. Diffusion away from the GNP un-quenches TAMRA fluorescence, similarly proportional to ligand concentration.[68]

### *1.2.3 Amplification Biosensors*

Aptamer biosensing by amplification utilizes couples the high binding selectivity of aptamers with the low limit of detection and exponential signal amplification of the polymerase chain reaction (PCR). The first approach is aptamer-based immuno-PCR (IPCR), which is essentially identical to the original IPCR with aptamers replacing the

antibodies.[69,70] In the original IPCR, solutions of unknown analyte concentration are fixed to a matrix or plate, after which a selective dsDNA-antibody conjugate is incubated. After binding and washing, PCR is performed, generating an amplicon only where analyte was present and proportional to its initial concentration.[69,71,72] Replacing antibodies with aptamers provides several benefits: (i) chemically synthesized aptamers display less batch-to-batch variability than antibodies, as discussed; (ii) this consistency is further enhanced as no second step dsDNA conjugation reaction is required when using a DNA aptamer, the appropriate sequence can simply be extended during aptamer synthesis; (iii) small soluble aptamers are less likely hinder PCR during high temperature denaturation steps or polymerase extension than larger more hydrophobic antibodies.

An alternative amplification-based aptamer biosensor approach is Proximity Ligation assay (PLA).[73–75] PLA is a conceptual amendment to sandwich enzyme-linked immunosorbent assays (ELISAs), requiring two aptamers, each containing a dsDNA extension and each binding distinct epitopes on the target ligand. Selective binding and washing is followed by enzymatic ligation, forming a template only where both aptamer sensors were present and bound to the same ligand molecule. PCR using a primer set that requires both dsDNA conjugate sequences ensures a genuine signal.[76] PLA selectivity is dramatically enhanced over IPCR, as it requires two independent aptamer-ligand interactions to stably form. The odds of non-specific interactions leading to amplicon formation are exceedingly low. PLA sensing requires either the ligand interacts as a homodimer, or that two unique aptamers are selected for different epitopes on the same ligand. Additionally, the ligand must be large enough to bind two aptamers, effectively

restricting PLA sensing to protein ligands. While this is restrictive in one respect, it does allow for PLA sensing of large complex structures such as whole bacterial and cancer cells.[77,78] The initial conception of PLA targeted the homodimer PDGF as proof of principle, and reported sensing *zeptomole* amounts of protein ($10^{-21}$ mol).[79]

### *1.2.4 Electrochemical Biosensors*

Aptamer biosensor applications are increasingly moving towards electrochemical (ECL) methods, as these techniques are dramatically more sensitive, they produce less optical artifacts and background spectra, the engineered chips and electrodes can be readily miniaturized, and the instrumentation is generally less specialized and less expensive. The majority of the general nucleic acid base pairing and aptamer folding principles discussed for optical sensing have been conceptually borrowed in ECL sensors. Instead of measuring the optical signals of fluorophores in response to aptamer-ligand binding, perturbations in the electrical properties of aptamer-conjugated conductive films and electrodes are measured in response to ligand association and aptamer folding. ECL approaches have been reported based on electrochemical impedance spectroscopy, amperometry, potentiometry, cyclic voltammetry, and electrogenerated chemiluminescence.[80–84]

While certain ECL approaches are conducive to label-free aptamer-ligand interactions, most approaches are based on a charge transfer between the electrode surface and a redox-active molecule linked to the aptamer terminus.[80] Highly specific electrochemical interactions between particular redox probes and selective electrode materials allows for ECL sensing to be performed in complex solutions including undiluted

serum, saliva, and various proteolyzed foodstuffs.[85–87] This flexibility of complex solution heterogeneity in conjunction with the simplicity of size and scalability is very conducive to diagnostic lab-on-chip development.[88] A recent informative review should be explored for more detailed discussion of aptamer-based ECL sensors.[44]

## 1.3 Chemically Modified Nucleotides

### *1.3.1 Limitations of Native Chemistry Nucleic Acids*

Despite the successes of native chemistry aptamer selection discussed in section 1.2, the chemical diversity of nucleic acid functional groups is limited, particularly when compared to protein. Proteins contain diverse functional groups not present in nucleic acids, with side-chains that are more reactive (carboxyl, thiol, and primary amine), have near-physiological $pK_a$ (imidazole), and are positively charged (ammonium and guanidino). Additionally, hydrophobic side chains (benzene, phenol, tryptamino, and alkyl) allow for strong entropically-driven interactions, tuning of dielectric environments which strengthen electrostatic interactions and shift $pK_a$s, and the ability to form pockets and packed cores with defined solvent boundaries. Additionally, while the flexibility of nucleic acids allows for highly complementary structures, more rigid polypeptide backbones may be beneficial for functional allostery, structural stability in wider environmental conditions (temperature, salt, etc.), or reducing the entropic cost of ligand binding. Inherent differences in backbone rigidity are exaggerated by protein side-chain interactions such as salt bridges and disulfide bonding not available to nucleic acids. Much of the chemistry that sets protein apart in native biological systems has directly inspired

synthetic nucleotide derivatives which can and have been readily incorporated into nucleic acids *in vitro.*

The principal difference between DNA and RNA, and source of the relative chemical instability of RNA, is the hydroxyl group at C2 of the ribose ("2'OH"). The 2'OH is oriented towards the backbone and partial hydrogen extraction, by acid-base catalysis via alkaline conditions or presence of ribonuclease, produces a strong nucleophile that engages in transesterification and auto-hydrolysis of the adjacent phosphodiester bond.[89,90] It is ironic this functional group is the primary source of RNA's chemical instability, as it is also the primary source of the structural stability that engenders RNA to be generally preferential over DNA for aptamer chemistry: the highly electronegative 2'OH is both a hydrogen bond donor and acceptor, engaging in intramolecular H-bonding and intermolecular hydration networks, while inducing a heavily shifted *3'endo* sugar pucker equilibrium contributing to substantial structural rigidity of an A-form RNA polymer.[91]

Emergent from the limitations of chemical diversity that exist within individual nucleobases are the structural limitations of massively combinatorial nucleic acid evolution *in vitro*. The functional diversity of successfully selected aptamers is enormous and expanding, however, certain aptamer functions may not be possible (at least within relevant timescales) even in a population of $>10^{15}$ combinations, a common starting population for a SELEX experiment. If a selection continues to fail to produce a genuine aptamer—and environmental variables such as time, temperature, salt, etc., are inflexible or have been exhaustively tested—one can increase sequence heterogeneity of the

starting library, thereby expanding the sequence-space sampled in the next selection attempt. This can be performed two ways, by modulating library length or library chemistry (or both). Randomized sequence lengths of 25-50 nucleotides are commonly employed for ligand affinity selection paradigms. While increasing sequence length does result in increasing heterogeneity and structural complexity, practical limits of concentration and volume quickly stifle heterogenic sampling beyond seemingly short lengths. To put the scale of sampling in perspective: with an average mass of ~339g/mol per nucleotide, the mass of a 100nt library that contains one copy of *every* possible sequence combination is ~$10^{36}$ kg, about one million times larger than the projected mass of the solar system. The fraction of that 100nt library that actually fits in a test tube is clearly an infinitesimally small slice of the theoretically available sequence-space. Yet, paradoxically, longer lengths such as these are required to provide a library necessary structural complexity with any statistical likelihood.[92] The inherent complications of scale at the onset of selection with long libraries are exasperated by additional factors that contribute to unselected and thus detrimental homogeneity. These include nonrandom library synthesis, off-target interactions, non-targeted compositional drift, and biased amplification of small or differentially stable structures during transcription and PCR.[14,93–95] One solution that mitigates these issues is retaining a shorter library length while modifying the nucleotide chemistry. This changes what area of sequence space is sampled without requiring literally astronomical masses, and can be as simple as switching from a library of DNA to one of RNA or include the incorporation of synthetic nucleotide analogs.

### *1.3.2 Chemical Identity*

The potential value of increased chemical diversity has not been lost on aptamer biochemists. Immediately upon the conception of SELEX and potential applications of aptamers, researchers began borrowing established nucleotide modifications while synthesizing novel ones, and incorporating them into aptamers for expanded function.[96] Broadly, chemical modifications of nucleotides fall into three structural classes based on nucleotide location: base, sugar, and backbone. Figure 1-3 shows a generic purine-pyrimidine dinucleotide, highlighting the numerous common positions of modification. The list of commercially available modifications is extensive and ever-growing, incorporating functional groups with ionizable, polar, aliphatic, aromatic, mirrored chirality, and bioorthogonal chemistries. Modified nucleotides are looked to for increased chemical, thermodynamic, and enzymatic stabilities, or to add chemical diversity to interfacial residues. Solid state synthesis is a requirement for use of many of these modifications, however the engineering of specific polymerase mutations allows for direct, template-based incorporation of select modifications during a SELEX experiment using standard molecular biology techniques.[97–99] A thorough analysis of specific modifications, or even a table of successfully implemented modifications, is far beyond the scope of discussion here, however many informative reviews can be found.[100–104]

Chemical modifications are incorporated into aptamer structures with two principal goals, to increase chemical stability or expand occupiable structural space. These goals are not mutually exclusive, and there are numerous examples of intentional and serendipitous overlap in observed functional outcome. Modifications for either intention can be incorporated throughout the duration of SELEX by addition to each passage's PCR

or transcription, or incorporated *post*-SELEX for aptamer optimization by rationally replacing specific residues with modified variants. The former approach allows the invisible hand of undirected selection to determine the optimal combination of number and location of synthetic nucleotides and thus overall aptamer structure, while the latter requires some subjective interpretation by an experimenter and typically involves multiplexed testing of various modifications of numerous residue positions.

### *1.3.3 Incorporation*

Modified nucleotides are incorporated into polymers by the same two methods of native chemistry nucleotides: solid phase phosphoramidite synthesis and enzymatic polymerization.[105] Within the context of *in vitro* evolution both approaches are typically employed during different phases. The vast majority of libraries are engineered exclusively by organic synthesis, which is conducive to generating random sequences by the simultaneous equimolar addition of every nucleotide. Synthesized initial libraries with constant defined sequences can be amplified by error prone polymerases to introduce stochastic diversity immediately prior to or during a selection.[106] This is commonly performed in directed evolution experiments of large natural ribozymes to create related families of sequences.[14] One alternative approach to traditional library synthesis is called genome-SELEX, and involves the generation of RNA transcripts from amplicons generated by randomized primer docking and PCR of genomic DNA for the identification of endogenous organism-specific RNA-ligand interactions.[107,108] This niche approach is much less used and has very focused applications.

Leslie Orgel's second axiom of evolution bluntly states, "Evolution is cleverer than you are."[109] In other words, evolution is a better "designer" of functional execution than a person can possibly be, taking into account more variables than a person is capable, analogous to the logic of contemporary machine learning algorithms. Organic synthesis of polynucleotides requires guided exogenous input—knowledge of either an explicit sequence or compositional bias. Both are antithetical to unguided selection based on pure biophysical favorability that *in vitro* evolutionary techniques such as SELEX attempt to capitalize on. Because of this, non-discriminatory enzymatic nucleotide polymerases are employed at the end of each selection cycle, providing the invisible hand of SELEX uncorrupted "control" over sequence progression. The relinquishment of control does, however, come at a cost. Only one modification per nucleobase per library is permitted for each selection experiment. The mechanism of nucleotide discrimination in polymerases is not specific to nucleobase recognition, but the planar geometry and favorable H-bonding of WC base pairing against a template polymer.[110–112] Therefore, if provided multiple differently modified flavors of the same nucleobase a polymerase will indiscriminately catalyze their incorporation at each complementary position in the template, producing a heterogeneous population that cannot be properly replicated and enriched during iterative SELEX passages. It is therefore necessary to clearly define the desired functional traits of a resulting aptamer prior to the onset of SELEX and commitment to a particular nucleotide chemistry. Upon the completion of SELEX, when a sequence or family of sequences is explicitly determined, experiments often return to organic synthesis for rational optimization.

### *1.3.4 Three Model Examples of Chemically Modified Nucleotides*

As discussed, the number of commercially available chemical modifications is vast and ever expanding, well beyond sufficient discussion here (though numerous informative can be found).[100–104] Three specific chemical modifications are worthy of attention here for their ubiquity, both to non-specialty labs toeing the waters of SELEX and to specialized research groups selecting aptamers for diagnostic or clinical use. The modified nucleotides are 2'-fluoro-2'-deoxypyrimidines (2'F), 2'-O-methyl-2'-deoxypurines (2'OMe), and 5-N-benzylcarboxamide-2'-deoxyuridine (Bn-dU) (Figure 1-3.A/B).

The earliest attempts to mitigate the chemical instability of structured RNA was the incorporation of modified nucleotides that replace the 2'OH with less basic functional groups, initially explored with the aforementioned 2'F pyrimidines and 2'-amino-2'-deoxypurines (2'$NH_2$).[113,114] Synthesis and polymer incorporation of 2'fluorinated nucleotides were reported decades before prior to their inclusion in structured RNA.[115,116] 2'F is comparable to 2'OH in size and electronegativity (though smaller and more negative), resulting in a more heavily favored 3'*endo* pucker. Additionally, replacing a hydroxyl with fluorine reduces the protecting group requirements during polymer synthesis, decreasing synthesis costs while increasing efficiency compared to other nucleotide modifications. For example, the natural Hammerhead ribozyme was found to retain partial catalytic activity even when heavily modified with 2'F pyrimidines, proving the modification to be sufficiently mimetic of natural nucleotides, stabilizing complex RNA structures against nucleases without completely disrupting their function.[117] The 2'F

modification was similarly observed to increase affinity and efficacy of antisense RNA oligonucleotides and effectiveness of siRNA gene silencers.[118] A series of studies on RNA-duplex stability by Egli and colleagues revealed the electron withdrawing strength of fluorine antiperiplanar to the glycosidic X bond is propagated across the nucleobase, resulting in increased inter-base H-bonding and more favorable base stacking.[119] This strong enthalpic favorability is complemented by comparatively minor entropic factors including pre-organized sugar conformation and substantially dehydrated minor groove, resulting from the 3'*endo* pucker and lack of H-bond donating of 2'fluorination.[120,121] The overall result is a substantial increase in duplex thermal stability (~1.8$^{o}$C per residue). 2'$NH_2$ is an H-bond donor *and* acceptor and displays a near-physiologic $pK_a$ ~6, making it initially quite attractive for incorporation into aptamers.[122] In direct contrast with 2'F, 2'$NH_2$ ribose heavily favors the *2'endo* pucker*,* destabilizing helical duplex and higher order RNA structure, and promoting premature termination during enzymatic incorporation.[117,123,124] For these reasons the modification was largely abandoned. In its place, 2'OMe—one of many naturally occurring post-transcriptional tRNA and rRNA modifications—was found to be a satisfying chemical replacement for its favorable 3'*endo* pucker and greater nuclease resistance (Figure 1-3.B).[125–127] Select mutations in T7 RNA polymerase have since been identified which increase both the rate and fidelity of enzymatic 2'OMe nucleotide incorporation to levels similar to 2'F, facilitating their ease-of-use in SELEX experiments and promoting their rise to the most widely used modifications in *de novo* aptamer selection.[97,98]

Great success has been found with complementary implementation of pyrimidine-specific and purine-specific modifications, most commonly 2'F and 2'OMe, respectively, as is the case in the extensively modified aptamer "Pegaptanib", the only therapeutic aptamer approved by the Food and Drug Administration (FDA) for clinical treatment of exudative age-related macular degeneration.[128,129] The disease results from dysregulated angiogenesis across the path of light in the vitreous humor of the eye, and is a major cause of blindness worldwide.[130] Pegaptanib was selected to bind the angiogenic growth factor, $VEGF_{165}$, and inhibit receptor interaction thereby suppressing blood vessel formation. The RNA aptamer was identified by traditional affinity SELEX using 2'F pyrimidine nucleotides from the onset, with 2'OMe purine incorporation by chemical synthesis during multiplexed optimization *post*-SELEX.[128] The final construct is modified at 25 of 27 residues (all pyrimidines and all but two purines) (Figure 1-3.B). Additionally, during pharmacokinetic optimization two terminal caps were added, a 5' 40kDa branched polyethylene glycol (PEG) moiety and a 3'-3' deoxythymidine, reducing exonuclease sensitivity and suppressing vitreous and renal clearance.[131] Together, these 4 combined chemical modifications provided the aptamer with half-lives of 131 hours and 9.3 hours in urine and vitreous humor, respectively (with detectable levels of $VEGF_{165}$ inhibition 28 days post-injection), dramatically longer than the unmodified, native chemistry equivalent.[128,129,131,132] The accomplishment of bringing Pegaptanib to market as a viable clinical tool is exciting for the future of aptamer therapeutics. However, from the vantage of nucleic acid biochemistry, Pegaptanib is exciting as an example of extreme aptamer modification. The original selection with only 2'F pyrimidine modifications yielded an

aptamer with a high affinity for $VEGF_{165}$ of 10pM.[128] A complete chemical overhaul, with modification at both termini (one with 40kDa of PEG) and 2'OMe replacement at all but two residues, dramatically enhanced chemical stability while only reducing the initial affinity by ~5-fold, to a final published affinity of 49pM.[128,129] Pegaptanib is composed of substantial WC duplex structure, which has been shown to be stabilized by these two specific 2' modifications, as discussed.

In contrast to modifications for the enhancement of chemical stability are modifications utilized for increasing the structural space an aptamer can occupy. This approach is best exemplified by a family of hydrophobic base-modified nucleotides incorporated into otherwise native DNA chemistry, which have found application for binding previously difficult protein targets. It would be impossible to discuss these modifications without first acknowledging the biotechnology firm developing them, which holds an effective monopoly on their implementation. Since its founding in 2000, SomaLogic, Inc, has been driving the development of modified aptamer selection against an ever-expanding swath of the human proteome (which includes Pegaptanib against $VEGF_{165}$, discussed above*). SomaLogic's principal goal is to produce ever-more fine-tuned aptamer-based lab-on-chip sensors for quantifying proteomic expression ratios in small volume serum samples for early detection of disease states, so called

*The original aptamer VP30.22 (now called "Pegaptanib") was selected by Dr. Larry Gold's startup, NeXstar (previously NeXaGen), which was bought by Gilead Sciences. The majority of the original personnel of NeXstar stayed behind to form SomaLogic. This same group of roughly a dozen researchers has remained the core team through three companies over 30 years, and can likely be credited with the partial discovery of aptamers and SELEX (through Gold) and the identification of >95% of all existing protein-binding aptamers.

"SOMAscans".[40,133] To this end, the company has successfully implemented the use of hydrophobic nucleotide modifications into their aptamer selection, improving selection success against difficult protein targets. SomaLogic has published several aptamer structures in peer reviewed journals, and while the specific sequence and protein target identities for the vast majority of their selected aptamers remain proprietary the company reports to possess selective aptamers against thousands of human and pathogen proteins.[40,134–137] The staple chemistry of their success is based on modifications at the C5 position of deoxyuridine that mimic hydrophobic and aromatic amino acid side-chains.[137] The four principle functional groups, each connected to uricil at C5 by an aminocarbonyl linker, are benzyl, naphthyl, tryptamino, and isobutyl (benzyl, or Bn-dU, is presented in Figure 1-3.C).[40]

Aptamers containing these modifications display extremely high affinities to their protein targets—commonly as low as 10pM—resulting from the aromatic-hydrophobic moieties at ligand interfaces, reminiscent of the amino acid biases of antibody CDRs (Complementarity Determining Regions).[138–142] The affinities of these aptamers are driven by extremely slow off-rate kinetics, inspiring the name "SOMAmers" (Slow Off-rate Modified Aptamers).[40,133] The published affinities of SOMAmers are the product of three traits: intermolecular hydrophobic interactions, structural stability, and shape complementarity.[13] Spectroscopic thermal denaturation analysis of pure WC duplex and structured SOMAmer sequence (both modified and unmodified) reveal both are stabilized by the presence of Bn-dU nucleotides.[143] While native DNA duplex with perfect complementarity is more stable than an engineered mismatch bulge containing Bn-dU,

that Bd-dU mismatch bulge is more stable than an equivalent bulge of native DNA.[143] Structural investigations suggest the source of this increased intramolecular structural stability is the formation of clusters of "benzyl-zippers", alternating aromatic $\pi$-stack interactions.[136,143] Both benzyl-zippers and edge-face interactions of the Bn-dU modification act as small hydrophobic cores which drive compaction and promote structural stability by entropic collapse similar to globular protein folding. Hydrophobic collapse and overall structural compaction additionally enhances nuclease resistance to SOMAmers*. Unlike Pegapatnib, which is modified at >90% of its residues, published SOMAmers contain a more modest degree of modification at 31-33% of their residue positions.[134–136] The remaining ~67% is unmodified DNA with conformational flexibility native DNA chemistry is allowed. The formation of zippered hydrophobic clusters within a flexible DNA background produces structure with exquisite shape complementarity to target proteins (exceeding even that of antibodies) and the capacity to fold into completely unique topologies only observed in SOMAmers.[13,137] There is much to be learned from the incorporation of hydrophobic modifications into nucleic acid structure. Regrettably, the majority of this intellectual property is held by a private company with little incentive for investment in structural determination or peer-reviewed publication of

*Interestingly, the compaction by hydrophobic collapse observed in SOMAmers has been proposed as a possible solution to the proposed statistical limitations of complexity in prebiotic nucleic acid structures. Very long RNAs sequences are required for catalysis of nucleic acid polymerization *in vitro (*Bernhardt 2012; Joyce 2020*).* How then could RNA polymerases evolve without the spontaneously and highly unlikely formation of large complex structures? Incorporation of hydrophobic residues in prebiotic nucleic acids (which are equally as likely—and equally as unfalsifiable—as any other nucleotide chemistry) can perhaps solve this problem through increased globularity and therefore decreased size of RNA structures required complex function (Wolk 2020).

their valuable aptamer collection.

### *1.3.5 Limitations of Modifications*

The benefits of modified nucleotide incorporation are hopefully apparent, however there are several caveats and perhaps two true drawbacks that need be considered. Native DNA is extremely inexpensive to synthesize in large scale, especially at the short lengths typical of aptamers. *In vitro* transcription of large quantities of native RNA is similarly inexpensive, and easy to produce and purify in-house. These conveniences are quickly lost upon incorporation of bulky modifications incompatible with polymerases, or multiple modifications of the same base in the same construct. As discussed, several mutant RNA polymerases have been described which can incorporate a *specific set* of modifications without a loss of fidelity, though these mutant polymerases have a strict list of accepted modifications. Organic synthesis by-passes this issue and provides the option of including multiple modifications of the same nucleobase in the same aptamer with position specific precision. However, synthesis becomes very expensive with increasing modification (especially in an RNA background), and, as discussed, is restricted to *post*-SELEX optimization of a previously selected aptamer.

There are two true detriments of selecting aptamers that incorporate and rely on modified nucleotides for function. The first is a general loss in reliability of structural prediction and molecular dynamic (MD) simulation**.** *In silico* structural prediction and MD simulations are valuable experimental approaches for proposing mechanisms, identifying critical residues, and approximating energetics of aptamer folding and ligand binding.[144] These tools are especially important for research programs that lack the specialized

expertise or funding for empirical structural determination. Prediction tools are assembled from information and trends observed in empirical structural determinations. Without extensive structure data, prediction software are inherently limited in accuracy and depth. The available tools for accurate structural prediction of nucleic acid polymers is therefore less developed than similar tools for proteins, which is both a result of and evidenced by the comparative dearth of resolved structures deposited in the Protein Data Bank (PDB)—as of April 2021, a PDB search yields ~3,500 nucleic acid structures as compared to the nearly ~170,000 unique protein structures. This is not to say that no 2D or 3D structural prediction software or MD force fields can process modified nucleotides, or that all modifications are equally difficult to model. Generally, the more extensively modified an individual structure is, or the more chemically disparate from native nucleic acid chemistry a particular modification is (e.g. bulky, aliphatic, or positively charged), the less reliable a purely theoretical model will be.[143] Predicted structures of all non-helical duplex nucleic acids, particularly so with modified residues, need to be considered with skepticism and tested appropriately.

The second detriment of modified nucleotide chemistry is a loss of potential expression in a cellular system by endogenous transcriptional machinery. This restriction actually applies to DNA-based aptamers as well, covering two out of three aptamer polymer chemistries. While aptamers can be introduced to the cytosol exogenously via the same techniques used for siRNAs or plasmid DNA, cell-to-cell variability, *in vivo* expression, and finely-tuned cellular manipulation are all best achieved by controlled in-cell transcription. Even if one could conveniently co-express precise levels of an

appropriate mutant polymerase, endogenous nucleotides will outcompete any successfully supplemented modified form, nullifying the attempt. Ultimately, if cellular manipulation is a desirable application, native RNA should be the chosen polymer. In fact, considering the chemical instability of RNA and the convenient availability of modifying it, perhaps cellular manipulation is the *only* condition purely native RNA is more fitting than DNA or a modified alternative.

## 1.4 Spinach Aptamer, Fluorescent Sensors, and G-Quadruplexes

### *1.4.1 An RNA Equivalent of GFP*

The impact Green Fluorescent Protein (GFP) has had on our understanding of molecular biological processes has been profound. With the advent of molecular cloning and recombinant DNA, GFP was elevated from an interesting trait of jellyfish tissue to an indispensable research tool for elucidating the fundamental principles governing subcellular protein trafficking, macromolecular complex function, subcellular organization, and organismal development.[145–150] GFP fluorescence is the result of the autocatalytic cyclization of $Ser^{65}$-$Tyr^{66}$-$Gly^{67}$, forming the fluorophore HBI (4-hydroxybenzlideneimidazolinone) which remains covalently incorporated within GFP's primary sequence, oriented in its excitable *cis*-isomer within the core of the protein's beta-barrel structure.[151–153] The potential value that could be obtained through a nucleic acid-based equivalent of GFP for the subcellular quantification and trafficking of endogenous RNAs was not lost on researchers, however no RNA or DNA sequence was known to produce fluorescent signals inherently. Linking GFP to sequence-specific RNA-binding proteins was an initially appealing option, however the necessary multi-protein

complexes are quite large, disrupting critical RNA-RNA/protein interactions.[154] Thus, *in vitro* evolution by SELEX was employed, to instead identify synthetic RNA sequences that directly bind and induce fluorescence of exogenous fluorophore ligands. Several aptamers were successfully identified, however their selectivity were poor and the fluorophores tended to be cytotoxic at high concentrations.[155–159] These issues were resolved with the identification of "Spinach" aptamer, which have reinvigorated interest in subcellular fluorescent investigation of RNA (Figure 1-4.A).[160]

### *1.4.2 Fluorescent Vegetable Soup*

The advantage of Spinach over previously selected alternative fluorescent aptamers is less related to any inherent property of the RNA sequences than it was to the character of the fluorophores they bound. Inspired directly by GFP, SELEX was performed against a synthetic analog of the GFP fluorophore HBI, DMHBI (3,5-dimethoxy-4-hydroxybenzylidene imidazolinone).[160] Conveniently, the resulting aptamer family not only bound and induced the fluorescence of DMHBI, but a small flight of DMHBI sister fluorophores.[160] These fluorophores all show low cytotoxicity at high concentrations, display high photostability and low background signal, bind Spinach variants with similar affinity, display individual excitation and emission wavelength combinations, and, finally, are conveniently compatible with the preset filters built into many optical instruments designed for use in GFP experimentation.[160–162] In just one SELEX experiment, a flight of fluorescent probes were identified that simultaneously paint most of the visible spectrum (Figure 1-4.B).

Due to the potential benefits of the Spinach aptamer, additional fluorescent aptamers with alternate or preferential features have been selected. Some have been identified through mutagenesis screening ("miniSpinach" and "Baby Spinach") or directed evolution of Spinach ("iSpinach", "Spinach2" and "Broccoli"), and others through completely independent SELEX ("Mango", "Corn", "Chili" and "DIRS2").[163–169] This family of aptamers have been used to investigate processes including subcellular localization of endogenous RNAs in bacterial and mammalian cells, viral RNA trafficking in infected cells, recruitment and dynamics of toxic CGG trinucleotide repeat-containing RNAs, formation and stability of large ribonuclear protein complexes, and cellular delivery of ncRNA by nanoparticles.[170–175]

Two independent crystallizations of Spinach in complex with DFHBI (3,5-difluoro-4-hydroxybenzylidene imidazolinone)—the brightest HBI derivative tested at the time—revealed the structural elements critical for its fluorescent function.[176,177] Spinach is composed predominantly of WC base pairing in a linear helical duplex, however an off center bulge is formed by an intricate, unique, mixed orientation, two-plane Guanine quadruplex motif (G-Quadruplex), upon which DFHBI docks in a *cis*-isomer stabilizing binding pocket (Figure 1-4.A).[178] Interestingly, with the exception of DIRS2, all of these aptamers utilize G-Quadruplex motifs to dock their fluorogenic ligands, exemplifying the inherent favorability of the G-Quadruplex motif and its capacity to imbue complex structure into otherwise simple primary sequence.[11] Unrelated selections of multiple fluorescent RNAs dependent on a G-Quadruplex additionally suggests the motif is more conducive for binding planar aromatic molecules than alternative nucleic acid structural

elements. This trend can be observed in the family of small-molecules that have been identified for their G-Quadruplex binding and stabilization in the therapeutic suppression of c-Myc expression and telomerase activity.[179]

### *1.4.3 Spinach Biosensors*

Spinach was selected with the original goal of subcellular visualization of endogenous RNA trafficking, for which it has clearly been put to good use. Even before to its structural elucidation, the potential application of Spinach for biosensor development was realized.[180] Sequence analysis of Spinach suggested the 5' and 3' termini ending in base paired duplex, conducive for the insertion of Spinach into alternate structures (Figure 1-4.A).[160]. Extant natural riboswitches display ligand-dependent conformational flexibility, and are generally utilized as 5'UTR-located translational regulators (a deeper discussion of riboswitches can be found in chapter 6).[38,181] Inserting Spinach into the labile ligand binding domain of a riboswitch destabilizes the Spinach duplex and G-Quadruplex domain, resulting in little or no background fluorescence. However, in the presence of ligand the riboswitch structure is stabilized, allowing the Spinach structure to form, promoting fluorophore docking and fluorescence.[182] This general approach has been utilized with several natural riboswitch structures. Riboswitches binding cyclic-di-GMP, cyclic-GMP-AMP, thiamine 5'-pyrophosphate (TPP), adenosine diphosphate (ADP), and S-adenosylmethionine (SAM) have all be engineered into ligand-dependent fluorescent biosensors via rational fusion with Spinach aptamer.[180,183–185]

Both Spinach and Broccoli aptamers have been engineered into split aptamers as well, allowing the observation of macromolecular assembly independent of a third-party

ligand interaction, in much the same way as split-GFP.[186] These generally remain in the proof of principle stage, observing nucleic acid assembly *in vitro* and *in cellulo*.[187–191] However, a few more complicated paradigms using the split aptamers have been performed to quantify ribozyme kinetics, validate synthetic nanoparticle assembly, or generate output signals of nucleic acid logic circuits.[192–196] Structural characterization of the original Corn aptamer revealed it to be a homodimer, with its fluorophore DFHO (3,5-dimethoxy-4-hydroxybenzylidene imidazolinone-2-oxime) sandwiched between the G-Quadruplex stacks of the monomers.[167] Interestingly though, specific interfacial mutations were tolerated yielding multiple complementary Corn pairs that only induce DFHO fluorescence when engaged in specific *hetero*dimeric binding. While Corn has yet to be widely exploited, the capacity to observe multiple association events simultaneously is promising.[197,198]

The aptamers in this family all bind exogenous fluorophores non-covalently, inducing fluorescence upon proper folded or specific dimerization. Their native RNA chemistry and nontoxic fluorophores allow *in cellulo* biosensing using endogenous expression machinery. Additionally, the dynamic binding and unbinding of isomerized HBI-derivative fluorophores provides Spinach family aptamers substantial resistance to photobleaching compared to GFP or other aptamer sensors.[161] Thus, these RNAs represent an essential fourth class of optical aptamer biosensors (section 1.2.2), distinct from those whose fluorescent capacity is dependent on modified nucleotides or the covalent attachment of fluorophores via *in vitro* evolved polymerases or organic synthesis.

### *1.4.4 The Tripartite Complex of Spinach*

The solution of Spinach structure revealed the critical G-Quadruplex motif is stabilized by a single potassium ion, in agreement with both expected cation identity and coordination stoichiometry for two-plane G-Quadruplexes in physiological conditions.[176,177,199–201] Successive investigations with the Spinach G-Quadruplex reveal the structure to have favorable coordination with strontium and lead as well as potassium, all three of which with significant preference over sodium and other cations tested (*unpublished).*[202] This is in agreement with cation preferences observed for other RNA and DNA G-Quadruplex topologies (generally $Sr^{2+}$>$Pb^{2+}$>$Ba^{2+}$>>$Ca^{2+}$>$Mg^{2+}$ for divalent cations and $K^{+}$>>$Rb^{+}$>$Na^{+}$>$Li^{+}$ for monovalent.[199] Studies with various cations have all confirmed Spinach fluorescence requires the association of all three constituent components. This paradigm—a tripartite complex—is composed of correctly folded RNA, coordinating one of several G-Quadruplex-stabilizing cations, docking a compatible HBI-derivative fluorophore (Figure 1-4.A/D). The absence or extraction of any one component results in a breakdown of tripartite complex and a loss of fluorescent signal (Figure 1-4.D).

## 1.5 G-Quadruplexes and Cation Sensing

### *1.5.1 G-Quadruplex-Cation Interactions*

The three interacting nucleobase edges (Watson-Crick, Hoogsteen, and Sugar), of all nucleobases in DNA and RNA, engage in *inter*- and *intra*- molecular H-bonding and form the basis of nucleic acid structural specificity (Figure 1-2.B and Figure 1-5.A).[203,204] Each edge can interact with the other two, and can do so in a *cis*- or *trans*- orientation that is

dependent upon base orientation around the glycosidic X bond (Figure 1-2.B). This yields twelve orientations in which two nucleobases can engage.[205] Multi-base combinations of edge interactions, aided by the conformational flexibility allotted to nucleic acids (section 1.1.1) (Figure 1-2), favor the formation of specific secondary structure elements within a nucleic acid sequence (e.g. the classic helical duplex is the product of a string of *cis*-WC/WC interactions). Other common motifs include hairpins, pseudoknots, kissing loops, three-way and four-way junctions (3WJ and 4WJ, respectively), triplexes, i-motifs, and quadruplexes. G-Quadruplexes are planar motifs formed by four guanines engaged in Hoogsteen-WC face H-bonding, and form spontaneously in guanine-rich sequences *in vivo* and *in vitro*.[206] G-Quadruplexes are often found in multi-plane stacks stabilized by aromatic $\pi$-$\pi$ interactions, with topologies that are qualitatively categorized by the directionality ("parallel, antiparallel, or mixed") and connectedness ("chair, basket, hairpin, hybrid, or unique") of their backbones.[207] G-Quadruplexes are both physiologically and therapeutically relevant structures, directly influencing gene expression and forming the structural basis of telomere DNA.[208,209] Their generalized cellular utility for expressional repression of several proto-oncogenes (notably kRAS, BCL2, c-KIT, and c-MYC) has inspired the development of G-Quadruplex-stabilizing small-molecule drugs for the treatment of various cancers.[210–214] Counterintuitively, however, active deconstruction of spontaneous non-telomeric G-Quadruplex formation appears to be a high priority of eukaryotic cells.[215,216] All of this highlights the motif's inherent favorability, and suggests a biological significance and depth of complex regulation that remains to be fully appreciated.

Overall G-Quadruplex geometry is balanced between favorable $\pi$-$\pi$ stacking, Hoogsteen-WC face H-bonding, symmetrical hydration structure, and repulsive electrostatics (polar carbonyls and negative phosphates). The balance of forces results in consistent structural geometry, independent of specific motif topology. O6 oxygen atoms of each of four guanines orient inward in a square planar arrangement with a consistent 30$^{o}$ twist and 3.3A rise between G-Quadruplex stacks (Figure 1-5.A).[207,217] As with any nucleic acid structure, the bulk of cation interactions are with the phosphate backbone. However, all G-Quadruplex structural investigations have shown specific cation coordination of guanine O6 atoms within the central pore to be essential to motif stability, contributing more free energy of interaction than H-bonding and $\pi$-$\pi$ stacking combined.[199,218,219] A handful of cations are now known to stabilize G-Quadruplexes, with binding preference dependent upon a series of elemental characteristics including atomic radii, electron valence, solvation energy, and apparent coordination number preference. Numerous comparative studies of $K^+$ and $Na^+$ binding in G-Quadruplexes have been explored as the two elements are the most highly concentrated cations in physiological conditions. $Na^+$, with an ionic radius of 0.95A, can move freely through the central pore and exist coplanar with an individual stack. Alternatively, $K^+$, with a radius so 1.33A, is restricted to inter-plane space, equidistant from each stack, binding in an *n*-1 stoichiometry. The freedom of $Na^+$ movement through the pore theoretically allows for enhanced quenching of O6 electronegative repulsion.[220] However, thermodynamic studies have revealed the driving force of $K^+$ preference is the higher energetic cost of $Na^+$ dehydration over $K^+$ dehydration, conceptually similar to the mechanisms of cation

filtration in voltage gated potassium channels.[221–223] $K^+$ preference is bolstered by studies suggesting $Na^+$ generally engages in fewer coordination numbers than $K^+$ in water, G-Quadruplexes, and channel selectivity filters.[224] These studies project both cations to form coordination numbers fewer than eight, suggesting they engage in dynamic equilibrium between the eight O6 guanine carbonyls within a stable G-Quadruplex complex. Additional divalent cations such as strontium ($Sr^{2+}$), lead ($Pb^{2+}$), and barium ($Ba^{2+}$) coordinate G-Quadruplexes more favorably than $K^+$, also reminiscent of potassium channel selectivity.[225] Each possess the favorable dehydration energetics and ionic radii alongside higher coordination numbers and the addition of 2+ charge neutralization within the core, greatly strengthening their binding G-Quadruplex structure and enhancing G-Quadruplex thermal stability.[226–229] Electrostatic repulsion between these divalent cations is enough to result in binding only at every other inter-plane space, in an n/2 stoichiometry.[230] Cation dependent G-Quadruplex formation and stability has been studied using many common biophysical techniques, however circular dichroism (CD) spectroscopy and UV-absorbance spectroscopy at 295nm are particularly convenient and widely employed for direct observation of G-Quadruplex structure.[231–233]

### *1.5.2 Example Cation Sensors*

In addition to their biological importance, G-Quadruplex motifs are disproportionately represented within the canon of structurally characterized aptamers. Thrombin Binding Aptamer (TBA) is *the* quintessential G-Quadruplex model aptamer, identified by SELEX to bind and inhibit human a-thrombin.[234–236] The minimalist palindromic DNA aptamer is just fifteen nucleotides in length (5'-GGTTGGTGTGGTTGG),

utilizing eight residues to form an antiparallel chair-type two-plane G-Quadruplex (Figure 1-5.B/C). While the aptamer binds its intended protein target with impressive affinity (40nM), TBA has garnered significantly more research attention as a model structure for investigating cation-dependent G-Quadruplex folding. Aside from the general chemical conveniences of short nucleic acids, several TBA-specific traits make it an attractive model: (i) in the absence of select G-Quadruplex-stabilizing cations TBA is completely unstructured; (ii) its folding is rapid, two-state, and reversible; (iii) it maintains a 1:1 G-Quadruplex:cation stoichiometry; (iv) its solution and crystal structure has been solved several times, and is amenable to additional structural studies including *in silico* modelling; and (v) its 5' and 3' termini end together, interacting in the top G-Quadruplex plane.

In addition to providing insight into G-Quadruplex formation and stability, this combination of traits—particularly the last one—has facilitated TBA construction and use as a quantitative biosensor for the cations it selectively coordinates. Takenaka and colleagues were the first to realize G-Quadruplex structures are an ideal platform for engineering cation sensors. Inspired by the termini-linked fluorophore/quencher, FRET, and excimer pairs of the then recently conceived dual-reporter aptamer beacons (section 1.2.2), they generated a set of fluorescent G-Quadruplex sensors.[54] The first attempts at developing selective $K^+$ sensors utilized minimalist human (5'-$(GGGTTA)_3GGG$) or Oxytricha (5'-$T_4G_4)_3GGG$) telomeric sequences, coined "PSO-1" and "PSO-2", respectively, and were visualized by FRET using terminal conjugation of 6-carboxyfluorescein (6-FAM) and 6-TAMRA donor and acceptor pairs.[237–239] PSO-1 displayed a narrow linear sensitivity

to from ~0.1-1mM $K^+$, while PSO-2 displayed weaker but broader sensitivity from ~5-20mM. Interestingly, both sensors showed sensitivity to competing $Na^+$, with CD suggesting a bimolecular homodimer G-Quadruplex structure formed upon $Na^+$ binding. To avoid this cation-dependent conformational heterogeneity, Takenaka and colleagues turned to TBA, whose short sequence length restricts its fold to its native antiparallel chair topology.[240,241] Replacing the FRET pair with a pyrene excimer pair (coined "PSO-py") enhanced signal-to-noise while retaining linear detection from ~1-10mM $K^+$, bringing the TBA sensor to its more or less final form (Figure 1-5.D).[242]

The success of cation sensing with TBA has spread from the Takenaka group, in some cases quite interestingly, while others require the reader to expand their definition of the word "novel". Fluorescence based $K^+$ sensing in conceptually similar manners with terminal fluorophores have been successfully executed.[243] TBA sensors have been used to visualize subcellular $K^+$ flux via cholesterol conjugation, and used for the first on-chip aptamer-based electrochemical $K^+$ detection (see section 1.2.4).[244,245] As expected based on generalized G-Quadruplex properties, TBA displays preferential binding to divalent cations $Sr^{2+}$, $Pb^{2+}$, and $Ba^{2+}$ with faster cation-association kinetics, increased melting temperatures, and subtly compacted quadruplex structure.[246,247] Fluorescently and GNP labelled TBA sensors have been demonstrated to detect $Pb^{2+}$ with detection limits in the nanomolar to low micromolar range.[248] Success with TBA has inspired additional G-Quadruplex-based cation sensing, much of which has been performed with alternative idealized, non-telomeric, guanine rich sequences.[249,250] These investigations have generally utilized a similar scheme of terminally conjugated FRET pairs, however reports

have been presented that exploit the preferential binding of specific small-molecule fluorophores to folded G-Quadruplex structures for cation sensing.[251–255] These last tools depend on the binding—and binding-induced fluorescence—of exogenous fluorophores non-covalently, forming cation-dependent tripartite complexes of their own similar to Spinach family aptamers.

## 1.6 Spinach as a Fluorescent Cation Sensor & *Post*-SELEX Aptamer Structural Recycling

### *1.6.1 Spinach and Lead(II)*

Spinach has been used as a split and fusion partner with alternate RNA structures for the fluorescent sensing of the presence of the ligands of those alternate RNAs, as discussed. For Spinach fluorescence to be used this way with accuracy, all three components of Spinach's own noncovalent tripartite complex must be present in constant concentrations. In intracellular assays with high concentrations of potassium and exogenously added DFHBI, this is a condition easily met, coincidentally and conveniently for researchers who tend not to discuss this potential limitation of fusion or split Spinach sensors. If, however, Spinach is not fused to another RNA for the detection of an alternate ligand, the non-covalent nature of tripartite complex formation is not a limitation but a potential boon. All three components, folded RNA, cation, and fluorophore, are required for the generation of fluorescent signal. However, holding just two of the three components constant allows for the detection of variable concentrations of the third, measured by fluorescent signal upon proper tripartite complex formation. *DasGupta et al* (2015) were the first to recognize this potential. Based on the general trends of

preferential G-Quadruplex-cation interactions, they investigated whether Spinach would both bind $Pb^{2+}$ and produce a fluorescent tripartite complex with DFHBI.[202] Fluorescence was observed, and G-Quadruplex formation was confirmed by CD. Taken one step further, they then asked if Spinach could be used to sense unknown $Pb^{2+}$ concentrations. Spinach displays a clear preference for $Pb^{2+}$ binding even in mixed metal solutions, and produces a fluorescent signal in $Pb^{2+}$ solutions as low as 10nM.[202] This report provided the proof-of-principle that the native sequence of Spinach, by the nature of its induced fluorescence of DFHBI, could be exploited for cation sensing without any chemical or structural modification.

### *1.6.2 In vitro Utility and Chemical Instability*

The work by *DasGupta et al* (2015) was focused in scope and discussion on $Pb^{2+}$ sensing.[202] However, within the context of additional previous successes of DNA G-Quadruplex cation sensing (most notable by Takenaka and colleagues), $Pb^{2+}$ sensing by Spinach opened the door broadly to direct *in vitro* sensing utility by this fluorescent RNA family. Unaddressed by *DasGupta et al* (2015), however, are any complications relating to the chemical instability of RNA that will inevitably arise when using RNA to sense a component of unbuffered and uncontrolled environmental solutions *in vitro*, principally alkaline sensitivity and the ubiquitous presence of stable ribonuclease enzymes.

Perhaps more concerning, however, is the chemical instability of RNA in the presence of $Pb^{2+}$, a phenomenon mechanistically described decades ago.[256,257] Similar to common enzymatic or alkaline RNA hydrolysis, $Pb^{2+}$-dependent degradation is attributed to the reactive 2'OH; deprotonation of the hydroxyl induces an auto-nucleophilic attack,

creating 2',3'-cyclic phosphate and 5'-OH cleavage products.[258,259] The regularity of this cleavage has been leveraged into a low resolution foot-printing assay for the elucidation of promiscuous cation binding pockets, or tertiary structural compaction more generally with preferential cleavage observed at loops, bulges, and single-stranded regions.[258,260–265] $Pb^{2+}$ foot-printing has additionally been used to observe differential flexibility resulting from competing protein, small-molecule, or alternate cation binding.[259] While undiscussed by *DasGupta et al* (2015), it is inherently paradoxical to perform structure-dependent sensing of an analyte with a sensor that is degraded by that very analyte being sensed. That is not to say Spinach cannot be *made into* a viable *in vitro* cation sensor, even for $Pb^{2+}$.

### *1.6.3 Post-SELEX Modifications for Enhancement of Original Function*

As might be expected, the majority of chemical modifications of aptamer *post*-SELEX are incorporated with the goal of enhancing the aptamer's original function, indirectly through increased chemical stability, or directly through increasing binding affinity. Successful examples range from a single modification to a complete chemical overhaul. One specific phosphorodithioate linkage modification at position 18 of another anti-α-thrombin RNA aptamer ("toggle-25t") increased it's affinity from ~2nM to ~2pM.[266] The same single modification increased anti-VEGF RNA aptamer affinity for its target protein by a similar ~1000-fold. Structural studies revealed in both cases the increased affinity was not due to a structural rearrangement in the aptamer, but a favorable S/$\pi$ interaction with an aromatic sidechain of the protein ligand. The alternative extreme case of modification has been performed, most notably with the aforementioned anti-$VEGF_{165}$

aptamer Pegaptanib, the quintessential example of *post*-SELEX chemical modification for functional enhancement (section 1.3.4). The aptamer was originally selected with 2'F pyrimidine nucleotides, but was further modified *post*-SELEX at all but two positions with 2'OMe purines, a 3'-3' terminal cap, and a 40kDa polyethylene glycol 5' cap. The result is an aptamer retains the sequence, structure, and affinity of its original selection while acquiring extreme resistance to all *endo/exo*-nucleases and renal clearance, resulting in an effective intra-vitreal therapeutic.

(Dave says "Discuss this") Bivalent and multivalent aptamer fusions are a less explored alternative approach to enhancing the original function, typically, though not exclusively, relegated to protein-binding aptamers. Fusions are engineered by extending the sequence of one into the other, forming a single transcript that when folded performs twice the function per stoichiometric unit. In cases where a protein ligand functions as a dimer or is membrane bound and thus restricted to spatial movement in two dimensions, favorable binding of a multivalent fusion aptamer via increased avidity can result in dramatically enhanced function. This approach has been employed to great effect for aptamers that target surface presented immunological proteins such as CTLA-4 and B-cell.[165,267–269] A subtle twist on this concept has been described, wherein two different aptamers that target two different epitopes of the same protein ligand are ligated together to enhance overall ligand binding through similarly enhanced avidity. This has been performed by fusing two different anti-thrombin aptamers.[270,271]

A separate and more intriguing fusion class joins two different aptamers, each retaining their original function, to produce a bivalent fusion with an emergent combined

function. The most nuanced use of these fusions are *in vivo* aptamer-controlled gene expression logic gates, most effectively executed by Cai and colleagues.[272] In this project three native RNA aptamers--which selectively bind eIF4G, theophylline, or tetracycline--were structurally manipulated to create ligand-binding dependent mRNA expression circuits that function by promoting or suppressing ribosomal recruitment to endogenous genes in mammalian cells.[39,272–275] Depending on the particular combination of sequence modification, binary logic gates were engineered that respond to the same ligands but can either upregulate or downregulate gene expression. The set of logic gates designed included AND, OR, NOT, NOR, XOR, NAND, and XNOR. The aptamers' ligand binding domains were held constant, and various 3' extensions were incorporated, complementary to 5'UTR sequences in the genes regulated, complementary to the extension on a different co-expressed aptamer, or both complements combined. The circuit design takes advantage of the ligand-dependent conformational stability of the tetracycline and theophylline aptamers, whose 3' extensions are differentially exposed upon ligand binding.[39,272,274] Descriptions of how these regulatory aptamer fusion constructs differentially promote or suppress expression in the different logic gate configurations can be found in the Figure 1-6. What is important to take away is how aptamer utility was altered by sequential modification. Native chemistry ribonucleotides were strategically incorporated into previously characterized aptamers, retaining original function while engendering new utility that is the product of their bivalent fusion, all without resorting to *de novo* selection of new aptamer structures.

### *1.6.4 Structural Recycling: Facilitation of Novel Function*

In contrast to a modification for enhancement of original function is a modification in pursuit of novel function that a particular aptamer was not originally selected to perform. Considering the size, structural simplicity, and binding selectivity commonly associated with aptamers, the capacity to perform two specific functions without intentional selection or sequence optimization is a rare aptamer trait, and one that is open to subjective interpretation.[11] This idea is perhaps best explained by example. The first, and heretofore *only,* examples of the use of modified nucleotides to reveal an alternate inherent aptamer function was that of terminal guanine-conjugated fluorophore-quencher, FRET, and excimer pairs of TBA for $K^+$ detection from Takenaka and colleagues, novelty which went unrecognized and undiscussed by the original authors.[240–242] TBA was selected to bind one ligand ($\alpha$-thrombin), but coincidentally requires specific binding of a second ligand. That second function is specific cation coordination by a G-Quadruplex motif integral to aptamer structure and upon which the intended primary function depends. By incorporating fluorophore- and excimer-conjugated nucleotides, Takenaka and colleagues leveraged specific cation binding into true cation sensing, opening up an entire functional paradigm for TBA that is completely independent of its originally selected function (the leveraging of cation binding has itself now branched into two separate functional uses of TBA, fluorescent sensing and G-Quadruplex modelling).

*DasGupta et al* (2015) revealed Spinach family aptamers to have a two-function capacity through their investigation of $Pb^{2+}$ fluorescent detection as—similar to TBA—Spinach aptamer was selected to specifically bind a fluorophore, not a cation.[160,202]

Acknowledgement of this fact was similarly unrecognized and undiscussed by the authors. It is entirely coincidental Spinach displays specific cation binding, and ever more serendipitous, in this case, that cation binding is essential for fluorescent emission. It is interesting, though perhaps not surprising that both TBA and Spinach secondary function relies on a two-plane G-Quadruplex motif for specific cation binding (section 1.7.1). It is very likely all G-Quadruplex aptamers can be exploited this way; for specific cation sensing, it seems likely G-Quadruplexes are the ideal (and perhaps only) generalized structural motif.

When properly formed, the tripartite complex of Spinach is inherently fluorescent, facilitating cation sensing and exploitation of secondary Spinach function without chemical modification. Can, however, a structural mimetic chemical modification be used in Spinach to enhance this secondary function? As discussed (section 1.7.2), it is inherently problematic to use a native chemistry RNA structure to observe $Pb^{2+}$ concentration, unless degradation *is* the readout. Commitment to *in vitro* application of a Spinach sensor opens up options of nucleotide modifications incompatible with endogenous cellular transcription machinery, which may be a viable solution to the paradox and $Pb^{2+}$ sensing by RNA. Numerous nucleotide modifications exist to increase the chemical stability of an RNA without (ideally) perturbing the overall structure (section 1.3.4), presumably a requirement in the case of Spinach. For a chemically modified nucleotide to be dropped in to the sequence of Spinach to perform this task, it should meet the following criteria: (i) be positioned at C2, stabilizing the construct against $Pb^{2+}$ and nuclease degradation; (ii) retain the favored *3'endo* pucker of native chemistry

ribose; (iii) perhaps not be located on guanine riboses to minimize any chance of disrupting G-Quadruplex formation; and (iv) be readily incorporated by RNA polymerase transcription for convenient and cost effective synthesis. Can Spinach aptamer be chemically modified without a loss of structural integrity? If so, with what modification(s), and to what degree of construct coverage? Can other structurally related members of the Spinach family be modified? Are the same modifications compatible? Are there unforeseen utilities of chemical modifications that *do* alter Spinach structure?

## 1.7 Dissertation Outline

This dissertation focuses on the effects of a specific nucleotide modification on cation association and tripartite complex formation in a highly related fluorescent aptamer family, how those effects can be leveraged for the development of RNA-based fluorescent cation sensors, and the implication of direct chemical modification for the discovery of novel aptamer function. In Chapter 2, the results of Spinach aptamer modification with 2'F pyrimidine nucleotides are discussed. The modification alters both Spinach's chemical stability in the presence of ribonuclease and high concentrations of $Pb^{2+}$, and Spinach's rate of tripartite complex formation with $Pb^{2+}$ and fluorophore DFHBI, assayed by the resulting fluorescent emission. In Chapter 3, the sequence relationship between Spinach aptamer and its highly related sister construct, Broccoli, is explored. Both constructs share substantial sequence identity, however subtle differences result in increased thermal stability and accelerated tripartite complex formation in Broccoli. A theoretical structural model is proposed therein, facilitating hypotheses for addressing

the discrepancies between Broccoli and Spinach tripartite complex behavior. Inspired by these differences, the same modified nucleotides (2'F pyrimidines) are incorporated into Broccoli aptamer. The resulting tripartite complex kinetics and apparent $K^+$ affinity are altered unexpectedly, and conducive to the development of a mixed chemistry chimeric fluorescent $K^+$ sensor. This work is presented in Chapter 4. Chapter 5 briefly summarizes the cumulative work presented in Chapters 2-4, and provides discussion of the work's implications for future cation sensor development using the modified Broccoli Chimera specifically and Spinach family aptamers generally. Chapter 6 of this dissertation draws upon the implications of structural recycling of extant aptamer structures for the discovery of novel function, inspired by the novelty of direct chemical modification of Spinach and Broccoli aptamers for fluorescent cation sensing. This discussion touches on mechanisms of natural evolution as compared to *in vitro* selection by SELEX, the merits of motif recycling, and the value of techniques that democratize and accelerate aptamer selection for the development of research, diagnostic, and therapeutic tools. Chapter 7 contains the published manuscripts resulting from the work presented in Chapters 2-4 of this dissertation, with the final formatting of their respective journals of publication.

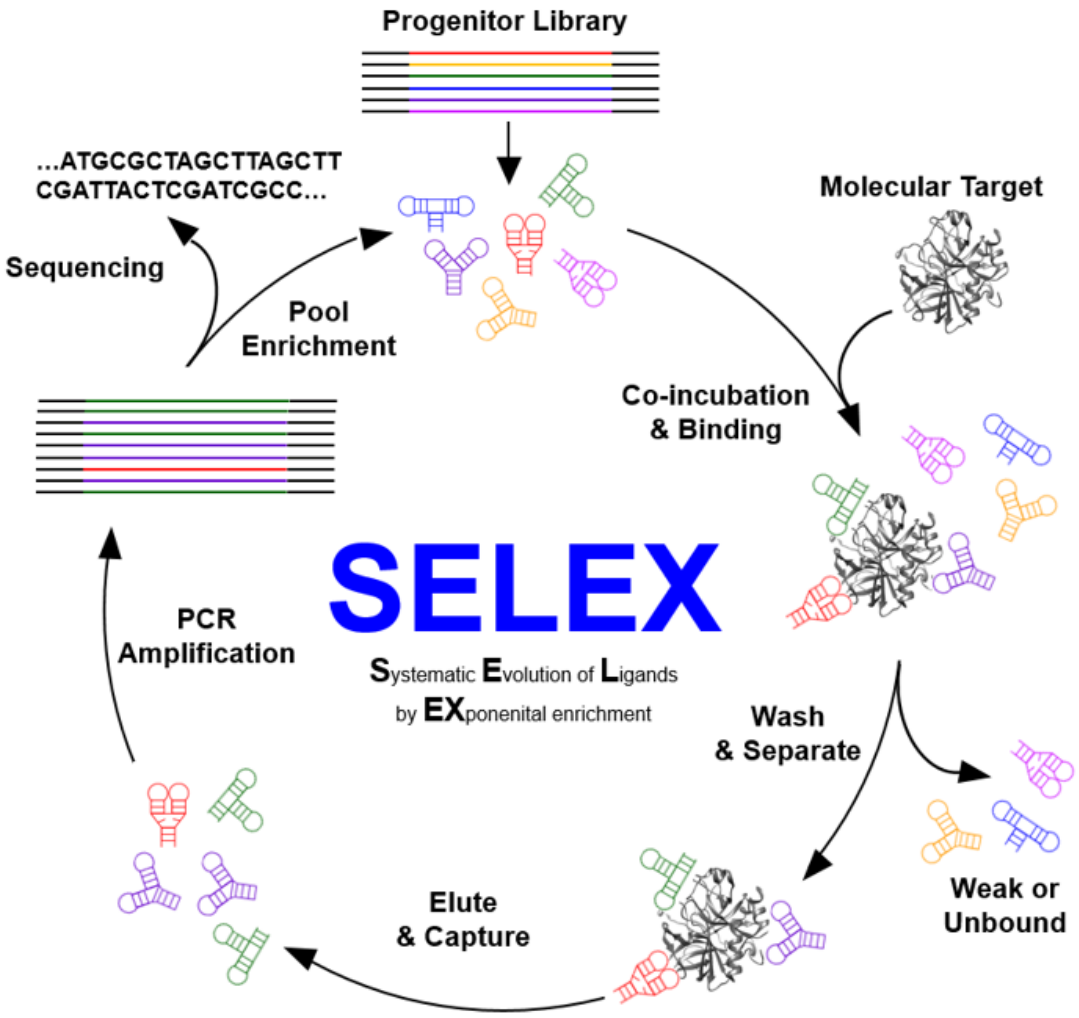



**Figure 1-1 SELEX Overview for Identification of Protein-Binding Aptamers**. A highly heterogeneous progenitor library of ssDNA or RNA structures is folded and co-incubated with a molecular target, in this case a protein. The technique relies on the existence of a small population of the initial pool displaying favorable binding by the inherent nature of its sequence. These few individuals favorably bound to the target molecule are separated from the weakly or unbound population, often by use of agarose beads or nitrocellulose filter binding. The entire population of captured aptamers are unbiasedly amplified by (reverse transcription) PCR. The resulting population displays considerably more homogeneity than the progenitor library. Successive passages, each with increasingly stringent selective pressure (e.g. reduced target molecule availability or increased temperature or salt), drive the population towards further and further homogeneity, until population sequencing yields few enough unique species to assay for binding.

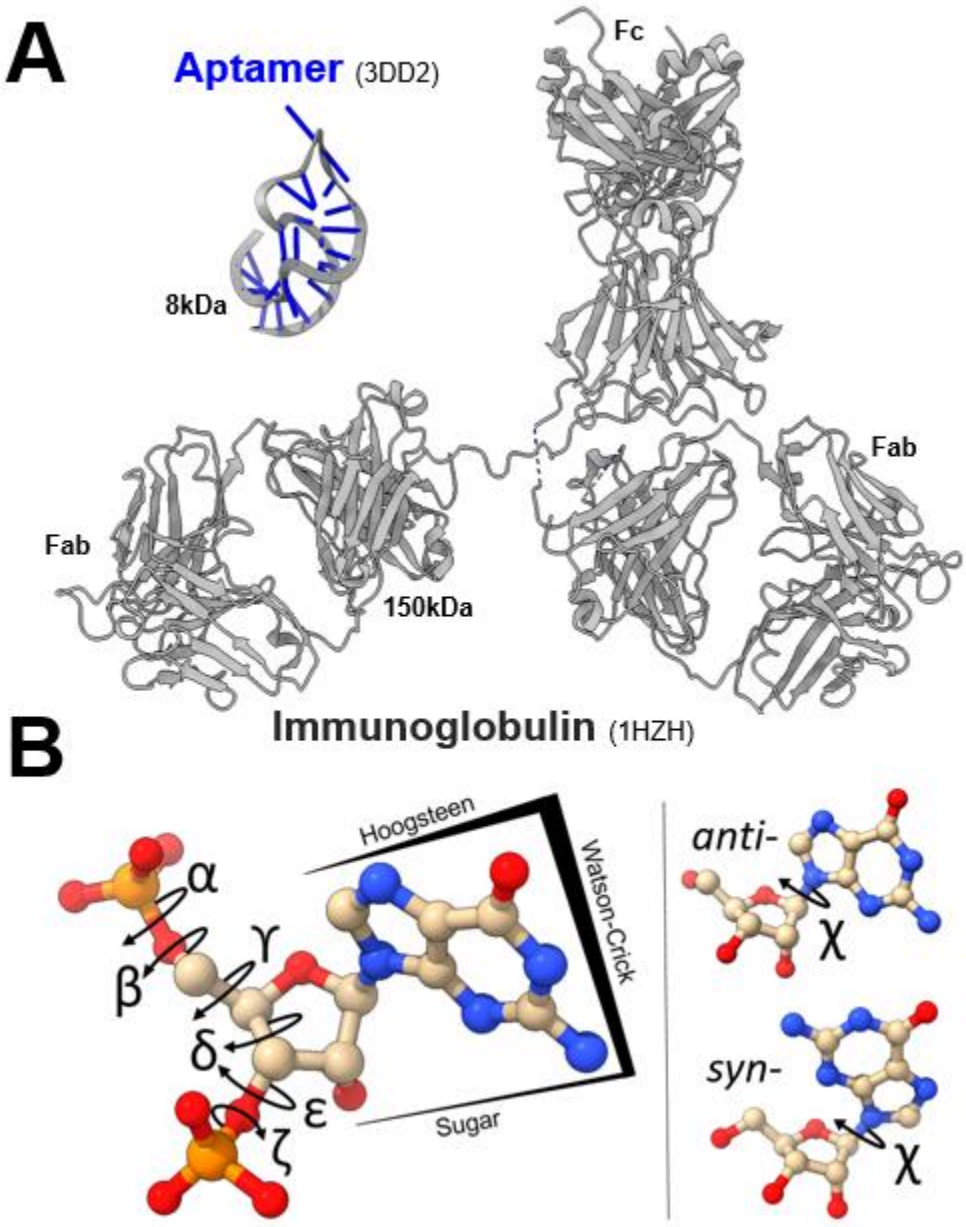



**Figure 1-2 Aptamer Structure and Conformation.** (A) A scaled size comparison of 8kDa aptamer ("toggle-25t" PDB 3dd2) compared to a 150kDa human antibody (IgG, PDB 1hzh). (B) *Left:* Individual torsion angles along the nucleic acid backbone are labelled, from alpha to zeta. The representative purine nucleobase edge nomenclature is similarly labelled. *Right:* Nucleobases can rotate freely around their single glycosidic bond, x. Extremes are shown, *anti-* and *syn-*.

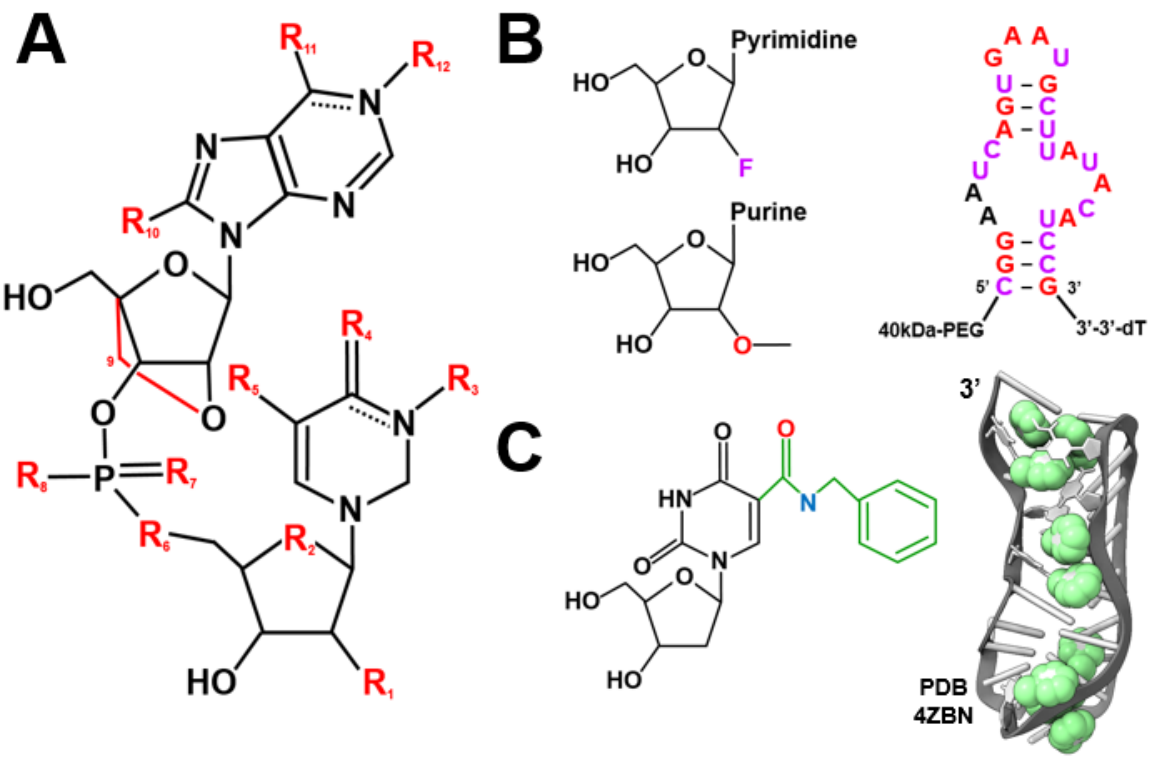



**Figure 1-3 Chemical Modifications of Nucleotides.** (A) Predominate locations of chemical modifications. Two nucleotides, generic purine and pyrimidine, are labelled with R1-12 showing the common locations of chemical modifications including the base, sugar, and phosphate backbone. (B) *Left*: Drawing of generic 2'-fluoro-2'-deoxypyrimidine and 2'-fluoro-2'-deoxypurine pair showing the C2 modified functional groups "2'F" and "2'Ome". *Right:* 2D representation of anti-$VEGF_{165}$ aptamer "Pegaptanib". Pink pyrimidines and red purines are modified with 2'F and 2'OMe, respectively. 5' PEGylation and 3'-3' dT cap are labelled. (C) *Left*: Drawing of 5-N-benzylcarboxamide-deoxyuridine (Bn-dU) utilized in numerous SOMAmer structures. *Right*: Cartoon 3D structure of crystallized anti-NGF (nerve growth factor) (PDB 4zbn). Space filled green functional groups correspond to the green benzylcarboxamide group in Bn-dU nucleotide.

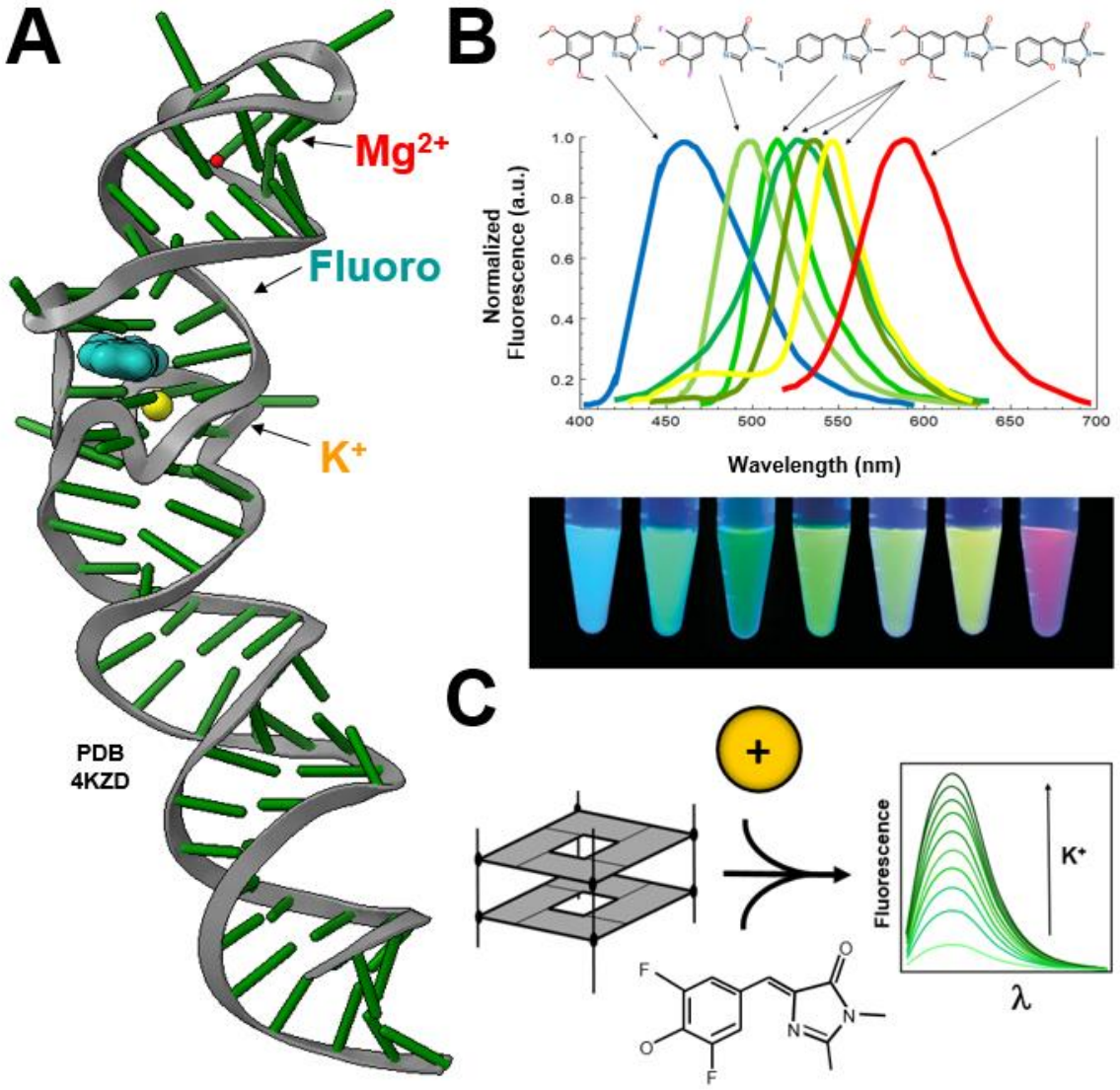



**Figure.1-4 Spinach Aptamer.** (A) 3D Cartoon representation of Spinach crystal structure (PDB 4kzd). RNA (green and grey) binds K+ (yellow) at a critical G-Quadruplex motif when properly folded, upon which DFHBI fluorophore (cyan) binds. (B) *Top:* Spectral overlap, in respective corresponding visible color, of emission profiles of HBI-derivative fluorophores in complex with Spinach RNA aptamer derivatives. *Bottom*: Image of the same Spinach-fluorophore combinations in microtubes, excited at their respective maximal wavelengths. (Adapted from Paige *et al* 2011). (C) The tripartite complex of Spinach family aptamers is composed of folded RNA G-Quadruplex, cation, and fluorophore. The proper association of these components produces fluorescent signal. Holding RNA and fluorophore components at constant concentrations allows for fluorescent detection of cation by emission upon tripartite complex formation.

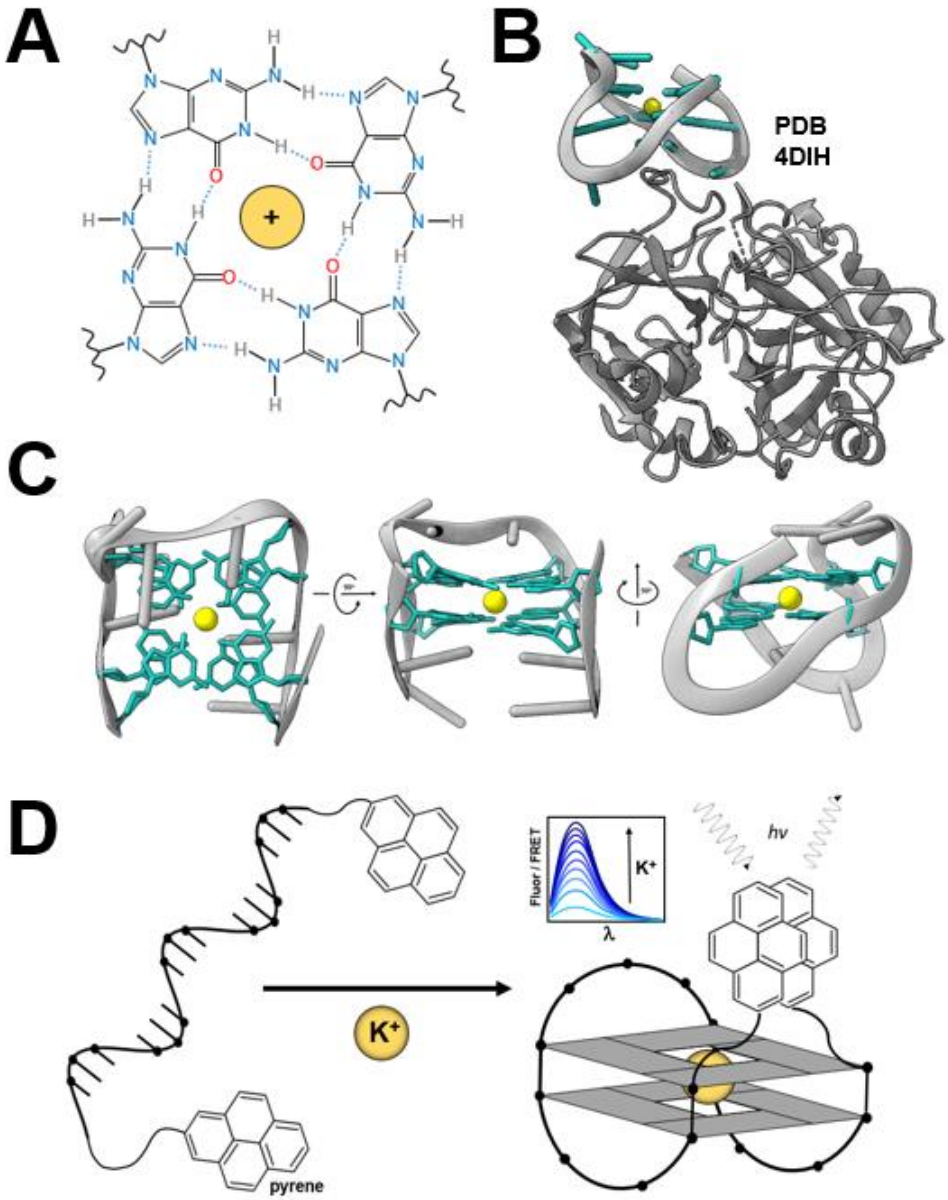



**Figure 1-5 G-Quadruplex Structure, TBA, and Functional Reprogramming.** (A) Chemical schematic of canonical G-Quadruplex structure, showing Hoogsteen face H-bonding, and location of central pore coordinating monovalent cation (yellow). O6 oxygen are red, purine nitrogen are blue, h-binding hydrogens are grey. (B) 3D cartoon representation of TBA in complex with alpha-Thrombin (PDB 4dih). Protein backbone is grey, nucleotides in aptamer are cyan, potassium ion in yellow. (C) TBA G-Quadruplex structure (PDB 4dih) coordinating K+, observed from three perspectives. G-Quadruplex guanines represented in cyan, potassium in yellow. (D) Basic outline of TBA-based K+ sensor function. IN the absence of cation TBA is unstructured, separating termini-linked pyrene excimer pair. Presence of K+ induced G-Quadruplex fold, binding together pyrene pair, allowing aromatic stacking and fluorescent detection proportional to K+ concentration from 2-10mM.

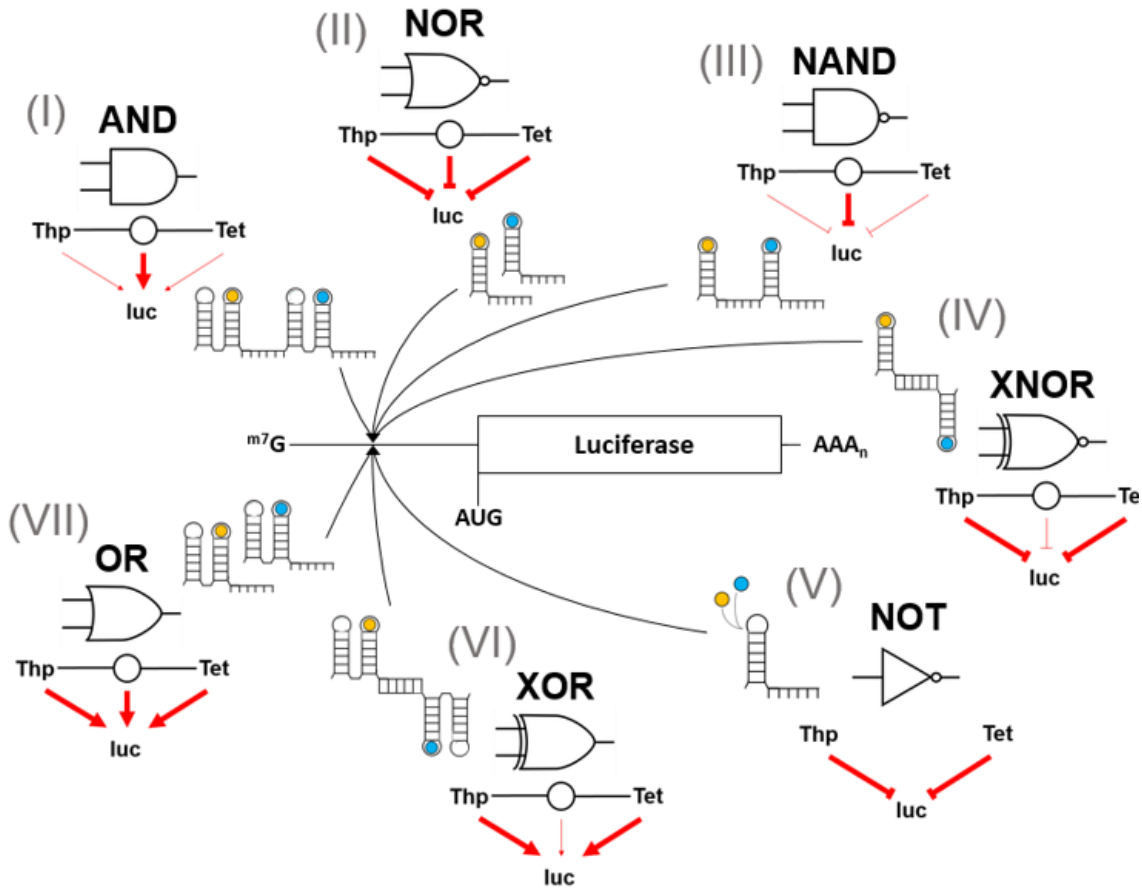



**Figure 1-6 Multivalent Aptamer Fusions *Post*-SELEX.** Three aptamers are used in combination, selectively binding eIF4G, theophylline, or tetracycline. Upregulation of reporter expression above baseline is dependent upon recruitment of eIF4G aptamer to 5'UTR, binding endogenous eIF4G and recruitment of ribosome above baseline. Downregulation below baseline is dependent upon recruitment of theophyilline or tetracyline aptamers (either in isolation or as a fusion construct) to the 5'UTR, inhibiting baseline ribosomal template reading. Blue dot represents theophylline, yellow dot represents tetracycline. Ligand binding induces conformational alteration in the respective aptamer, changing the exposure of the 3' extension sequence, thus altering favorability of base pair with different complementary sequence partners. (I) AND gate (Upregulation); if either drug is present, reporter is baseline; if both drugs are present, reporter is promoted. (II) NOR gate (Downregulation); if either drug is present, reporter is suppressed; if both drugs are present, reporter is suppressed. (III) NAND gate (Downregulation); If either drug is present, reporter is baseline; if both drugs are present reporter is suppressed. (VI) XNOR gate (Downregulation) If either drug is present, reporter is suppressed; if both drugs are present, reporter is baseline. (V) NOT gate (Downregulation); If either drug is present, reporter is suppressed. (VI) XOR gate (Upregulation); If either drug is present, reporter is promoted; If both drugs are present, reporter is baseline. (VII) OR gate (Upregulation); If either drug is present, reporter is promoted; if both drugs are present, reporter is promoted. Figure adapted from *Liu et al*, eLife, 2018.

# 2

# A Ribose Modification of Spinach Aptamer Accelerates Lead(II) Cation Association *in vitro*

This chapter has been reformatted from its original publication in *Chemical Communications*:

**Savage, J. C.**; Shinde, P.; Bächinger, H. P.; Davare, M. A.; Shinde, U. A Ribose Modification of Spinach Aptamer Accelerates Lead(II) Cation Association in Vitro. *Chem. Commun. (Camb).* **2019**, *55* (42), 5882–5885. https://doi.org/10.1039/c9cc01697j.

J.C.S. and U.S. designed the research; J.C.S. performed all experimentation; P.S. assisted in experimental execution; H.P.B. and M.A.D. provided thoughtful discussion and manuscript editing; J.C.S. wrote the paper; J.C.S. and U.S. edited and strategized the paper.

## 2.1 Abstract

Spinach aptamer fluorescence requires formation of a tripartite complex composed of folded RNA, a GFP-like fluorophore, and selective cation coordination. 2'F pyrimidine modified Spinach has retained fluorescence, increased chemical stability, and accelerated cation association via increased G-quadruplex dynamics, thereby reducing readout time and enhancing Spinach utility for aqueous $Pb^{2+}$ detection.

## 2.2 Discussion

The fluorescent RNA aptamer "Spinach" contains two stacked G-quadruplex planes stabilized by cation binding (Figure 2-1A).[176,177] Spinach was selected to bind 3,5-difluoro-4-hydroxybenzylidene imidazolinone (DFHBI), a synthetic derivative of green fluorescent protein chromophore, using the aptamer selection technique SELEX: systematic evolution of ligands by exponential enrichment.[5,6,160] Two structural studies revealed a unique non-parallel G-quadruplex that is coordinated with $K^+$ ions.[176,177] When folded, DFHBI docks on the G-quadruplex motif via nine hydrogen bonds that stabilize the *cis*-conformation, facilitating fluorescence.[177] Folded RNA, coordinated cation, and docked DFHBI are all essential to form the tripartite complex necessary for fluorescence.

Interactions between $Pb^{2+}$ and stacked G-quadruplex motifs are well documented.[199,201,276] One $Pb^{2+}$ atom docks between two successive G-quadruplex planes stabilizing the structure.[199,201,276] Physiologically, G-quadruplex are observed in coordination with $K^+$, however several other cations stabilize the structures.[199,201,252] $Pb^{2+}$ also interacts with non G-quadruplex motifs, facilitating catalysis of synthetic nucleic

acids.[24,277] Specific and selective interactions between nucleic acids and metal cofactors is critical for DNA and RNA structure and catalysis in natural and synthetic structures. Both are of interest in understanding the evolution of the central dogma of molecular biology and the role of nucleic acids therein.[278,279]

Spinach was proposed as an alternative tool for fluorescent detection of soluble $Pb^{2+}$.[202] The convenience of nucleic acid sensors is attractive, yet the chemical instability of RNA—especially for environmental sample analyses—complicates its utility. RNA has an additional complication of $Pb^{2+}$-catalyzed degradation.[280] Difficulties of using RNA sequences as experimental tools *in vitro* has led to a widespread use of modified nucleotides.[281,282] We hypothesized that modified nucleotides may provide similar avenues for chemical improvement to Spinach for use in *in vitro* applications. Of the various available modifications, ribose C2 fluorination (2'F) is among the most convenient and economical.

Ribose with a fluorine substitution favors a *3'endo* pucker, allowing for structural rigidity comparable to 2'OH RNA.[16] Yet, a 2'F modification abrogates nucleophilic attack on the phosphodiester backbone at modified positions, increasing chemical stability compared to 2'OH RNA.[16] A mutant T7 RNA polymerase was used to transcribe "2'OH WT" or "2'Fpyr" Spinach, the latter transcribed with equimolar ratios of 2'OH purine and 2'F pyrimidine triphosphates (see Supplement).[98] 2'Fpyr contains a 2'F ribose on 34 residues (40%) of the sequence (Figure 2-1A/B).

To examine whether 2'Fpyr retains functional properties similar to 2'OH WT, fluorescent and circular dichroism (CD) spectroscopies were employed. Despite having 34

fewer potential donor hydrogen bonds, 2'Fpyr docks DFHBI and facilitates its fluorescence, albeit with loss of ~35% (Figure 2-1C), and takes an overall structure similar to 2'OH WT (Figure 2-1D). Though reduced sharpness and intensity at 210, 224, and 264nm in the spectra suggests diminished G-quadruplex compaction.[202] Additionally, 2'Fpyr Spinach displayed similar cation selectivity as 2'OH WT (Figure 2-S1). Together, the data provide functional validation of 2'Fpyr, showing it to be comparable to 2'OH WT despite the chemical changes throughout the structure.

We next investigated the relative changes in chemical and structural stability between the constructs. The ubiquity and stability of ribonucleases complicate the use of RNA as a probe in physiological or environmental samples. We therefore examined the comparative stabilities of the constructs in the presence of purified RNase A, wherein 2'OH WT is more than 99% degraded in 5 minutes, while 2'Fpyr is less than 50% degraded in 5 hours (Figure 2-2A). $Pb^{2+}$ catalyzed RNA hydrolysis is well established, though the mechanism is unclear.[263,280] To test relative resistances to $Pb^{2+}$-catalyzed degradation, we exposed the two constructs to 500μM $Pb^{2+}$ at 37°C (Figure 2-2B). 2'Fpyr displayed increased resistance to $Pb^{2+}$, degrading ~3 fold slower than 2'OH WT. The relative structural stability of the G-quadruplex was measured by decrease in fluorescence signal with temperature, which is proportional to reduced tripartite complex stability, as loss of RNA structure or binding of either ligand results in complete loss of fluorescence. 2'F chemistry reduces G-quadruplex melting by ~7°C (Figure 2-2C). While previous investigations suggest that electron withdrawing 2'F ribose modifications increase the stability of Watson-Crick bound RNA duplexes, the specific constraints of 2'F on G-

quadruplex structures have yet to be explicitly explored.[119,120] Our data show that while 2'F enhances chemical stability of Spinach, the thermal stability of this G-quadruplex orientation is reduced.

Preliminary fluorescence studies revealed a consistent increase in signal after $Pb^{2+}$ addition, even tens of minutes after initial scans. To explore this, we compared binding kinetics of the tripartite components required for fluorescent readout. First, the kinetics of DFHBI association were investigated by mixing 10X concentrated DFHBI with a solution of Spinach RNA folded with $Pb^{2+}$, with fluorescence immediately recorded (Figure 2-S2A). The resulting kinetic fits are similar, with DFHBI saturating 2'OH WT and 2'Fpyr at similar rates, which—for 2'OH WT—agrees with previously published data.[161] Reciprocally, constructs were folded with DFHBI in an "Apo" state without $Pb^{2+}$, then mixed with 10X concentrated $Pb^{2+}$ and monitored similarly. 2'OH WT required more than 200 minutes to reach fluorescent saturation. In contrast, 2'Fpyr saturated within 40 minutes. $k_{obs}$ values were acquired from monoexponential fits for a range of $Pb^{2+}$ concentrations (Figure 2-3A). Third, RNA was folded without either DFHBI or $Pb^{2+}$, both were then added and fluorescence recorded. This resulted in a slow rate of fluorescent gain similar to the $Pb^{2+}$ addition (data not shown), suggesting cation binding is rate limiting for fluorescence. Linear regression of $k_{obs}$ values from $Pb^{2+}$ concentrations plotted against corresponding sums of $Pb^{2+}$ and RNA concentrations provides $k_{on}$ and $k_{off}$ (slope and y-intercept, respectively) (Fig. 3C).[161] We obtained kinetic values for 2'OH WT and 2'Fpyr, $k_{on}=(4.77\times10^{2}\ M^{-1}\ s^{-1})$, $k_{off}=(5.17\times10^{-3}\ s^{-1})$, and $k_{on}=(3.02\times10^{3}\ M^{-1}\ s^{-1})$, $k_{off}=(5.93\times10^{-2}\ s^{-1})$,

respectively. These yield a calculated $K_D$ (termed $K_D^{Calc}$) equaling 10.8□M, and 19.6□M, respectively, consistent with previous work.[202]

Prior studies presenting 2'OH WT Spinach fluorescence have consistently reported experimental time-scales shorter than 15 minutes.[202] Our data suggests such an approach would produce artificially suppressed fluorescence readouts. To estimate the magnitude of this suppression we determined a series of theoretical $K_D$ (termed $K_D^{Thr}$) values for various time points within the kinetic fits (Figure 2-S2B) (see Methods). The various $K_D^{Thr}$ values were then compared to the $K_D^{Calc}$, with the difference between the two presented as a percent error of $K_D^{Calc}$. In 10 minutes, $K_D^{Thr}$ for 2'Fpyr was found to be within 2% of $K_D^{Calc}$. In the same time, 2'OH WT $K_D^{Thr}$ deviated from $K_D^{Calc}$ by 196%. 2'OH WT did not produce a $K_D^{Thr}$ within 10% error of its $K_D^{Calc}$ for 180 minutes. Accelerated association kinetics is advantageous for the detection of $Pb^{2+}$ in samples with unknown RNase capacity. More importantly, however, is the increased internal consistency of 2'Fpyr fluorescence over a shorter time scale. Despite the potential limits of RNA for $Pb^{2+}$ detection, fluorination produces interesting kinetic results with potential utility as of yet unexplored.

To understand reasons for the differences in kinetics of $Pb^{2+}$ binding between the two constructs, we employed molecular dynamic (MD) simulations. The structure of 2'OH WT Spinach (PDB 4kzd) was used to generate a model of 2'Fpyr (see Supplement).[177] The two constructs were subjected to a 10 nanosecond simulation using AMBER ff14 RNA force field in both Apo and $Pb^{2+}$-coordinated states.[283] When coordinated with $Pb^{2+}$ both 2'OH WT and 2'Fpyr showed similar conformational rigidity over the 10ns simulation

(Figure 2-4). However, in the Apo state 2'Fpyr showed drastic local reduction in G-quadruplex compaction compared to its $Pb^{2+}$ bound state and the Apo state of the 2'OH WT (Figure 2-4B/C and Figure 2-S3). MD results agree with CD analysis (Figure 2-1D), thermal stability (Figure 2-2C), and the accelerated $Pb^{2+}$ binding kinetics (Figure 2-3B/C). Molecular breathing of the G-quadruplex residues could facilitate faster association by reduced steric occlusion and alterations in localized electrostatics.

The 2'OH groups of RNA contribute to both entropic and enthalpic aspects of RNA stability. Electrostatic repulsion between 2' and 3' oxygens shifts the ribose pucker equilibrium in favor of *3'endo* orientation, providing rigidity to intricate RNA architectures DNA cannot favorably maintain. The 2'OH is additionally capable of donating one and accepting two H-bonds, interacting with nucleobases and backbone phosphates in long range interactions promoting favorable non-Watson-Crick interactions not observed in DNA. Electronegativity of the hydroxyl contributes to the highly negative charge of RNA, enriching and complicating electrostatic interactions involved in folding and ligand binding. 2'F ribose similarly favors *3'endo* pucker conformation enabling comparable entropic rigidity. Additionally, organic fluorine are weak hydrogen bond acceptors and may replace a 2'hyroxyl in limited capacity.[284–286] Yet 2'fluorine moieties cannot donate a hydrogen bond, and the increased electronegativity of fluorine may alter localized electrostatics in ways incompatible with certain secondary or tertiary interactions. Given these considerations, it is surprising a direct replacement of 2' chemistry in Spinach retains secondary structure and fluorescence without sequence optimization.

The Spinach tripartite complex presents an interesting paradigm, bringing

together concepts of folding, equilibrium dynamics, affinity, and cooperativity. Our experimental data—CD, thermal stability, kinetic association—and MD simulations suggest cation association drives G-quadruplex stability and is the rate limiting step of fluorescent complex formation. MD simulations of 2'OH WT suggest G-quadruplex instability is independent of cation presence, which is unexpected given our observations that Spinach does not fluoresce without a coordinating cation. This raises interesting questions about possible long-range interactions between a coordinated cation and DFHBI through the G-quadruplex plane, independent of G-quadruplex structural stability. While revealing, our data does not completely elucidate the relationship between cation coordination and DFHBI fluorescence. Additional investigations of fluorescence-independent biophysical and computational approaches are merited.

In this work 2'F is limited to pyrimidines, and not present on the guanine nucleotides in the G-quadruplex, and the retention of ~60% of 2'OH chemistry in modified 2'Fpyr leaves the construct vulnerable to spontaneous alkaline or $Pb^{2+}$ dependent hydrolysis at those positions. The dramatic increase in chemical stability upon just pyrimidine fluorination suggests complete modification of the construct with alternative 2' chemistry at every position may further increase chemical stability of the structure. However, given differences in hydrogen bond capacity and electrostatic character of a completely modified RNA structure, it seems unlikely that total modification would be possible without sequence optimization.

Generally, our data demonstrate residues proximal to G-quadruplex have significant influence on the properties of the final motif. Hence, it may be possible to

rationally tune G-quadruplex properties and ligand binding by altering proximal residues with modified chemistries, or in combination with sequence alterations. This will be increasingly feasible with characterization of novel G-quadruplex conformations in future SELEX experiments. More broadly, the exploitation of modified nucleotides opens up new avenues for the investigations of G-quadruplex biology, and further broadens the expansive toolbox of nucleic acid biosensors.[284]

To our knowledge, this work represents the first investigation of 2'F nucleotides in a synthetic fluorescent RNA structure. Since the initial publication of Spinach, SELEX experiments have produced additional fluorescent RNAs: optimized Spinach derivatives such Spinach2, iSpinach, and Broccoli, and independent structures Mango, and Corn.[163–167,287] Interestingly, all of these have converged on derivatives of this stacked G-quadruplex motif, highlighting its importance as an ideal orientation for fluorophore docking. Our studies suggest that G-quadruplex folding, fluorophore docking, altered kinetics, and affinity for cation association with modified nucleotides is a promising approach for functional modification of alternative RNA aptamers with stacked G-quadruplexes. Optimized Spinach derivatives share conserved sequences in the G-quadruplex core. Despite this, the constructs have varying ligand affinities, thermal stabilities, and DFHBI fluorescent maxima. The comparative properties of these related structures raise interesting questions about the long-range influence of Watson-Crick duplexes, G-quadruplex proximal residues, and altered electrostatics on the tripartite complex formation and fluorescence. This growing class of synthetic fluorescent aptamers remains underutilized as *in vitro* tools, and although it remains to be seen

whether fluorination of these perturbs their folding, fluorescence, or dimerization, our findings with Spinach necessitate exploration. Increasing the chemical stability of these RNAs, while retaining their fluorescence for use *in vitro* through modified nucleotide chemistry, is a valuable expansion of their experimental utility.

While 2'Fpyr displays reduced thermal stability compared to 2'OH WT (Figure 2-2C), fluorescent-based kinetic studies revealed $Pb^{2+}$ associates with 2'Fpyr seven fold faster than 2'OH WT (Figure 2-3). The increased chemical stability allows for sensor utility over longer time-scales despite the observation that faster association kinetics enables $Pb^{2+}$ visualization on a shorter time frame, before RNA hydrolysis can become an issue. Accelerated kinetics of 2'Fpyr result in internally consistent fluorescent signal within 10 minutes, whereas 2'OH WT requires greater than 90 minutes before the kinetics are within 50% error of reliable saturation conditions (Figure 2-S2B). MD simulations suggest 2'F reduces compaction of the G-quadruplex in the absence of $Pb^{2+}$ (Figure 2-4C and Figure 2-S3). Increased movement of 2'Fpyr G-quadruplex Apo state explain experimentally observed changes in CD and fluorescence spectra, reduction in thermal stability, and accelerated binding kinetics with $Pb^{2+}$.

Accurate and rapid detection of heavy metal toxins is a public health necessity, with soluble lead of particular importance.[288–292] Classical approaches for lead quantification—Atomic Absorbance Spectrophotometry (AAS) and Inductively Coupled Plasma Mass Spectrometry (ICP-MS)—are associated with prohibitive costs.[293,294] Fluorescence is a comparatively simpler approach.[295,296] We present a chemical improvement of Spinach for the detection of soluble $Pb^{2+}$. 2'Fpyr displays a dramatic

increase in chemical stability in the presence of ribonuclease and $Pb^{2+}$, both of which pose a significant liability for RNA utility. We propose pyrimidine fluorination is a chemical improvement of Spinach for *in vitro* $Pb^{2+}$ detection, with 2'Fpyr Spinach displaying increased chemical stability and internally consistent fluorescent signal over shorter time scales.

## 2.3 Materials & Methods

***In vitro* Transcription**: RNA transcripts were prepared using linear dsDNA PCR templates, using the forward strand sequence (T7 promoter is underlined): 5'—GCGCGCGAATTCTAATACGACTCAC

TATAGGAGGACGCGACCGAAATGGTGAAGGACGGGTCCAGTGCGAAACACGCACTGTTGAGT

AGAGTGTGAGCTCCGTAACTGGTCGCGTC—3'. Reactions assembled at room temperature, in the following order, to final concentrations of: water, 1X T7 RNAP transcription buffer (NEB), 2.5-5% DMSO, 24mM $MgCl_2$, 4mM each rNTP (2'OH purines and 2'OH pyrmidines from NEB, 2'F pyrimidines from Trilink Biotechnologies), 15-25ng/μl dsDNA, 5mM DTT, 1U/μl murine RNase inhibitor (NEB), 2.5mU/μl yeast inorganic pyrophosphatase (NEB), 5U/□l WT T7 RNA polymerase (NEB) or 1.25U/μl mutant T7 R&D polymerase (Lucigen, Inc). RNA was transcribed at 42°C for ≥4 hours, and observed on a 4% agarose TBE gel. Target transcript band was purified by gel excision and electro-elution in dialysis tubing, and dialyzed against ≥5,000X volume of 10mM Tris (pH 8.0) for ≥2 hours at room temp. RNA was cleaned with three successive mixes with *acidified* phenol:chloroform:isoamyl alcohol (Ambion), then precipitated in 0.3M sodium acetate (pH 5.2) with either ≥50% isopropanol or ≥75% ethanol at -20°C for ≥2 hours. RNA was pelleted by centrifugation

(16.1k *x g* for ≥30 minutes at 4°C), washed with ice-cold 70% ethanol and centrifuged again (16.1k *x g* for ≥10 minutes at 4°C). RNA pellets dried at 37°C and suspended in RNase free 20mM Tris (pH 8.0), quantified by UV-absorbance at 260nm or fluorescence (Qubit, Invitrogen), and stored at -80°C. Based on current reagent prices and our average recovery efficiency at a low scale of transcription, we calculate 2'Fpyr modified RNA to be ~2-4 fold more expensive per unit mass than 2'OH unmodified. Unless otherwise stated, all RNA experimentation was performed in a standard base buffer solution of 20mM Tris (pH 8.0), 5mM $MgCl_2$, 20□M DFHBI. Prior to experimental use of frozen samples, RNA was folded following a basic protocol: Dilution in buffer then successive thermocycler incubations at 85°C (20sec), 50°C (1min), addition of DFHBI, 37°C (1min), 25°C (≥30min).

***RT.PCR****:* RNA constructs were reverse transcribed and amplified using NEB OneTaq OneStep RT.PCR kit following manufacturer protocol. PCR cycling conditions as follows: 48°C (30min), 95°C (1min), 25cyles at 95°C (20sec) 63°C (30sec) 68°C (30sec), 68°C (10min), 4°C. RT.PCR amplicons were gel purified as described above and ethanol precipitated, before ligation into a TOPO.pca2.1 TA sequencing vector (Invitrogen) and used for transformation of DH5α *E. coli*. Sanger sequencing was performed commercially by Eurofins Scientific.

**Liquid Chromatography - Mass Spectrometry**: LC-MS experimentation and analysis performed commercially by Novatia, LCC, using a Dionex UltiMate HPLC and Thermo Finnigan LTQ mass spectrometer.

**Circular Dichroism Spectroscopy**: Measurements were performed on an Aviv 202 CD Spectrometer at 20°C. RNA samples were folded and prepared similarly to fluorescence

based experiments. 5μM RNA samples were scanned in 1mm quartz CD cuvette from 320 to 185nm by 0.25nm steps, with a 2 second integrated read per step. Data presented as average of three individual scans. Control scans of DFHBI in buffer subtracted from RNA spectra. Values plotted in GraphPad Prism 5.

**Chemical Stability Measurements**: *RNase-dependent:* 0.5μM RNA was folded as described in standard spinach buffer without lead. RNA was incubated with purified RNase A (Qiagen) at a final concentration of 1.25mU/μl at 25°C. 10μl samples were removed from pooled incubation at designated time points quenched with 10mM EDTA, and 2X DNA loading dye (Thermo Scientific). Samples were run on a 3% agarose TBE gel, stained with Sybr Gold nucleic acid dye. Band intensities were quantified by densitometry with Image Lab software and plotted in GraphPad Prism 5. *$Pb^{2+}$-dependent:* 0.5μM RNA was folded as described in standard Spinach buffer solutions, and incubated with or without 500μM $Pb^{2+}$ for 4 hours at 37°C. 10μl samples were removed from pooled incubation at designated time points, quenched with 10mM EDTA, and frozen. Samples were run on a 3% agarose TBE gel, stained with Sybr Gold nucleic acid dye. Band intensities were quantified by densitometry with Image Lab software, and plotted in Graph Pad 5.

**Thermal Stability Measurements**: Thermal melting was observed using an ABI StepOne Real Time PCR Thermocycler. 0.5μM RNA was folded as described in 20mM Tris (pH 8.0), 5mM $MgCl_2$, 50μM $PbCl_2$, 20μM DFHBI, 1% DMSO, and loaded in 20μl volumes in triplicate wells in an ABI brand 96 well QPCR plate. Fluorescence was measured in the Sybr Green Ex/Em channel from 4°C to 70°C at a 2% gradient. Control scans of DFHBI in buffer

subtracted from RNA spectra. Fluorescent values averaged and plotted in GraphPad Prism 5.

**Fluorescence Measurements**: All fluorescent measurements were performed on a Photon Technologies International QM-1 steady state fluorescent spectrophotometer. Samples were prepared at stated concentrations and excited at 468nm, at 20$^{o}$C held constant by a circulating water bath, in a 2mm quartz fluorescence cuvette. Kinetic experimentation was performed by rapid, in-cuvette mixing of folded Apo-state RNA in buffer and 10% vol. of 10X $PbCl_2$ in water. Emission at 501nm was read every 29.5 seconds with a 0.5 sec shutter exposure. $K_D^{Calc}$ values were determined using $k_{on}$ and $k_{off}$ determined under steady state fluorescence conditions, at 200min and 40min for 2'OH WT and 2'Fpyr, respectively. $K_D^{Thr}$ values were determined by excluding data beyond stated time points, and refitting to generate new theoretical $k_{obs}$ at those time points. Control scans of DFHBI in buffer subtracted from RNA spectra. Fluorescence data was plotted and fit in GraphPad Prism 5.

**Modelling and Molecular Dynamic Simulations**: Spinach aptamer (PDB 4kzd) without co-crystalized Fab was used as the template for modelling and MD simulations. 2'Fpyr construct was modelled by direct chemical replacement of the C2 hydroxyl group with fluorine on every pyrimidine ribose using YASARA *Structure* modelling software. A revised AMBER ff14 RNA force field was used for energy minimizations, using an explicit solvent TIP3P water model. Nucleotide bases were frozen and the backbone was subject to an energy minimization, after which the backbone was frozen and the bases were minimized. The structure was then subject to a final, full-structure minimization with no constraints.

Potassium was replaced with a lead atom, in both constructs, and full-structure energy minimizations were performed with no constraints. MD simulations were run for 10ns, with snapshots taken every 100ps, using the YASARA md_run macro with explicit solvent conditions at 298K with a physiological intracellular NaCl concentration. Hydrogen bonding atoms present in the bases of the starting aptamer structure were identified and the distances between them were catalogued for every snapshot. A distance over 2.5Å was defined as a broken hydrogen bond. YASARA hydrogen bond interpolation was then used to count the number of hydrogen bonds present within the aptamer over time, within the G-Quadruplex, duplex sections, and the transition between the duplex and the G-Quadruplex core. These values were obtained for every snapshot. Hydrogen bond values were plotted in Sigma Plot, and structural images presented with Chimera.

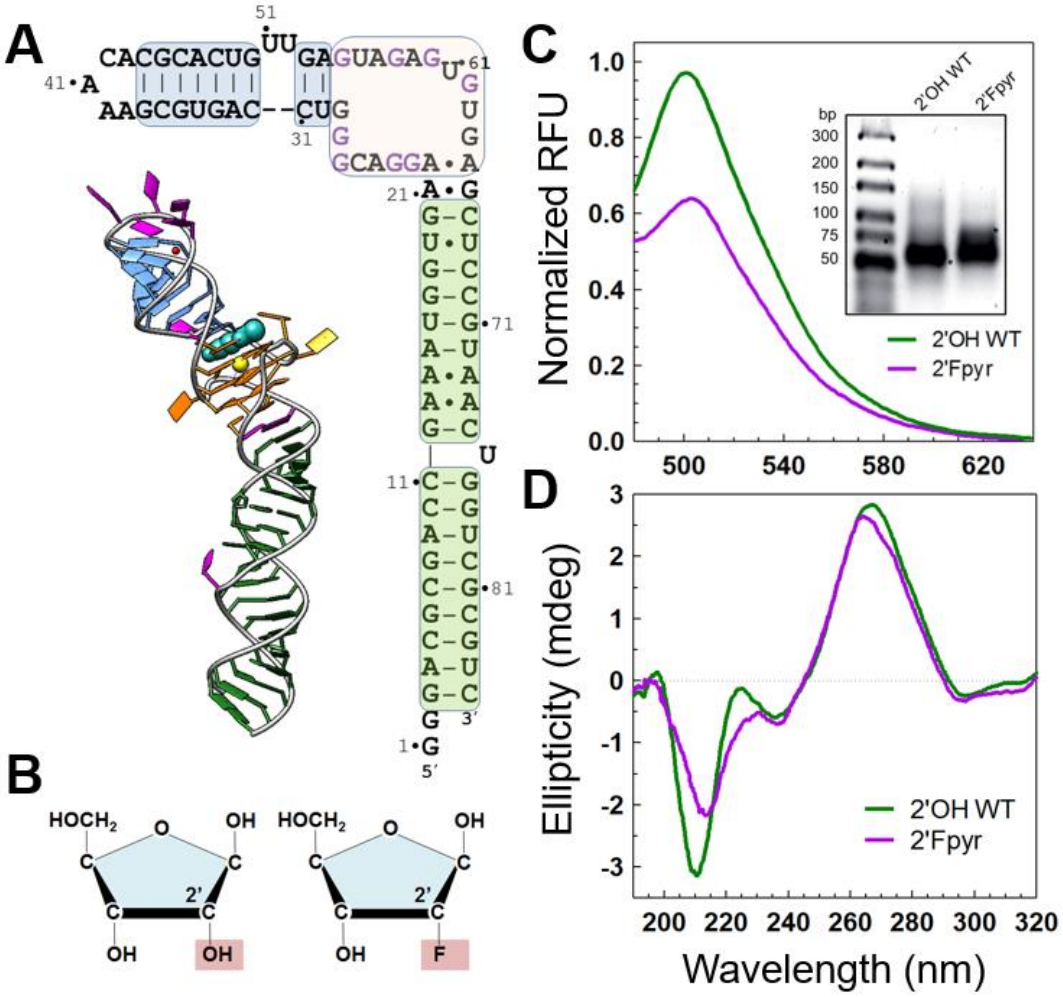



**Figure 2-1 Spinach Aptamer Accepts Fluorinated Pyrimidine Nucleotides.** (A) Spinach construct: cartoon of spinach. Flattened projection is color coordinated with 3D structure (adapted PDB 4kzd): green and blue are stem loops; orange is G-quadruplex; cyan molecule is DFHBI; yellow is Pb2+. (B) Ribose fluorination: modified ribose differ by a fluorine at C2. (C) Fluorescence: 2'OH WT and 2'Fpyr facilitate DFHBI Ex/Em at 468/501 nm. Gel inset depicts RNA integrity. (D) Circular dichroism: comparison of CD spectra showing structural similarity.

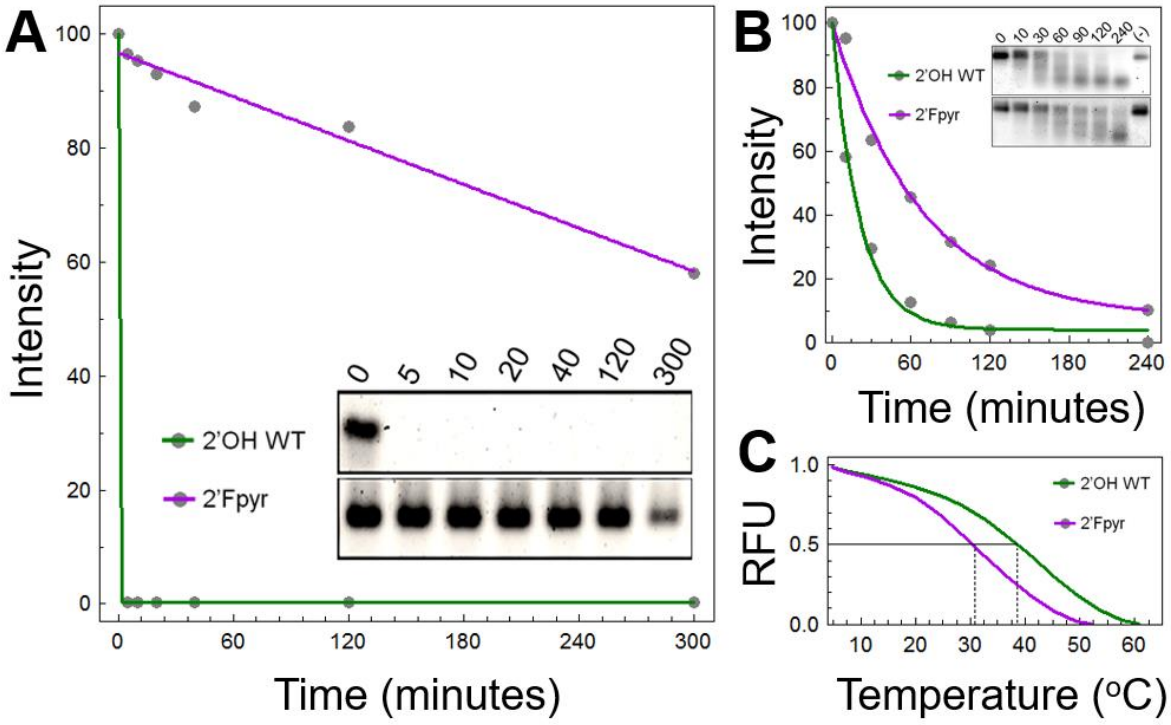



**Figure 2-2 Comparative Stability of Modified Spinach.** (A) RNase A-dependent stability: gel inset shows time-dependent aptamer degradation of single experiment (minutes). Graph plotted from densitometry quantification of three replications. (B) $Pb^{2+}$-Dependent stability: gel inset shows time-dependent aptamer degradation of single experiment (minutes). Negative control is 240 minute incubation without $Pb^{2+}$. Graph plotted from densitometry quantification of intact aptamer band in three replications. (C) Thermal stability: melt curves are readouts of tripartite complex stability

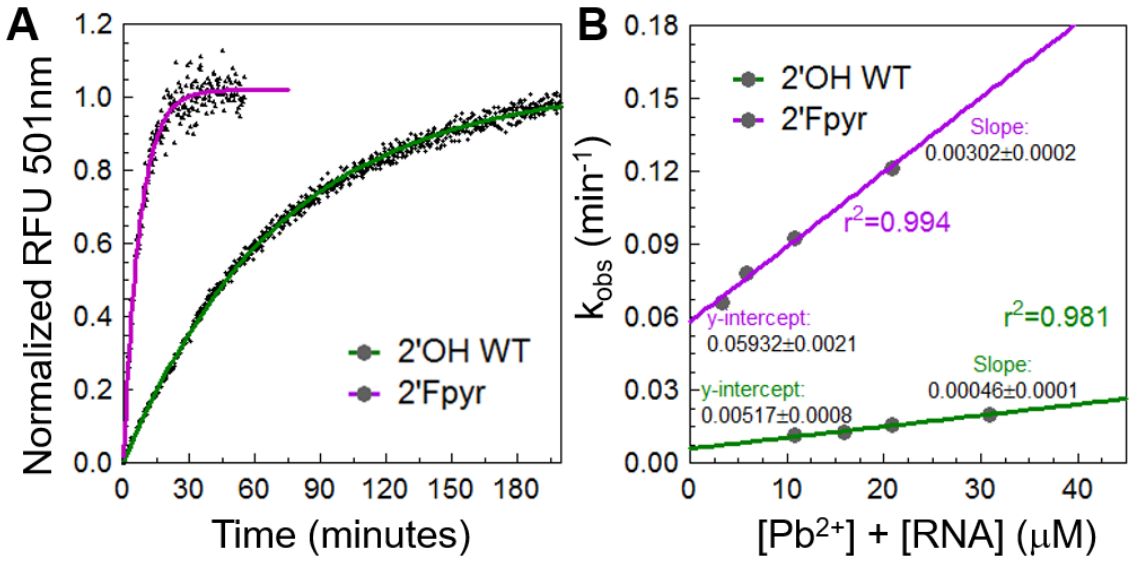



**Figure 2-3 Tripartite Complex Kinetics of Spinach with $Pb^{2+}$.** (A) Binding kinetics of $Pb^{2+}$: fitted fluorescence profile of 20mM $Pb^{2+}$ mixed with folded 0.5mM Apo RNA. (B) $K_{obs}$ vs. [$Pb^{2+}$ + RNA]: $k_{obs}$ values of multiple $Pb^{2+}$ concentrations plotted against the sum of $Pb^{2+}$ and RNA in each experiment. Linear regression provides $k_{on}$ values (slope) and $k_{off}$ (y-intercept), from which $K_D$ is calculated.

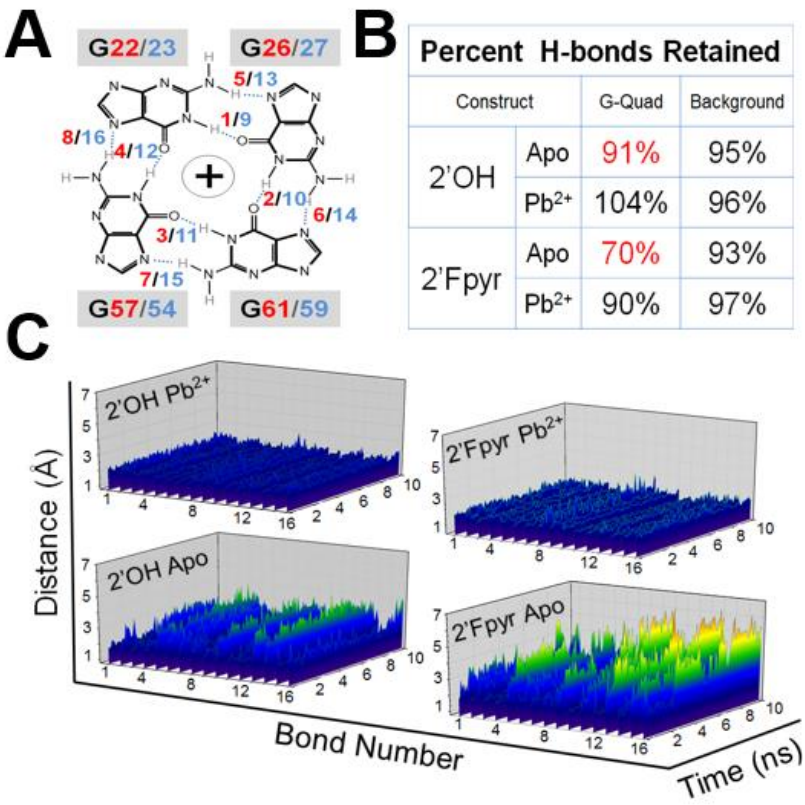



**Figure 2-4 G-Quadruplex Molecular Dynamics.** (A) G-quadruplex H-bonds: schematic of 16 guanine H-bonds of the two G-quadruplex planes. Red and blue correspond to bottom and top planes, respectively. (B) Percent retention: H-bonds retained after 10 ns simulation as percentage of starting count. H-bonds extended beyond 2.5Å in the simulation are considered broken. Background structure H-bonds calculated as the difference between total and G-quadruplex. (C) 3D visualization of distances: left plots are 2'OH WT, right are 2'Fpyr. Top depict RNA simulations in coordination with $Pb^{2+}$ and bottom are Apo.

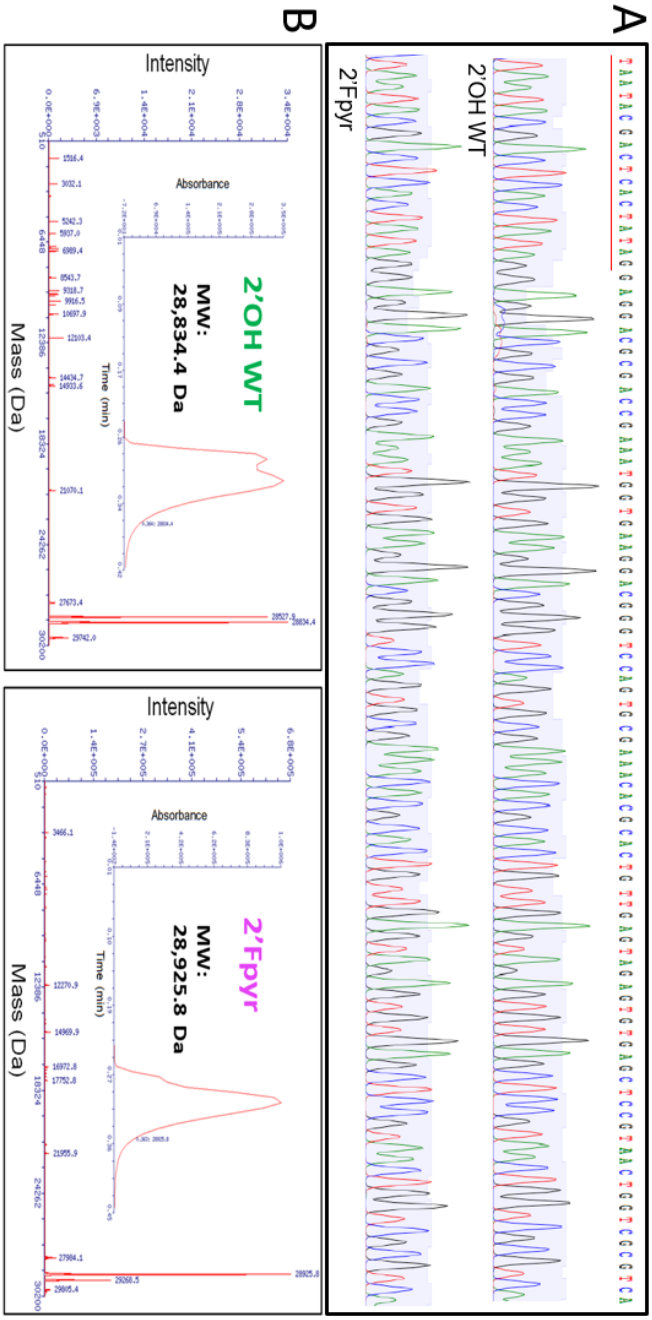



**Figure 2-S1 Mass Spectrometry and Sequence Analysis of Spinach.** (A) Sequence Identity: cDNA of Spinach constructs were sequenced by Sanger sequencing validating integrity of the transcripts generated by T7 RNA polymerase variants and modified nucleotides (T7 promoter sequenced is underlined in red). (B) LC-MS: Spinach Constructs analyzed by Liquid Chromatography—Mass Spectrometry to validate sample identity. 2'OH WT sample composed of two species differing by a single 5' terminal guanosine, likely due to T7 RNA polymerase slipping. Mass shift between 2'OH WT and 2'Fpyr constructs corresponds to fluorination of pyrimidine residues.

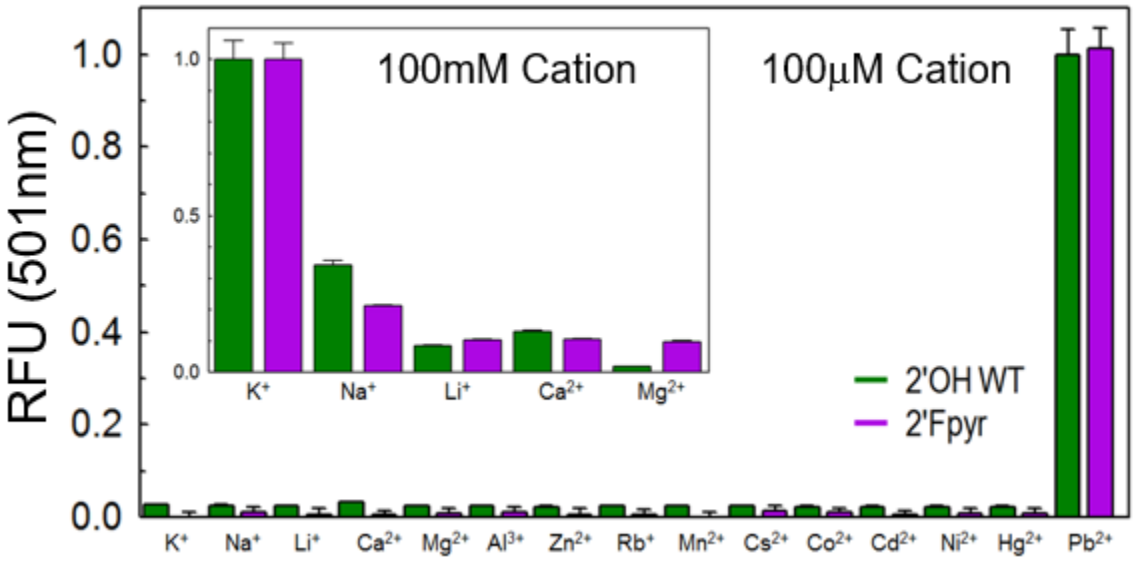



**Figure 2-S2 Comparative Cation Selectivity of Modified Spinach.** Spinach constructs were folded in the presence of DFHBI and various cations. Both constructs show similar fluorescent yield with each cation, showing the strongest preference for $Pb^{2+}$ in the micromolar range. Inset: physiologically relevant cations were also investigated in 100mM concentrations. Both constructs show similar preferences. All cation solutions prepared from $Cl^-$ salts for consistent background.

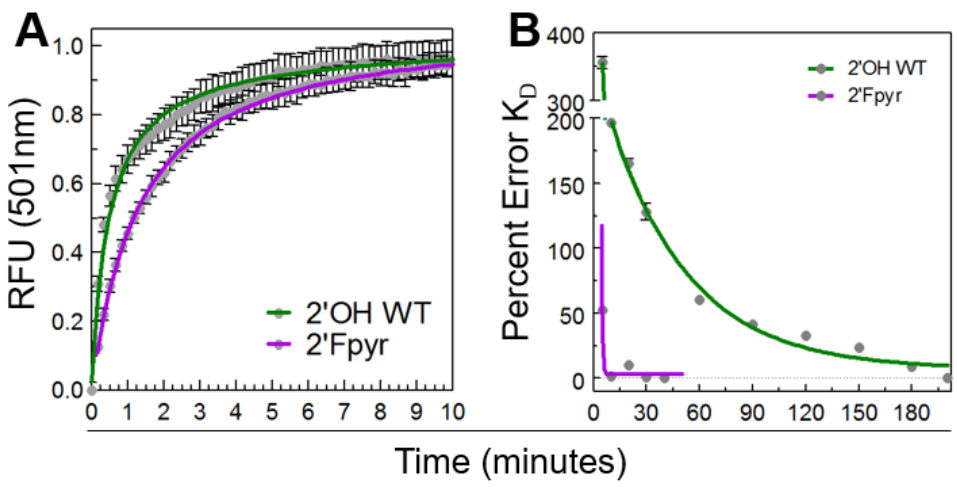



**Figure 2-S3 Supplemental Tripartite Complex Kinetics in Spinach.** (A) Binding Kinetics of DFHBI: Fluorescent gain after 20μM DFHBI mixing with folded 0.5μM RNA pre-incubated with Pb2+. Colored lines are monoexponential fits. (B) Percent error of fitted $K_D$ vs. Time. $K_D$ values determined ($K_D^{Thr}$) at specified time windows 5, 10, 20, 30, 40 (2'Fpyr and 2'OH WT), 60, 90 120, 150, 180, 200 min (2'OH WT only). Deviation from calculated $K_D$ ($K_D^{Calc}$; calculated at saturation for 2'Fpyr and 2'OH WT, 40 and 200min, respectively) at each time point is presented as percent error of $K_D^{Calc}$, and is calculated as ($|K_D^{Thr} - K_D^{Calc}| / K_D^{Calc}$)*100. Lines are monoexponential fits

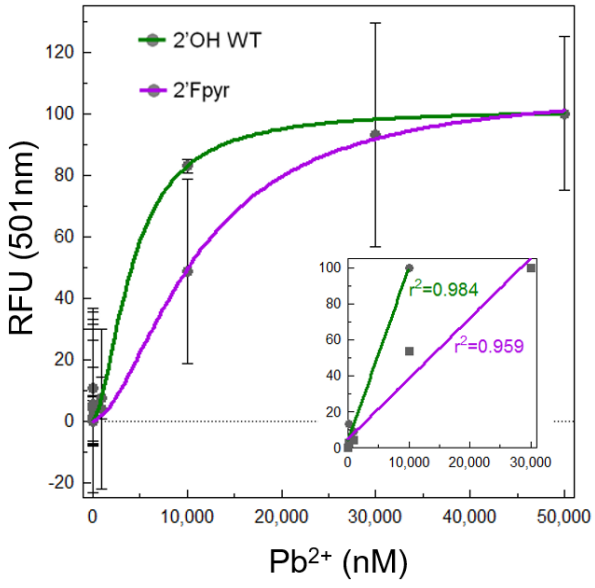



**Figure 2-S4 Detection Limit of $Pb^{2+}$.** $Pb^{2+}$ was detectable by fluorescence down to 10nM by both Spinach constructs. Inset: Linear regression of $Pb^{2+}$ detection by both constructs for $Pb^{2+}$ concentrations between 50nM and 30μM.

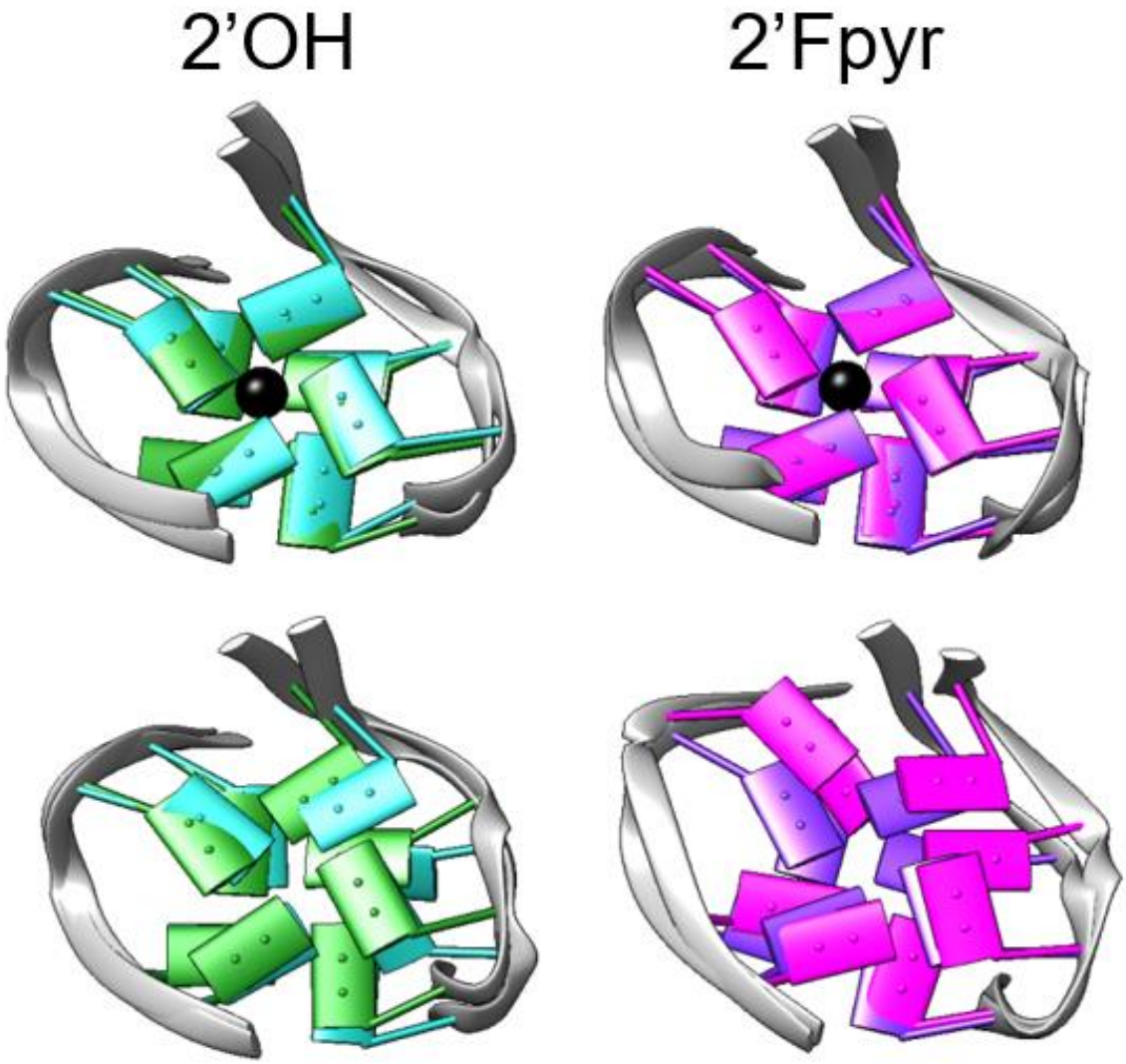



**Figure 2-S5 Visual Comparison of G-Quadruplexes in Molecular Dynamic Simulations.** Images correspond to 2'OH WT and 2'Fpyr both in coordination with $Pb^{2+}$ and in Apo states. Green and purple correspond to 0ns, and cyan and magenta correspond to 10ns.

# 3

# Subtle Sequence Variations Alter Tripartite Complex Kinetics and G-Quadruplex Dynamics in RNA Aptamer Broccoli

This chapter has been reformed from its original publication in *Chemical Communications*:

**Savage, J. C.**; Davare, M. A.; Shinde, U. Subtle Sequence Variations Alter Tripartite Complex Kinetics and G-Quadruplex Dynamics in RNA Aptamer Broccoli. *Chem. Commun.* 2020, *56* (17), 2634–2637. https://doi.org/10.1039/c9cc09375c.

J.C.S. and U.S. designed the research; J.C.S. performed all experimentation; M.A.D. provided thoughtful discussion and manuscript editing; J.C.S. wrote the paper; J.C.S. and U.S. edited and strategized the paper.

## 3.1 Abstract

Though extensively utilized, the fluorescent RNA aptamer Broccoli is poorly characterized with an unknown structure. Spectroscopic and kinetic investigations of tripartite complex formation reveal surprising differences between Broccoli and Spinach aptamers despite extreme sequence conservation. Our studies highlight how subtle sequence variations impart functional consequences of G-Quadruplex—cation interactions in RNA.

## 3.2 Discussion

The RNA aptamer "Spinach", selected *in vitro* by SELEX [4–6,160], docks to DFHBI *(3,5-difluoro-4-hydroxybenzylidene imidazolinone*) and induces DFHBI's fluorescence through a unique, cation-coordinated, stacked two-planed G-quadruplex motif, that coalesces into a tripartite complex of RNA + DFHBI + Potassium, or "RDK".[161,176,297] Spinach has been minimized in successive *in vitro* selections producing a family of smaller and generally more stable iterations including Spinach 2, Spinach mini, baby Spinach, and iSpinach.[163,164,170,176,287,297] "Broccoli", one of the smallest DHFBI-binding RNAs, has been utilized *in vitro* and *in vivo,* but not extensively characterized.[165,170] Despite the assumption that all Spinach RNAs share the same G-Quad motif, Broccoli has remained in question and its structure remains unknown.[298] Here we investigate the relationship between kinetics and structure of RDK formation in Broccoli and Spinach. The two aptamers display notable differences in kinetics of complex formation, yet they share extensive sequence identity and secondary structure, as observed by sequence

alignment, and Circular Dichroism (CD) and fluorescence spectroscopies. To understand our experimental observations we built a homology model of Broccoli, analyzed its structure, and subjected it to molecular dynamic (MD) simulations. Our data suggest that both aptamers adopt the same unique G-Quad motif, but with altered stability imparted by subtle sequence differences, resulting in differential RDK formation and altered complex stability.

Broccoli and Spinach RNAs display comparable fluorescence under high concentrations of KCl and DFHBI (100mM and 20□M, respectively), with both showing maximal excitation and emission at 468 and 501nm, respectively (Figure 3-1A). The fluorescent complex of Broccoli has an increased thermal stability over Spinach, as assayed by fluorescence (Figure 3-1B), which may be explained by the higher affinity for $K^+$ Broccoli displays over Spinach (Figure 3-1C). Interestingly, the kinetics of complex formation upon rapid mixing of concentrated KCl and DHFBI with RNA are notably different (Figure 3-1D). Fig.1E shows the linear regression of $K_{obs}$ values for RNA→RDK determined from the fit of five KCl concentrations with constant DFHBI (Figure 3-1D; *inset*). The slope of the regression provides $K^+$ association forming RDK, which is observed to be roughly 2.5-fold faster for Broccoli than Spinach (Figure 3-1E).

Researchers have treated Spinach and Broccoli as unrelated constructs, as per their nomenclature and 2D structural representations. However, the fact they induce fluorescence of the same ligand with similar photophysical properties and cation dependency necessitates deeper analyses of their sequence and structure. Sequence alignments show the G-Quad motif within iSpinach and Spinach share >92% sequence

identity (Fig.2A; S1). Only one residue differs in the G-Quad motif ($U_{17}$ and $A_{27}$ in iSpinach and Spinach, respectively). The superimposed X-ray structures of the two have a room mean square deviation (RMSD) of 0.63Å within the G-Quad motif (Figure 3-2B and Figure 3-S2 and Figure 3-S3). Broccoli also shares 91% and 95% sequence identity within the Spinach and iSpinach G-Quad motifs, respectively, suggesting Broccoli also adopts the same unique structure (Figure 3-2A and Figure 3-S2). To experimentally compare Broccoli and Spinach, CD spectroscopy was utilized.[201,202,232,233] The CD spectra reveal the two RNAs adopt nearly identical secondary structures, validating the similarity suggested by the sequence alignment (Figure 3-2A/C and Figure 3-S1). Furthermore, the ensemble of peaks at 295, 260, and 240nm provide experimental evidence of a G-Quad motif, independent of Spinach as a crystallized reference.[176,233,287,297] Together, the alignment and spectroscopic analysis strongly suggest Broccoli contains a G-Quad, in an orientation likely identical to Spinach.

Given the similarities based on sequence alignments and spectroscopic analyses, it is surprising that the kinetics observed in RDK complex formation between Broccoli and Spinach are different (Figure 3-1D/E). In pursuit of a mechanistic explanation of these observed differences a homology model of Broccoli was generated using iSpinach as the template (see Methods). iSpinach was chosen for four reasons: (1) Broccoli is more similar to iSpinach than Spinach in sequence length and identity (Figure 3-2A and Figure 3-S1); (2) a crystal structure of iSpinach exists, which (3) is nearly identical to Spinach, particularly at the G-Quad motif (Figure 3-S3); yet, (4) the reported thermal stability and

$K^+$ affinity of iSpinach are more similar to values obtained for Broccoli than for Spinach (Figure 3-1).[164]

Considering their sequence conservation, it is unsurprising the Broccoli model is superimposable upon the crystal structures of Spinach and iSpinach with RMSDs of their aligned G-Quad motifs <1.0Å (Figure 3-2B and Figure 3-S2 and Figure 3-S3). The G-Quad motif of Spinach and Broccoli is flanked by two A-form Watson-Crick (WC) duplex helices, one ending in both 5' and 3' termini and the other ending in a hairpin loop (Figure 3-3A). Structural analysis of Broccoli and Spinach reveal three base pair interactions located in the termini duplex that may, in part, explain the observed difference in thermal stability and cation affinity presented in Figure 3-1. The 3 differing residues around the G-Quad in Broccoli result in two canonical WC duplex base pairs with increased hydrogen bonding compared with Spinach ($A_4$,$C_7$,$G_{54}$ and $U_{18}$,$A_{21}$,$A_{64}$, in Broccoli and Spinach, respectively; Figure 3-3A). This suggests that, while substantially shorter than Spinach, the termini duplex of Broccoli is likely more stable. The effect of alternate base interactions may be the most dramatic at the first base-pairing downstream of the G-Quad motif in the termini duplex ($A_{21}$-$A_{64}$ in Spinach and $C_7$-$G_{54}$ in Broccoli). $C_7$-$G_{54}$ in Broccoli compresses the transition region between G-Quad and termini duplex in Broccoli, zipping up the duplex immediately adjacent to the G-Quad motif with a CG WC pair, increasing its stability with a minimum predicted enthalpic contribution of -5.8kcal/mol at that position.[299]

To generate hypotheses and rationalize observed differences in kinetics of $K^+$ association between Broccoli and Spinach, 50ns MD simulations with or without KCl were performed in YASARA modelling suite using a modified Amber14ff RNA force field (Figure

3-3B).[283] Broadly, the MD simulations show Broccoli to be more stable than Spinach (Figure 3-3B). Yet both Broccoli and Spinach display an increase in the conformational flexibility of their G-Quad motifs in an Apo state (without $K^+$ coordination), with Broccoli displaying more destabilization (Fig3-4A). Interestingly, what movement is seen is predominately just two H-bonds from one guanine residue, Broccoli $G_{49}$ and Spinach $G_{59}$. The dynamics of $G_{49}$ in Broccoli are drastically increased over the movement of equivalent $G_{59}$ in Spinach, with Broccoli $G_{49}$ flipping out of the G-Quad plane at ~5ns and remaining open for the duration of the Apo simulation (Figure 3-4A/B). $G_{49}$ flipping is particularly striking when the absolute distance of its movement is compared to the remaining G-Quad H-bonds of both structures (Figure 3-4B/C). In the Apo state, the total G-Quad motif of Spinach and Broccoli is rigid (Figure 3-4C; *top*). The 7 G-Quad guanines without $G_{49}$ of Broccoli display subtly increased conformational rigidity (Figure 3-4C; *middle*), and the majority of the G-Quad movement observed for Broccoli in Figure 4B may be attributed to $G_{49}$ (Figure 3-4C; *bottom*). The remaining residues, particularly $G_{17}$ and $G_{46}$, provide the site for DFHBI docking, and the subtly increased stability of these residues in Broccoli (Figure 3-4C; *middle*) may explain its increased Apo fluorescence (Figure 3-4D). Together it appears $G_{49}$ acts as a $K^+$ "gatekeeper", resulting in faster $K^+$ association and increased DFHBI fluorescence in the Apo state (Figure 3-1D and Figure 3-4D).

Despite the differences between Broccoli and Spinach derivatives suggested by their nomenclature and presented 2D structures, the two aptamers share much in common. They both promote the fluorescence of the same ligand with similar photophysics in a $K^+$-dependent manner, in the same concentration range. Sequence

analysis reveals the two share >90% sequence in the G-Quad motif suggesting structural similarity, which is experimentally corroborated by CD. Despite these similarities the two display differences in stability and kinetics of RDK formation.

Our model allows for investigations into the relationship between sequence, structure, and function in Broccoli and Spinach, and provides the means for *in silico* rationale of experimental observations (Figure 3-1). Two phenomena have been uncovered that may work synergistically. The first is canonical WC base pairing in the termini duplex stem of Broccoli that cinches the strands closer to the G-Quad motif through G-C hydrogen bonding, which may restrict the conformational flexibility of the motif, increasing the RDK melting temperature as observed (Figure 3-1B). The second is the movement of $K^+$ gatekeeper $G_{49}$ to an open state seen in the Apo MD simulations, which could explain the accelerated $K^+$ association kinetics (Figure 3-1D/E) by increased accessibility to the inside of the G-Quad motif. This hypothesis is consistent with recent work on lead association with 2'F pyrimidine modified Spinach RNA.[300] Subtle perturbations in local chemical environments—either due to modification of pyrimidine ribose or minor changes in the primary sequence—produce conserved G-Quad motifs with altered dynamics and kinetic stabilities that result in changes of cation association rates. These data posit further questions on the capacity for rationally tuning of cation selectivity into related RNA structures for modified function which is yet to be explored. Broadly, our work is further complemented by another recent publication[298], wherein a single pyrimidine residue in Broccoli tunes the fluorescent emission of fluorophore DFHO by up to 20nm. Together these demonstrate how subtle differences in chemistry and

sequence can produce changes in biophysical properties of RNA, independent of structure. These observations are likely not unique to this RNA family and may have broad implications for a wide array of G-Quad structures.

G-Quads are critical in numerous biochemical systems including gene regulation, metabolite sensing, and diseases ranging from cancer to viral infection to Alzheimer's, and are attractive for exploitation in synthetic biology and for the development of chemical tools.[210,216,301–307] Despite their importance in RNA biology, the majority of *in vitro* characterizations of G-Quads rely on telomeric DNA; G-Quads are inherently weak in their absorbance and fluorescence.[231,308,309] but telomere sequences are almost exclusively composed of G-Quad stacks, magnifying their signal strength with increasing sequence length. While this makes telomeres a spectroscopically convenient G-Quad model, telomeric DNA does not provide comparable insight into equivalent RNA sequence.[310] Spinach family RNAs are predominately used for subcellular visualization of larger complexes[170,311], yet their utility as models for G-Quad and RNA-cation investigations are heretofore unrealized. This RNA family shares a unique combination of model features: (i) structure-dependent fluorescent by a small molecule ligand with high affinity and specificity; (ii) two distinct cation cofactors of differing valences and hydration states ($K^+$ and $Mg^{2+}$); and (iii) three distinct structural motifs (Watson-Crick duplex, hairpin, and G-Quad). These features provide a system for RNA investigations from various functional and chemical perspectives within a constant background for direct experimental comparisons.

Our results provide mechanistic explanations for the increased thermal stability and accelerated rate of tripartite complex formation of Broccoli compared to Spinach. These investigations reveal how subtle sequence differences result in canonical WC base pairings, which stabilize the G-Quad motif and increase dynamics of gate-keeper $G_{49}$, promoting cation association. This work utilizes DFHBI fluorescence to investigate RDK tripartite complex formation; however, these investigations posit new questions about component interactions which require alternative, fluorescence-independent biophysical approaches for direct detection of cation hydration, order of association, cooperativity, or folding on sub-second timescales. Such investigations will provide proof-of-principle for biochemical investigations of RNA using this model system, a more complete picture of Broccoli-cation complex biophysics, and strengthen our understanding of the relationship between sequence, structure, and function in RNA.

## 3.3 Materials & Methods

**In vitro Transcription and purification**: RNA transcripts were prepared using linear dsDNA PCR templates. Sequences of the forward DNA strand are as follows (T7 promoter is underlined, construct in bold font): Spinach 5'—GCGCGCGAATTCTAATACGACTCACTATAGGAGGACGCGACCGAAATGGTGAAGGACGGGTCCAGTGCGAAACACGCACTGTTGAGTAGAGTGTGAGCTCCGTAACTGGTCGCGTC—3'; and Broccoli 5'—GCGCGCGAATTCTAATACGACTCACTATAGGGAGACGGTCGGGTCCAGGCACACAAAAATGTTGCCTGTTGAGTAGAGTGTGGGCTCC—3'. Prior to transcription, PCR templates were gel purified, phenol:chloroform cleaned, and desalted, similar to RNA clean up ( described

below). Transcription reactions were assembled at room temperature in the following order, to final concentrations of: water, 1X T7 RNAP transcription buffer (NEB), 24mM $MgCl_2$, 4mM each rNTP (NEB), 15-25ng/μl dsDNA, 5mM DTT, 1U/μl murine RNase inhibitor (NEB), 2.5mU/μl Yeast Inorganic Pyrophosphatase (NEB), 5U/μl T7 RNA polymerase (NEB). RNA was transcribed at 42°C for 4-10 hours. Transcriptions were quenched with DNase I treatment, and observed on a 4% agarose sodium borate gel, stained with Sybr Gold nucleic acid dye. Target transcript bands were purified by gel excision and electro-eluted into 3.5kDa MWCO dialysis tubing. RNA was concentrated with a 3.5KDa MWCO concentrating spin column (Millipore) and cleaned with acidified phenol:chloroform:isoamyl alcohol (Ambion). RNA was then desalted into 5mM Tris (pH 8.0) using a Zeba 7kDa MWCO desalting spin column (Thermo), and quantified by UV-absorbance at 260nm and fluorescence (Qubit HS RNA kit; Invitrogen), and stored at -80°C. Prior to experimental use of frozen samples, RNA was folded following a basic protocol: Dilution into desired buffer, then successive thermocycler incubations at 85°C (20sec), 50°C (10sec), 37°C (10sec), addition of DFHBI, 4°C (≥30min).

**Fluorescence:** All fluorescent measurements were performed on a Photon Technologies International QM-1 steady state fluorescent spectrophotometer. All samples were read under the following solution conditions unless otherwise stated: 0.5μM RNA, 20mM Tris (pH 8.0), 100mM KCl, 5mM MgCl2, 10μM DFHBI. Samples were 20°C held constant by a circulating water bath, excited at 468nm in a 2mm quartz fluorescence cuvette. Kinetic experimentation was performed by rapid, in-cuvette mixing of folded Apo-state RNA in buffer and a 10% volume of 1M KCl in water. Emission was then read at 501nm every

29.75 seconds with a 0.25 second shutter exposure. KCl titrations were performed by folding RNA in the presence of various KCl concentrations in individual reaction tubes. Samples were then incubated on ice for 2hr to ensure equilibrium was reached. Fluorescent of each sample was individually scanned (468/501nm). Fluorescence data was plotted and fit in GraphPad Prism 5.

**Circular Dichroism:** Measurements were collected on an Aviv 202 CD Spectrometer at 20°C. RNA samples were assayed in 10mM Cacodylate (pH 7.4), 100mM KCl, 5mM $MgCl_2$. RNA was folded as described above. 135ng/□l RNA samples were scanned in 1mm quartz CD cuvette from 320 to 190nm by a 0.2nm step, with a 3 second integrated read per step. Data presented as average of three individual scans. Control scans of buffer were subtracted from spectra. Values plotted in GraphPad Prism 5.

**Thermal Stability:** Thermal melting was observed using an ABI StepOne Real Time PCR Thermocycler. 0.5□M RNA was folded as described in 20mM Tris (pH 8.0), 5mM $MgCl_2$, 100mM KCl, 20□M DFHBI. 20□l sample volumes were scanned in triplicate wells in an ABI brand 96 well white opaque QPCR plate. Fluorescence was measured in the Sybr Green Ex/Em channel from 4°C to 70°C at a 2% gradient. Control scans of DFHBI in buffer were subtracted from RNA spectra. Fluorescent values averaged and plotted in GraphPad Prism 5.

**Homology modelling:** Initial homology models for Broccoli were produced using automated 3D RNA structural prediction servers, RNA Composer[312], ModeRNA[313], and SimRNA[314]. While these servers provided 3D predictions consistent with their structural templates, a complete and fully automated predictive model for Broccoli was only

partially generated. Hence, a local multiple sequence alignment (MSA) of the target RNA, Broccoli, with derivatives of Spinach, and multiple structural templates 4TS2, 4KZD, and 5OB3 was obtained using ClustalW, and edited to ensure optimal base complementarity indicated by arrows pointing towards the hairpin (Fig.2A). The G-Quad motif for Broccoli was generated from 5OB3 based on the MSA that established this region of the target RNA is almost identical within all available structural templates. An extensive fragment library search identified a loop of 16-nucleotides from a zinc finger RNA (1UN6, residues 13-28) as an optimal fragment to bridge the G-quad motif in Broccoli from residues 20-35[315]. The model was energy minimized using a modified Amber14ff RNA force field in YASARA modelling suite, and the quality of the model was assessed using RNAssess[316], a computational server that discriminates between potentially correct and incorrect conformations by comparing RNA 3D models with the reference structures.

**MD simulation:** Generated Broccoli model and Spinach (4KZD) without co-crystalized Fab were used in MD simulations. A revised AMBER ff14 RNA force field283 was used for energy minimizations, using an explicit solvent TIP3P water model. MD simulations were run for 50ns, with snapshots taken every 100ps, using the YASARA md_run macro with explicit solvent conditions at 298K with a physiological intracellular NaCl concentration. Hydrogen bonding atoms present in the bases of the starting aptamer structure were identified and the distances between them were catalogued for every snapshot. Hydrogen bond values were plotted in Sigma Plot, and structural images presented with Chimera.

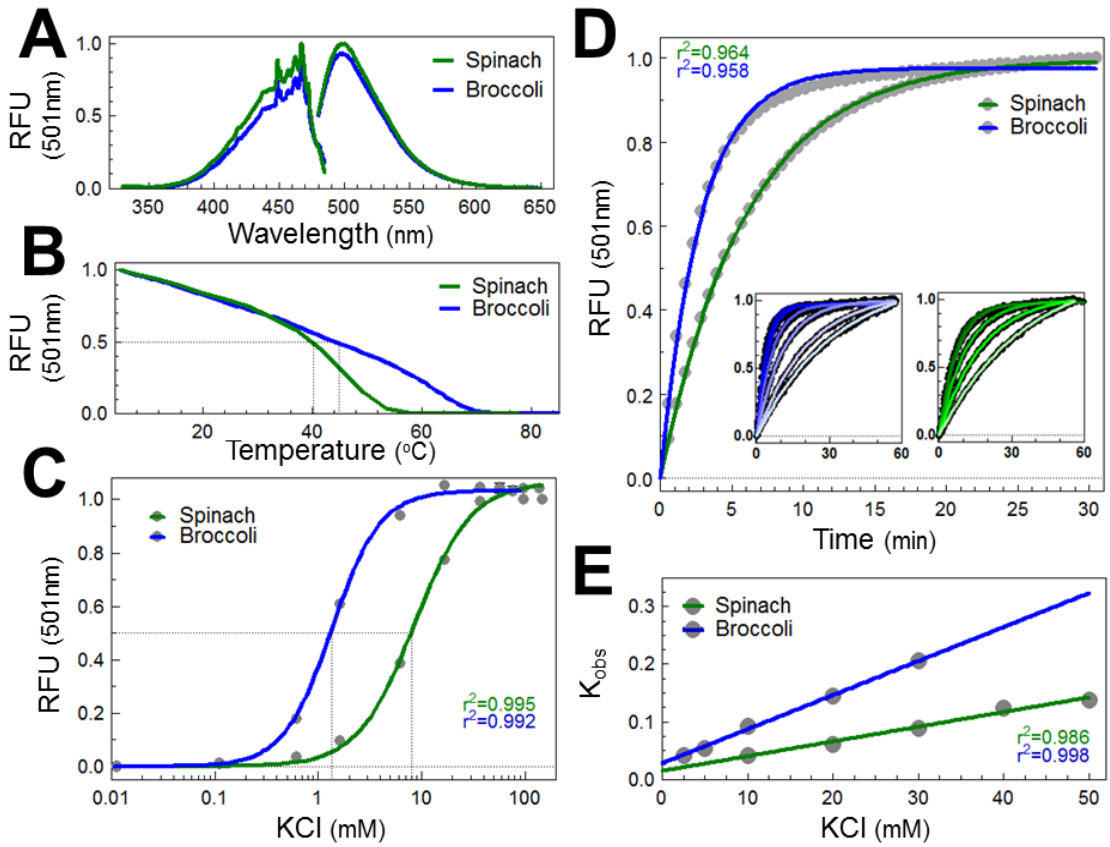



**Figure 3-1 Fluorescence, Stability, and Tripartite Complex Kinetics in Spinach and Broccoli.** (A) Fluorescence spectrum: fluorescence of DFHBI induced by Spinach and Broccoli show similar properties including Ex/Em wavelengths (468/501 nm). Constructs display 45% difference in intensity. (B) Thermal melting: RDK fluorescence as a function of temperature. Broccoli displays increased thermal stability (Tm = 45°C) compared to Spinach (Tm = 40°C). (C) K+ titration: RNA incubated with 20mM DFHBI and varying [KCl], and allowed to reach equilibrium. Fluorescent Em at 501nm was recorded, and values fit for approximate affinity for $K^+$. Broccoli shows a B10-fold increase in affinity for $K^+$. (D) Comparative kinetics: Apo folded Spinach or Broccoli was rapidly mixed with concentrated KCl and DFHBI in a cuvette and scanned for fluorescence upon RDK formation. Colored lines are fit. Inset: Identical approach performed with multiple concentrations of KCl, with fixed concentrations of RNA and DFHBI. All scans normalized and fit. (E) Comparative kinetics continued: $K_{obs}$ values for each KCl concentration from fits of (B) inset were plotted against respective KCl concentrations. Slope of the linear regression equals the $K_{on}$ for complex formation. Spinach and Broccoli $K_{on}$ = 0.0025 $M^{-1}$ $min^{-1}$ and 0.0059 $M^{-1}$ $min^{-1}$, respectively

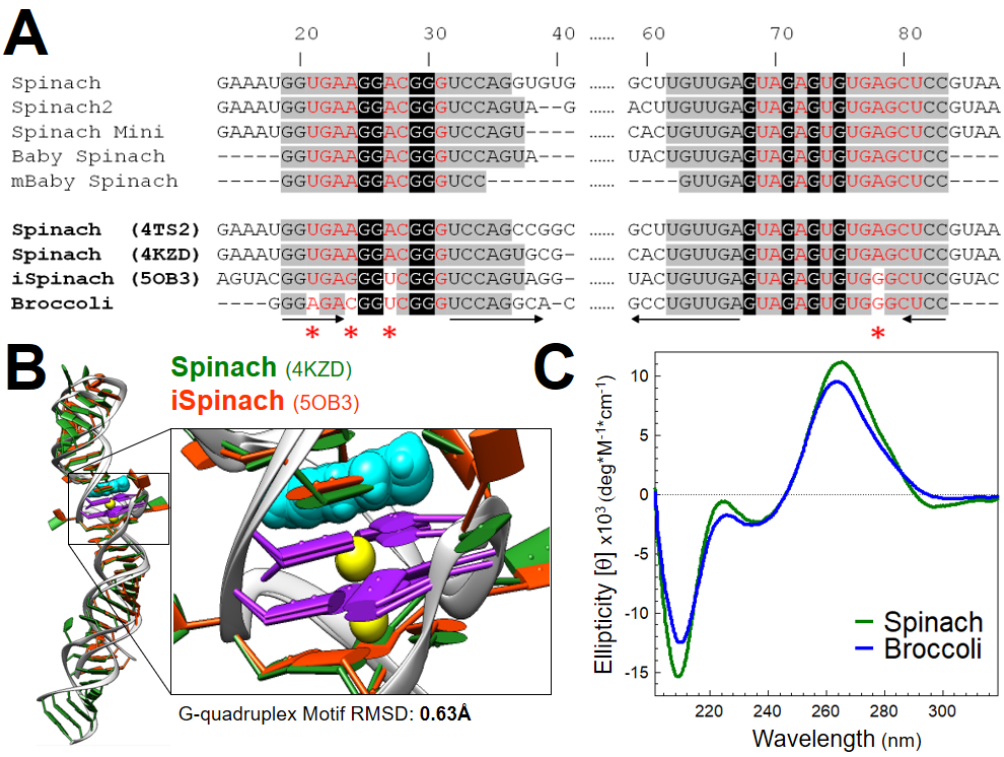



**Figure 3-2 Analysis of Spinach Family Aptamers.** (A) Sequence alignment: Comparison of Spinach derivative family members and Broccoli. Color coding: Red letters/grey background are G-Quad core motif residues; White letters/Black background are G-Quad guanine residues; Black Lettering/Grey background are conserved non-G-Quad residues; arrows indicated complementary regions that point towards the hairpin. Bottom and bold sequences correspond to crystalized sequences (or in the case of Broccoli, modelled here). (B) Spinach v iSpinach comparison: The crystal structures of Spinach (4KZD) and iSpinach (5OB3) are aligned. RMSD of the alignment at the G-Quad core motif show a considerable degree of structural similarity, suggesting iSpinach to be a satisfactory template for generating a homology model of Broccoli. G-Quad guanine bases are purple, DFHBIs are cyan, and $K^+$ ions are yellow. (C) Circular Dichroism: Spinach and Broccoli both show a clearly similar spectral signature by CD, suggesting Broccoli contains the G-Quad core motif shared by all other Spinach derivatives. Additionally, signal intensities at wavelengths 295, 260, and 240nm confirm the presence of a G-Quad independent of Spinach as reference structure.

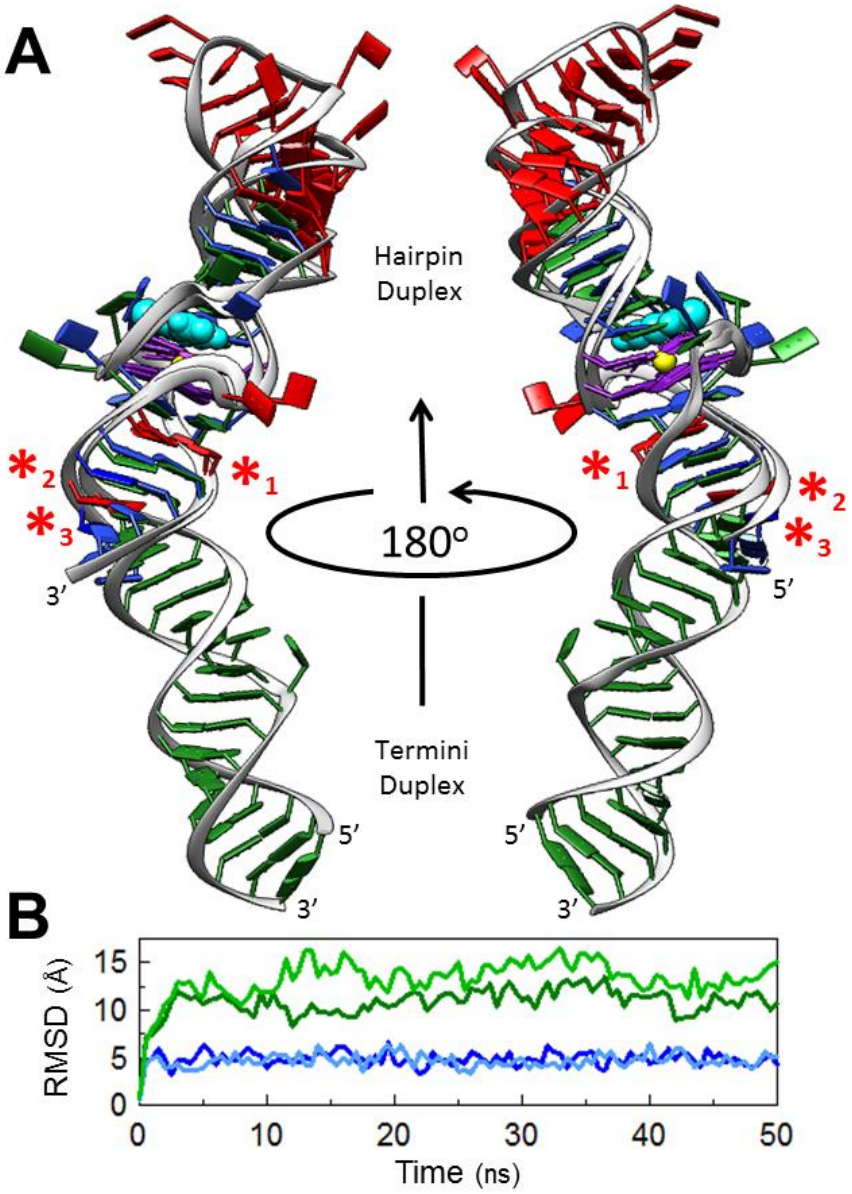



**Figure 3-3 Structural Comparison of Spinach and Broccoli.** (A) Broccoli and Spinach base-pairing comparison: Broccoli and Spinach Structures are aligned (model and PDB ID 4KZD, respectively). Color coordination is consistent with previous figures; Broccoli in blue, Spinach in green, G-Quad guanines in purple, DFHBI's in cyan, $K^+$ ions in yellow, Non-conserved residues in red. *1 is $C_7$-$G_{54}$ in Broccoli and $A_{21}$-$A_{64}$ in Spinach; *2 is $A_4$-$U_{57}$ in Broccoli; *3 is $U_{18}$-$U_{67}$ in Spinach. (C) Time course of MD simulation: Spinach structure (4KZD) and Broccoli model were subjected to a 50ns MD simulation. RMSD was used as a metric for structural stability over the time window, with Broccoli displaying increased structural stability. Green and blue lines represent Spinach, and Broccoli, respectively; light and dark hues are Apo and potassium bound, respectively.

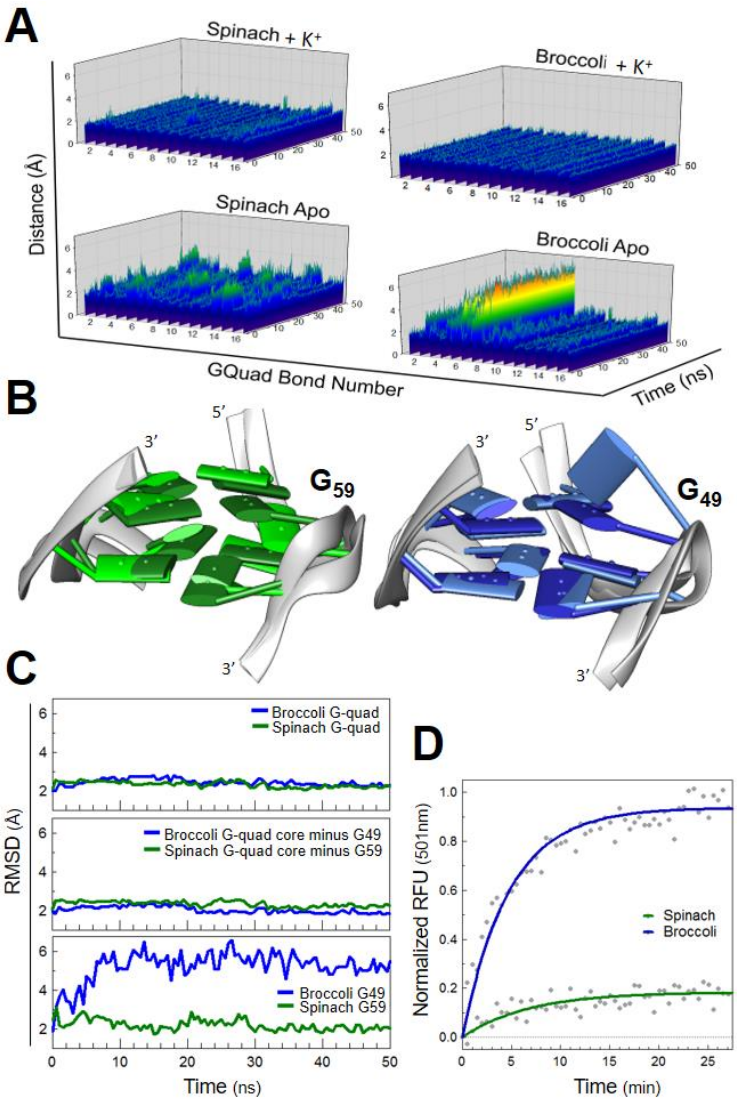



**Figure 3-4 Molecular Dynamic Simulations of Spinach and Broccoli.** (A) Structural visualization of MD simulations: The G-Quad guanine residues of Spinach (left, green) and Broccoli (right, blue) are presented. T=0 and T=50ns of each are aligned (dark shades are $T_0$, light shades are $T_{50}$). The most drastic movement for both structures corresponds to the same guanine ($G_{59}$ of Spinach and $G_{49}$ of Broccoli). The movement of $G_{49}$ in Broccoli is much more pronounced. (B) Residue specific H-Bond dynamics in the G-Quad: The 16 H-bonds from the 8 Guanine residues of the stacked G-Quad guanines are plotted as a function of stretch distance and time, in Spinach and Broccoli, both with and without $K^+$ ("Apo"). As can be seen in (A), Broccoli shows more base movement than Spinach in the Apo state, however the majority of the movement is restricted to one base, $G_{49}$. The $G_{49}$ trap door may explain the accelerated $K^+$ association kinetics observed in Figure 1 by providing an opening for $K^+$ association. (C) Cumulative RMSD of G-Quad guanines: Stability of the G-Quad guanines *without* potassium was looked at via RMSD over the same MD simulation. Both RNAs display rigid G-Quad structures throughout the simulation (*Top*). Broccoli appears to have a subtle increase in G-Quad stability when $G_{49}$ is removed, i.e. remaining 7 guanines (*middle*), whereas $G_{49}$ alone in Broccoli appears to account for most or all of the net G-Quad dynamics, in contrast to $G_{59}$ in Spinach which is quite static (*bottom*). (D) RNA induced DFHBI association without Potassium: DFHBI was spiked into Apo RNA and fluorescence gain was observed. Although low background fluorescence was observed for both, Broccoli displayed 5 fold higher signal than Spinach.

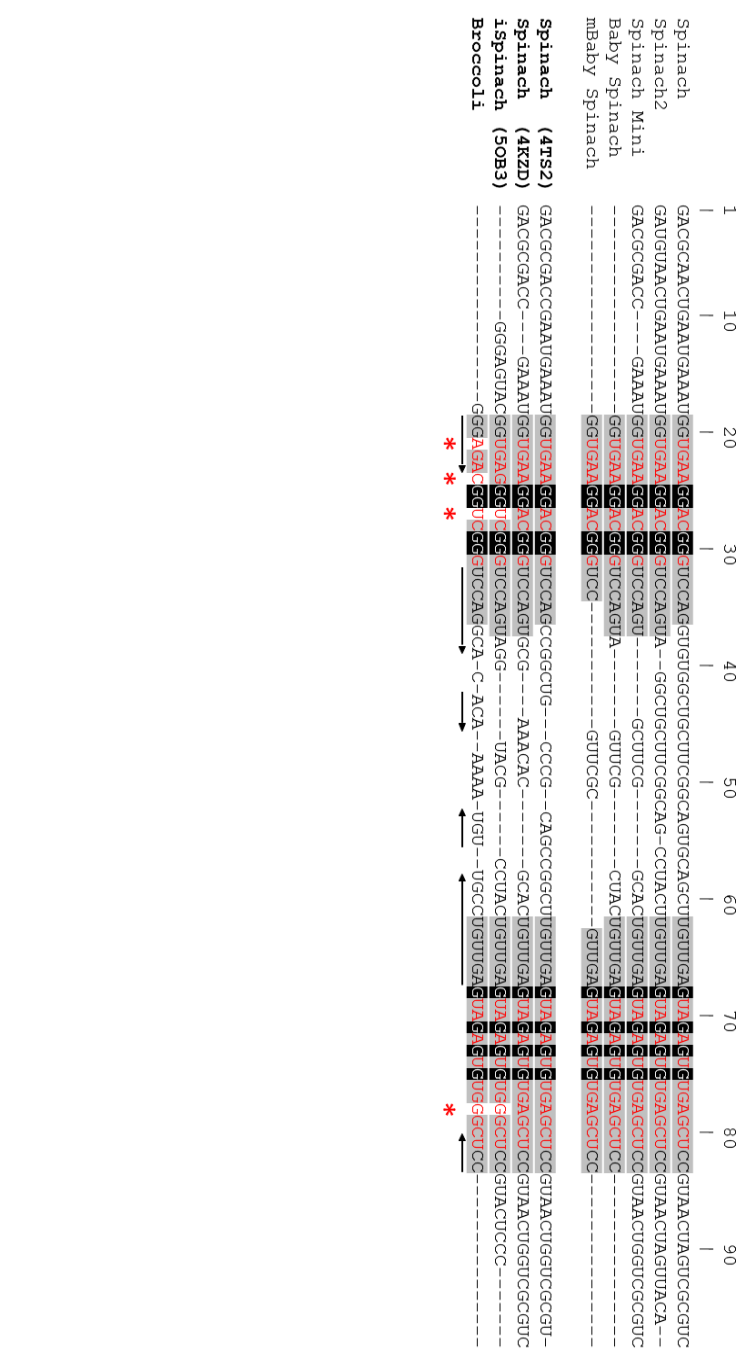


**Figure 3-S1 Full sequence alignment of Family Aptamers.** Comparison of Spinach derivative family members and Broccoli. Color coding: Red letters/grey background are G-Quad core motif residues; White letters/Black background are G-Quad guanine residues; Black Lettering/Grey background are conserved non-G-Quad residues; arrows indicated complementary regions that point towards the hairpin. Bottom and bold sequences correspond to crystalized sequences (or in the case of Broccoli, modelled here).

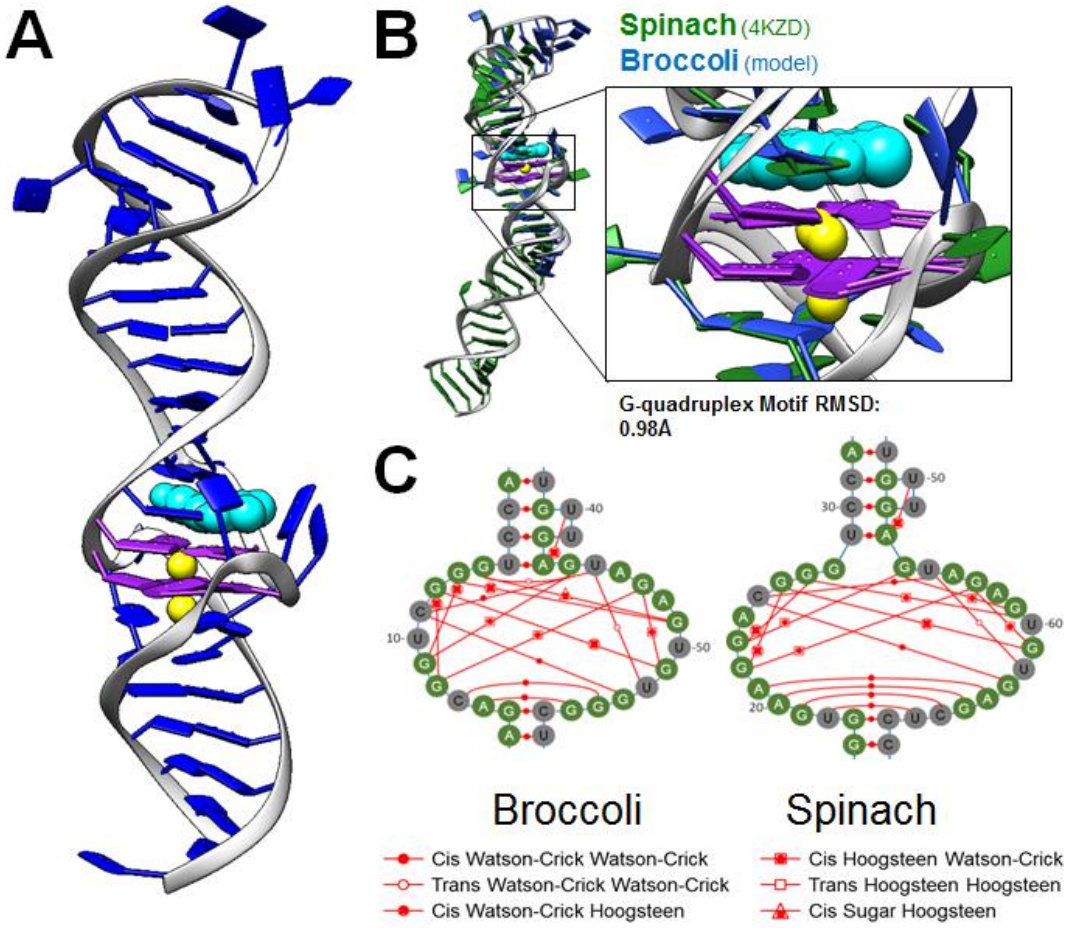



**Figure 3-S2 Supplemental Structural Comparison of Spinach and Broccoli.** (A) Broccoli model: Cartoon representation of the homology model of Broccoli; DFHBI, G-Quad guanine bases, and K+ ions are colored cyan, cyan, and yellow, respectively (B) Spinach v Broccoli comparison: The crystal Structure of Spinach (4KZD) and model Broccoli are aligned. The two show significant structural similarity, as would be expected based on the similarity between Spinach and iSpinach. G-Quad bases are purple, DFHBIs are cyan, and K+ ions are yellow. (C) Residue interaction map: The G-Quad core residues are presented in a 2D flattened projection, with corresponding nucleotide interactions.

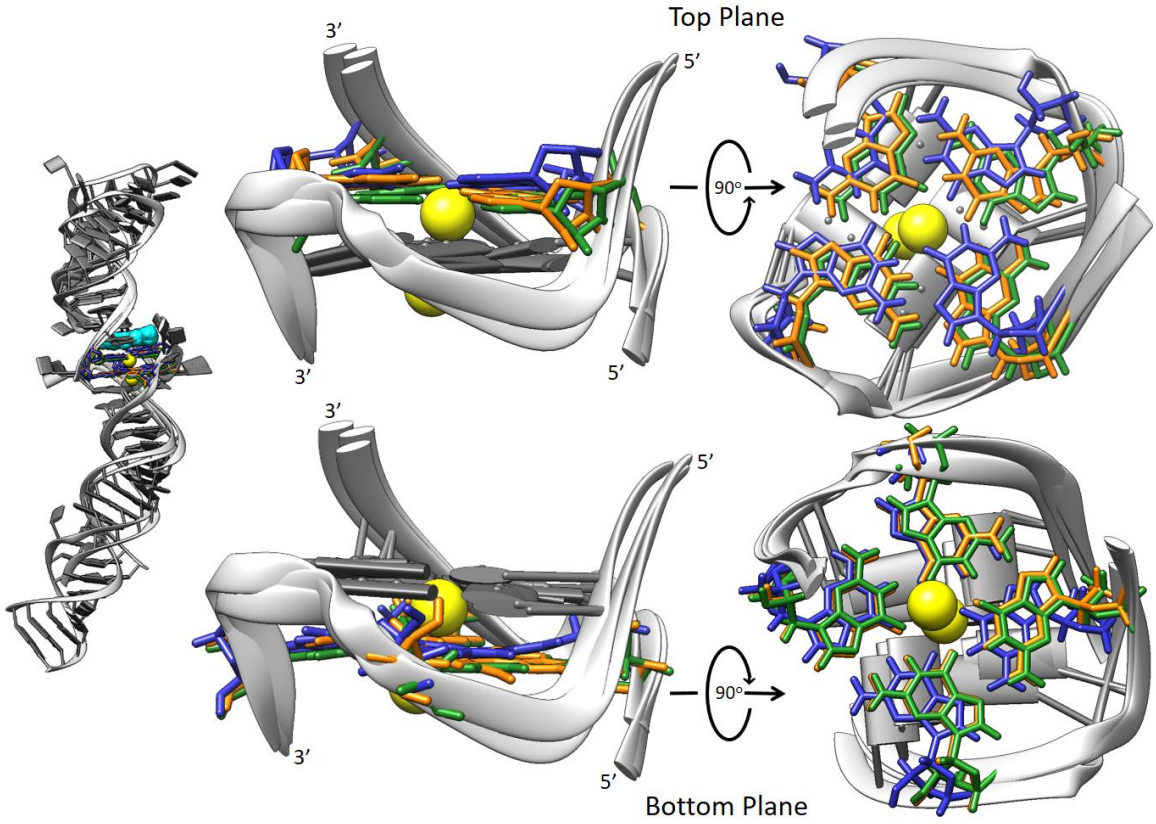



**Figure 3-S3 G-Quadruplex Structural Alignment.** Comparison of G-quadruplex planes of Spinach (4KZD), iSpinach (5OB3), and Broccoli model, colored green, orange, and blue, respectively. Top and bottom planes are individually colored against a grey neutral background for clarity.

# 4

# A Broccoli Aptamer Chimera Yields a Fluorescent $K^+$ Sensor Spanning Physiological Concentrations

This chapter has been reformed from its original publication in *Chemical Communications*:

**Savage, J. C**.; Shinde, P.; Yao, Y.; Davare, M. A.; Shinde, U. A Broccoli Aptamer Chimera Yields a Fluorescent $K^+$ sensor Spanning Physiological Concentrations. *Chem. Commun.* 2021, *57* (11), 1344–1347. https://doi.org/10.1039/d0cc07042d.

J.C.S. and U.S. designed the research; J.C.S. performed all biochemical experimentation; P.S. performed computation analysis; Y.Y. assisted with kinetic experimentation. M.A.D. provided thoughtful discussion and manuscript editing; J.C.S. wrote the paper; J.C.S. and U.S. edited and strategized the paper.

## 4.1 Abstract

The RNA aptamers Broccoli accept 2'fluorinated (2'F) pyrimidine nucleotide incorporation without perturbation of structure or fluorescence in the presence of potassium and DFHBI. However, the modification decreases Broccoli's apparent affinity for $K^+$ >30-fold. A chimera of Broccoli RNAs with mixed chemistries displays linear fluorescent gain spanning physiological $K^+$ concentrations, yielding an effective RNA-based fluorescent $K^+$ sensor.

## 4.2 Discussion

The RNA aptamer "Broccoli" was identified by SELEX to be an improved version of the aptamer "Spinach", the first fluorescent RNA equivalent of green fluorescent protein.[160,165] Fluorescent emission is contingent upon the association of three individual components, forming a "tripartite complex". This complex is composed of (1) a unique RNA G-quadruplex fold stabilized by (2) selective cation coordination, which promotes (3) the docking of fluorophore DFHBI (3,5-difluoro-4-hydroxybenzylidene imidazolinone) or select derivatives (Figure 4-1A).[176,177,317] Fluorescence of Broccoli or Spinach is absolutely contingent upon the proper association of *all* tripartite complex components. Though the structure of Broccoli has not been solved crystallographically, sequence alignment, spectroscopic studies, and computational modelling suggest Broccoli shares significant structural similarity with Spinach aptamers.[165,176,287,297,298,317]

While biological systems required billions of years to select riboswitch aptamers capable of sensing ions and metabolites, *in vitro* evolution experiments by SELEX allow

for rapid identification of a wider range of aptamer structures and functions.[3,4,6,318–320] Incorporation of synthetic nucleotide derivatives further expand the chemo-structural space RNA can adopt outside the biological constraints of enzymatic nucleotide synthesis and polymerization. Most published aptamers do not contain modified nucleotides as such chemistry restricts intracellular applications. However, the use of modified nucleotides in *de novo* aptamer selection and established aptamer tuning for *in vitro* utility remains underexploited and unexplored.

It has been previously demonstrated Spinach aptamer can be transcribed with 2'fluorinated pyrimidine nucleotide incorporation without perturbing fluorescence or structure.[300] Though numerous modified ribonucleotide chemistries are available, 2'fluorination is commonly used as it is inexpensive, accepted by certain RNA polymerases without loss of fidelity, increases polymer resistance to nucleases, and is well published.[97,98,119,120,300] Despite their sequence similarities at the G-quadruplex, Broccoli is smaller overall with increased thermal stability compared to Spinach, making it more attractive for *in vitro* applications.[165,317] Hence, we have explored the use of modified chemistry in Broccoli aptamer and discovered its distinct biophysical properties open new avenues for *in vitro* biosensor applications.

Broccoli shares significant structural similarity with Spinach aptamers, including the G-quadruplex motif unique to this family (Fig4-1A). Akin to Spinach, Broccoli transcribed with 2'F pyrimidine nucleotides—without any change in primary nucleobase sequence—folds correctly, binds DFHBI and $K^+$, and displays dramatically increased chemical stability in the presence of RNase A. (Figure 4-1 and Figure 4-S1). RNAs

transcribed *in vitro* with 2'hydroxylated (2'$OH_{py/pu}$) or 2'fluorinated pyrimidine (2'$F_{py}$) nucleotides (Figure 4-1B) were purified and compared by fluorescence and circular dichroism (CD) spectroscopies (Figure 4-1C and Methods). Fluorescence with 100mM KCl yielded spectra of 2'$F_{py}$ that display no change in excitation and emission wavelengths, albeit with a 25% reduction in fluorescent intensity compared to 2'$OH_{py/pu}$ (Figure 4-1C). CD reveals both Broccoli constructs share similar secondary structures (Figure 4-1C, *inset*), which are nearly superimposable on the spectrum of Spinach.[300,317]

Interestingly, despite initial similarities, obvious differences between Broccoli and Spinach emerge at sub-saturating $K^+$ concentrations as observed upon titration (Figure 4-1D). Broccoli 2'$OH_{py/pu}$ displays a higher affinity for $K^+$ than Spinach 2'$OH_{py/pu}$.[165,317] However, Broccoli 2'$F_{py}$ displays a lower affinity for $K^+$ than Spinach 2'$F_{py}$ (Figure 4-1D). The resulting apparent affinities between Spinach constructs differ less than 2-fold, but Broccoli constructs differ greater than 30-fold. The subtle sequence variation between Broccoli and Spinach imparts significant biophysical differences in cation association and tripartite complex stability, which is further exaggerated with 2'fluorinated pyrimidine incorporation.

To examine changes in affinity in finer detail, $K^+$ association kinetics were assayed as described (Figure 4-2; Methods).[161,300] Kinetics were performed with each RNA construct in isolation (Figure 4-2A). Fits of these data provide individual apparent association constants ($K_{obs}$), which each contain discrete component binding ($k_{on}$) and unbinding ($k_{off}$) rates within it. To determine those individual components, $K_{obs}$ values are plotted against their corresponding KCl concentrations. Linear regression provides $k_{on}$

(slope) and $k_{off}$ (y-intercept) (Fig. 2B). The results show pyrimidine 2'fluorination in Spinach increases $k_{off}$ with little effect on $k_{on}$, whereas the same modification in Broccoli both increases $k_{off}$ and reduces $k_{on}$. Individual kinetic values, and calculated dissociation constant $K_D$, roughly concur with $K^+$ affinities approximated in Figure 4-1D (Fig. 2B *inset*).

In attempt to rationalize the kinetic observations, a structural model of 2'$F_{py}$ Broccoli was constructed and molecular dynamic (MD) simulations were performed to directly compare RNA constructs, 2'-chemistries, and $K^+$-coordination states (Methods).[300,317] Root mean square deviation (RMSD) analysis provides insightful structural information, however it does not allow for analysis of structural convergences that may occur *during* a simulation. To capture this information we performed a principal component analysis (PCA) of G-quad motifs from all time-points of each simulation (Figure 4-S2). Projection plots of the Spinach and Broccoli structures onto the subspace spanned by their respective first two principal components indicate the presence of several distinct conformational populations dependent upon the construct, chemistry, and coordination state. 2'OH Broccoli appears to sample a reduced set of distinct populations than 2'OH Spinach, in both $K^+$-coordinated and Apo states, and is consistent with our experimental observations that 2'OH Broccoli is more stable than 2'OH Spinach. 2'$F_{py}$ Spinach and Broccoli display the same population distribution when $K^+$-bound, yet 2'$F_{py}$ Broccoli samples more distinct structural populations in its Apo state. We interpret this as evidence that 2'$F_{py}$ Broccoli is less stable than 2'$F_{py}$ Spinach, which is consistent with our experimental data (Figure 4-1D and Figure 4-2). This work corroborates to the

Broccoli structural model and provides an additional approach for future *in silico* characterization of sequence-optimized constructs.

The results hitherto establish the same chemical modification in these two highly similar RNAs does not yield effects proportional to their structural relatedness. While Fluorescent and CD spectroscopies show both Spinach and Broccoli RNAs retain their structural identity, and all four constructs display similar $K^+$ binding profiles upon titration, binding kinetics and apparent affinities of $K^+$ are altered drastically (Figure 4-1C/D and Figure 4-2 and Figure 4-S2). Such functional difference between Broccoli and Spinach in response to pyrimidine modification led us to investigate how the biophysical behavior of Broccoli could be exploited in a way Spinach could not.

The combined detection limits of the two individual Broccoli constructs span physiological $K^+$ concentrations of ~0.15 to 100mM, more than 20-fold the concentration the Spinach constructs cover (Figure 4-1D). We postulated a *mixture* of $2'OH_{py/pu}$ and $2'F_{py}$ Broccoli RNAs might produce a signal with a broad range of linear fluorescent gain, equivalent to the combined upper and lower detection limits of the individuals. Such results could open an avenue for development of an RNA-based fluorescent $K^+$ sensor, a tool with diagnostic potential for dysregulation of potassium-dependent homeostatic functions including blood pressure, muscular contraction, circadian rhythm, and others.[321]

Three ratios of Broccoli constructs were mixed—2:1, 1:1, 1:2—and KCl titrations were performed (Figure 4-3A). The widest detection range with a linear regression fit of $R^2>0.90$ was attributed to the 1:1 mixture (Figure 4-3A *inset*). To engineer the mixture

into a single construct which can be maintained under non-equilibrium conditions, complementary 3' extensions with unfavorable self-annealing were designed and added to each Broccoli RNA (Figure 4-3B and Figure 4-3C and Methods). The resulting construct contains two units of identical nucleobase sequence but is a hybrid of the two individual's chemistries, or a chimera. The proposed Chimera contains two distinct G-quadruplex motifs, each forming a separate tripartite complex with a respective $K^+$ ion and DFHBI pair. We postulated that annealing of the overhangs would not alter the biophysical behavior of the individual tripartite complexes, given the highly modular nature of RNA structural elements. As hypothesized, chimera formation does not change the $K^+$ titration profile or fluorescence intensity as compared to an un-annealed mixture, suggesting an absence of cooperativity in cation association after annealing (Figure 4-3C *upper right,* and Figure 4-3D) The Chimera displays KCl association kinetics that approximately split the difference between $2'OH_{py/pu}$ and $2'F_{py}$ monomers (Figure 4-S3). Similarly, as expected, the fluorescent properties of DFHBI in the $2'OH_{py/pu}$ and $2'F_{py}$ Broccoli subunits of the Chimera remain unchanged with excitation and emission wavelengths at 468 and 501nm, respectively. Thus both individual emissions coalesce into a single readout signal (Figure 4-3C *upper right*). The Chimera produces maximal fluorescence at pH 8, and displays Three ratios of Broccoli constructs were mixed—2:1, 1:1, 1:2—and KCl titrations were performed (Figure 4-3A). The widest detection range with a linear regression fit of $R^2$>0.90 was attributed to the 1:1 mixture (Figure 4-3A *inset*). To engineer the mixture into a single construct which can be maintained under non-equilibrium conditions, complementary 3' extensions with unfavorable self-annealing were designed and added

to each Broccoli RNA (Figure 4-3B and Figure 4-3C, and Methods). The resulting construct contains two units of identical nucleobase sequence but is a hybrid of the two individual's chemistries, or a chimera. The proposed Chimera contains two distinct G-quadruplex motifs, each forming a separate tripartite complex with a respective $K^+$ ion and DFHBI pair. We postulated that annealing of the overhangs would not alter the biophysical behavior of the individual tripartite complexes, given the highly modular nature of RNA structural elements. As hypothesized, chimera formation does not change the $K^+$ titration profile or fluorescence intensity as compared to an un-annealed mixture, suggesting an absence of cooperativity in cation association after annealing (Figure 4-3C *upper right,* and Figure 4-3D) The Chimera displays KCl association kinetics that approximately split the difference between $2'OH_{py/pu}$ and $2'F_{py}$ monomers (Figure 4-S3). Similarly, as expected, the fluorescent properties of DFHBI in the $2'OH_{py/pu}$ and $2'F_{py}$ Broccoli subunits of the Chimera remain unchanged with excitation and emission wavelengths at 468 and 501nm, respectively. Thus both individual emissions coalesce into a single readout signal (Figure 4-3C *upper right*). The Chimera produces maximal fluorescence at pH 8, and displays greater than 70% fluorescent signal intensity in up to 50% cell culture serum for over 60 minutes (Figure 4-S4). Linear regression of KCl titration shows the Chimera to be suitable for visualizing $K^+$ over a wide concentration range, with signal changes observable as low as 70μM (Fig. 3D *inset*). Additionally, the Chimera displays sufficient selectivity and $K^+$ preference over alternate cations, confirming its potential as a $K^+$ sensor (Figure 4-3C).

The sensor was tested in an *in vitro* cell system to observe bulk $K^+$ flux upon hypotonic lysis of adherent HEK293T cells (Figure 4-4). Cultured cells at various degrees

of confluence were lysed in a $K^+$-free hypotonic buffer containing DFHBI and Broccoli chimera. Lysis-induced efflux in $K^+$ concentration were observed (Figure 4-4 *upper inset*) and interpolated from a standard curve (Figure 4-4, and Methods). Changes in calculated $K^+$ concentrations are directly proportional to the observed cell density prior to hypotonic lysis, validating sensor applicability in an *in vitro* cellular system (Figure 4-4 *upper inset*).

In this work only select riboses were altered, not the nucleobase sequence. 2'Fluorinated pyrimidines result in retained structure and fluorescence but with dramatically altered cation association. Numerous alternately modified nucleotides are commercially available, and may prove more appropriate for certain applications.[103] Additionally, chemical synthesis—as opposed to transcription—allows for combining chemistries with site-specific incorporation.[322] The fluorescent property and small size of Broccoli's G-quadruplex makes it possible, albeit tedious, to test every combination of base and chemistry at each position. On the basis of these results we posit two avenues of aptamer modification—sequence and chemistry—can be used synergistically for finer adjustment of functional tuning for new applications. There are several examples of previously reported aptamer fusions.[180,183,323] However, these fusions are the result of connecting the primary sequences of two aptamers together to be transcribed as a single unit. There are advantages of this approach, though it does restrict the fusion product to a single nucleotide chemistry. This work represents, to our knowledge, the first example of a mixed chemistry chimera, generated *in vitro* via multiple synthesis reactions, expanding the aptamer toolkit.

Functionally, the RNA Chimera is a valuable addition to the available methods for fluorescent $K^+$ sensing, which contains several small molecule chelators, a single protein-based construct employing FRET or a split-eGFP, and three DNA constructs, which are the most chemically similar (Table S1).[237,324–331] The DNA sensors PSO-1,2 and TBA also employ G-quadruplex folds for $K^+$ docking, though, unlike the Chimera, the fluorescence of these DNA sensors is based on termini-linked FRET pair interaction upon $K^+$-induced G-Quadruplex folding.[237,238,331,332] While this detection scheme has a lower background signal than the Chimera, the linear range of function $K^+$ detection of these tools is within a single order of magnitude, ~1-10 and ~100-200mM.[237,332] The complementary molecular geometry of nucleic acid G-Quadruplexes and desolvated potassium is sufficiently favorable and selective to suggest any nucleic acid sensor engineered for detecting $K^+$ in the millimolar range will likely employ the motif. Tuning such sensors will be heavily based on modulating the number of G-Quadruplex planes and optimizing proximal residues, changing the number of bound potassium and their binding affinity. In the case of this Chimera, additional tuning can be achieved by optimizing construct preference for alternate DFHBI derivatives with differential affinity, excitation/emission, brightness, and photo-bleaching.[161,333]

Considering the transcriptional limitations discussed above, visualizing *subcellular* $K^+$ gradients with this Chimera may prove difficult. However, it is reasonable to presume future re-programming experiments could yield a Broccoli-derived RNA sensor with similar properties to the Chimera but composed exclusively of hydroxylated nucleotides. Such work has been explored conceptually for the Spinach aptamer for sensing of $Pb^{2+}$

ions.[202] Transcribing Spinach with modified nucleotides was demonstrated to accelerate $Pb^{2+}$ association, increase response fidelity, and enhance chemical stability, all without affecting $Pb^{2+}$ preference over $K^+$ by 1000-fold.[300]

Broccoli has been shown to accept various DFHBI derivatives, producing fluorescent complexes in a range of colors.[160,162,298,334] Sequence optimization to increase fluorophore selectivity of Chimera constituents could yield chimeras with FRET capacity or simultaneous sensing of multiple targets.[160,162,298] Similarly, because individual units contain two distinct G-quadruplex motifs that form independently, a re-tuned Broccoli chimera could be exploited to investigate basic RNA-cofactor biophysical interactions and tripartite complex formation studies. Chimeras of mixed chemistry provide vast opportunity for custom-tailored sensor development while expanding the aptamer canon for functional modification and rational tuning.

Nucleic acid aptamers have ligand affinity with selectivity that rival antibodies. Using aptamers to detect or "sense" their molecular targets in ways similar to antibodies has been commonplace for some time. However nucleic acid sensors for the accurate *quantification* of their ligands are less developed, for which fluorescent aptamers are an ideal system. We have found the unaltered sequence of Broccoli aptamer allows for quantification of $K^+$ when mixed with a modified $2'F_{py}$ construct of the same sequence. A chimera of the two annealed RNAs produces a stoichiometric unit that can be maintained in non-equilibrium conditions. While the properties of the monomeric subunits are unaltered, their individual signals coalesce into a single readout spanning a linear range of $K^+$ detection of nearly three orders of magnitude. This platform provides a promising

foundation for future ion-selective optimization, which could yield a flight of RNA sensors for the real-time fluorescent quantification of cellular cation flux.

The authors thank Drs. Dave Farrens, Matt Thayer, and Larry David for their thoughtful discussion and generous use of scientific equipment. This work was supported by funds from the NIH and the Medical Research Foundation of Oregon.

## 4.3 Materials & Methods

**In vitro Transcription and purification:** RNA transcripts were prepared using linear dsDNA PCR templates. Sequences of the forward DNA strand are as follows (T7 promoter is underlined, construct in bold font): Spinach 5'—GCGCGCGAATTCTAATACGACTCACTATA**GGAGGACGCGACCGAAATGGTGAAGGACGGGTC CAGTGCGAAACACGCACTGTTGAGTAG**AGTGTGAGCTCCGTAACTGGTCGCGTC—3'; and Broccoli 5'—GCGCGCGAATTCTAATACGACTCACTATA**GGGAGACGGTCGGGTCCAGGCACACAAAAATGTT GCCTGTTGAGTAGAGTGTGGGCTCC**—3'. Prior to transcription, PCR templates were gel purified, phenol:chloroform cleaned, and desalted, similar to RNA clean up (described below). Transcription reactions were assembled at room temperature in the following order, to final concentrations of: water, 1X T7 RNAP transcription buffer (NEB), 24mM $MgCl_2$, 4mM each NTP (2'OH purines from NEB, 2'F pyrimidines from Trilink Biotechnologies), 15-25ng/µl dsDNA, 5mM DTT, 1U/µl murine RNase inhibitor (NEB), 2.5mU/µl Yeast Inorganic Pyrophosphatase (NEB), 5U/µl WT T7 RNA polymerase (NEB) or 1.25U/µl mutant T7 R&D polymerase (Lucigen, Inc). RNA was transcribed at 42$^{o}$C for 2-5

hours. Transcriptions were quenched with DNase I treatment, and observed on a 4% agarose sodium borate gel, stained with Sybr Gold nucleic acid dye. Target transcript bands were purified by gel excision and electro-eluted into 3.5kDa MWCO dialysis tubing. RNA was concentrated with a 3.5KDa MWCO concentrating spin column (Millipore) or Speed Vac, and cleaned with acidified phenol:chloroform:isoamyl alcohol (Ambion). RNA was then desalted into 5mM Tris (pH 8.0) using a 7kDa MWCO desalting spin column (Zeba, Thermo Scientific), and quantified by UV-absorbance at 260nm and fluorescence (Qubit HS RNA kit; Invitrogen), and stored at -80$^{o}$C. Prior to experimental use of frozen stocks, RNA was folded following a basic protocol: dilution into reaction buffer, then successive thermocycler incubations at 95$^{o}$C (20sec), 50$^{o}$C (10sec), 37$^{o}$C (10sec), addition of DFHBI, 4$^{o}$C (≥30min).

**Fluorescence:** Single sample fluorescent measurements were performed on a Photon Technologies International QM-1 steady state fluorescent spectrophotometer. All fluorescent measurements were performed with the following solution with various cation additions: 0.5µM RNA, 20mM Tris (pH 8.0), 5mM $MgCl_2$, 20µM DFHBI. Samples were read at 20$^{o}$C, held constant by a circulating water bath, and excited at 468nm in a 2mm quartz fluorescence cuvette. KCl titrations were performed by folding RNA in the presence of various KCl concentrations in individual reaction tubes. Samples were then incubated on ice for 2hr to ensure equilibrium was reached.

**Chemical Stability Assays:** *RNaseA-dependent:* 0.5µM RNA was incubated with 1.25mU/µl purified RNase A (Qiagen). At stated time points, 5µl of reaction was removed, quenched with 10mM EDTA in 2X gel loading dye, and frozen at -20$^{o}$C. Samples were

analyzed via 4% agarose gel stained with Sybr Gold nucleic acid dye. *pH-dependent:* 0.5µM Chimera was assayed for fluorescent yield at 9 pH points in a universal buffer composed of 20mM acetate, 20mM HEPES, 20mM Borate, 100mM KCl, 5mM $MgCl_2$, 20mM DFHBI. Samples were incubated at stated pH for 30minutes, then analyzed simultaneously in a SpectroMax M2$^e$ Microplate reader (Molecular Devices), Ex/Em at 468/501nm, respectively. *Serum-dependent:* 0.5µM Chimera was assayed for fluorescence over time in increasing concentrations of cell culture serum, 0, 10, 25, and 50%. Serum was composed of a 1:1 mixture of Bovine Growth Serum and Fatal Calf Serum (Atlanta Biologicals). Fluorescence was measured in a SpectroMax M2$^e$ Microplate reader (Molecular Devices). Samples were Top-read at 25$^o$C, every 60 seconds for 60min with Ex/Em at 468/501nm. Data plotted in GraphPad Prism 5.

**Tripartite complex association kinetics:** Kinetic fluorescence was performed by rapid, in-cuvette mixing of folded Apo-state 0.25µM RNA in buffer with a 10% volume of 10X KCl in water. Emission was read at 501nm every 29.75 seconds with a 0.25 second shutter exposure. Association curves of individual KCl concentrations were fit in GraphPad Prism 5 using the following built in equation for specific binding:

$$Y = [(B_{max})(X)]/(K_D+X) \qquad (1)$$

This fit assumes any non-specific binding does not contribute to fluorescence. Individual kinetic fits provided K observed ($K_{obs}$) rate values. To determine the individual kinetic parameters of the $K_{obs}$ (binding and unbinding, or $K_{on}$ and $K_{off}$, respectively), $K_{obs}$ values for each KCl concentration were plotted against their respective KCl concentration. Linear

regression of the resulting plot provides $K_{on}$ and $K_{off}$, via the slope and y-intercept, respectively.

**Circular Dichroism:** Measurements were collected with an Aviv 215 CD Spectrophotometer at 20°C. RNA samples were assayed in 10mM Cacodylate (pH 7.4), 100mM KCl, 5mM $MgCl_2$. RNA was folded as described. 135ng/µl RNA samples were scanned in 1mm quartz CD cuvette from 320 to 190nm by a 0.2nm step, with a 3 second integrated read per step. Data presented as average of three individual scans. Control scans of buffer were subtracted from spectra. Values plotted in GraphPad Prism 5.

**Construct Design:** Complementary 3' extensions were engineered into Broccoli construct by PCR. The two sequences added to the 3' terminus of dsDNA T7 Broccoli constructs: (1) 5'-**AGACAGACAGT**-3' and (2) 5'-**ACTGTCTGTCT**-3'.

**Cell Culture:** HEK293T cells were cultured in DMEM with 10% fetal bovine serum, penicillin and streptomycin, and supplemented 5mM glutamine, at 37°C with 5% CO2. Cells were maintained between 40-80% confluent between passages.

**Hypotonic Lysis:** HEK293T cells were cultured to various degrees of confluence in 96-well, clear-bottom, black-walled, tissue culture plates. Cells were quickly and gently washed once with 1X PBS then once with hypotonic buffer (20mM Tris pH 8.0, 5mM $MgCl_2$, 0.01% Triton X-100), before 200ul of hypotonic lysis solution containing 3µM DFHBI, 0.5µM Broccoli chimera RNA, and 80mU/µl murine RNase Inhibitor was added to each well. Plate immediately placed in SpectroMax $M2^e$ Microplate reader (Molecular Devices). Samples were bottom-read at 25°C, every 60 seconds for 60min with Ex/Em at 468/501nm.

Positive control wells of varying KCl concentrations with no cells were simultaneously read, as were negative control wells containing either no cells, no RNA, or neither of both. Positive control samples used fit to a simple nonlinear regression model in GraphPad Prism 5:

$$Y = Y_{intercept} + X*Slope \qquad (2)$$

Unknown samples values were interpolated from the resulting plot in GraphPad Prism 5.

**Molecular Dynamic Simulations:** Broccoli structural model adapted from Savage et al, 2020.[335] 2'fluorinated Broccoli structural model was created in the YASARA software package as described.[300,335] Simulations were performed for 50ns at 298K with an improved Berendsen thermostat for temperature control and a "densostat" for pressure control, with an integration time step of 2.5fs. Ionic concentration was 0.9% w/v NaCl concentration; electrostatics were evaluated using Particle Mesh Ewald with an 8Å cut off for long range interactions.[336] A modified AMBER14 force field for RNA was employed.[283]

**Model Analysis:** *G-Quadruplex bonding analysis:* The distances between hydrogen bond donors and acceptors for each canonical hydrogen bond within the two-plane G-quadruplex motif was analyzed for each snapshot and plotted as a time trace. Distances within 2Å were considered bound via hydrogen bonding, and interaction distances outside this range were considered unbound. *Principle Component Analysis:* The core residues of the G-quadruplex and DHFBI binding pocket were extracted for every simulation snapshot and aligned to the starting broccoli model, to remove any translations and rotations that occurred over the course of the simulation. The MD

trajectories for the aptamer with and without $K^+$ were combined to form a single trajectory upon which a principal component analysis was conducted using ProDy.[337] Cluster analysis was conducted on the projections of all snapshots onto the first three principal components using Wolfram Mathematica FindCluster. The structures corresponding to cluster averages were then back-calculated using the coordinates of individual elements within the cluster in order to find potential intermediate states.

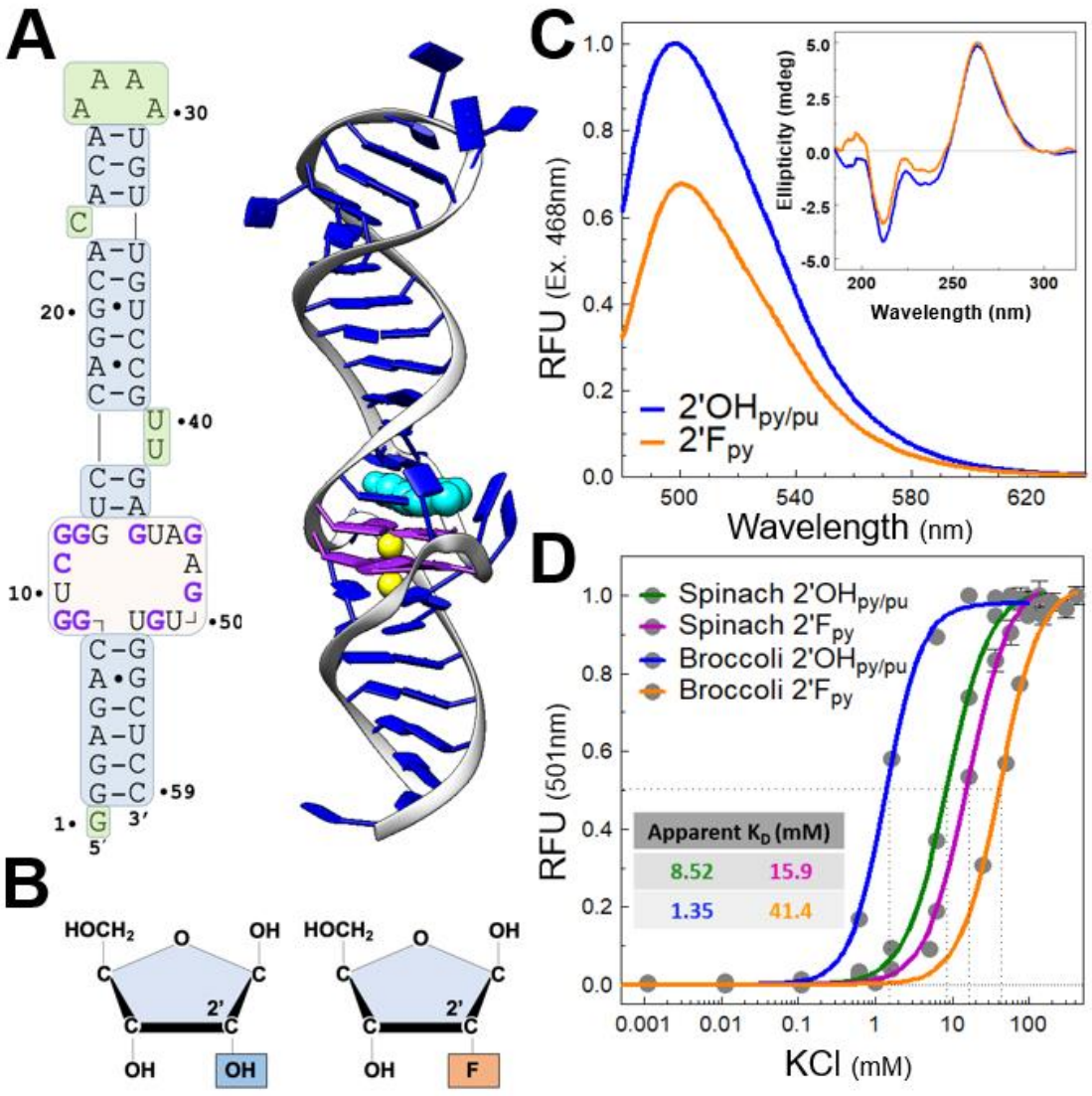



**Figure 4-1 Broccoli Aptamer Accepts Fluorinated Pyrimidine Nucleotides.** (A) a 2D projection with sequence, and a 3D structural model of Broccoli RNA. In both representations, G-quadruplex guanines are purple, coordinating K+ ions are yellow, and DFHBI is cyan. Broccoli structure adapted from Savage et al., 2020.5 (B) 20 hydroxylated and 20 fluorinated pyrimidines differ only by the functional group present at the ribose C2 position. (C) Fluorescence comparison between 2'OHpy/pu and 2'Fpy Broccoli RNA with 20mM DFHBI and 10mM KCl. (C; inset) Circular dichroism (CD) spectra of 2'OHpy/pu and 2'Fpy Broccoli show the two constructs share nearly identical secondary structures. (D) Fluorescence-based KCl titrations with RNAs provides an approximate $K_D$ for K+ at the G-quadruplex. Despite the structural similarity between Spinach and Broccoli, the two RNAs display dramatic differences in $K^+$ association when 2'Fpy nucleotides are incorporated, with a <2-fold difference between Spinach constructs, but >30-fold between Broccoli constructs. *(inset)* Apparent $K_D$ values for $K^+$ determined from fits.

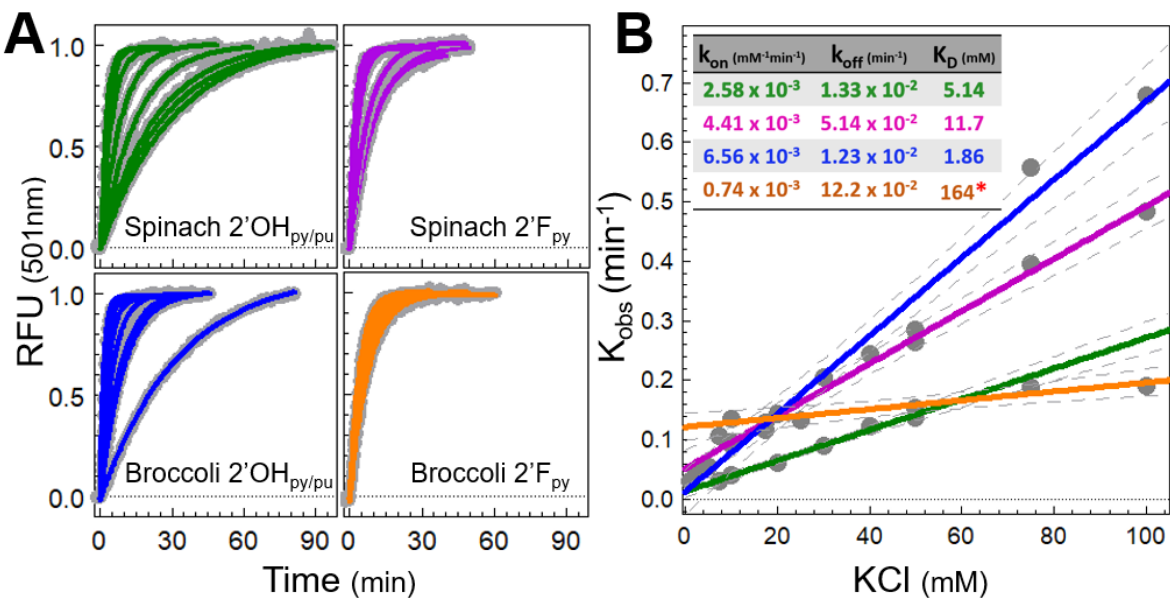



**Figure 4-2 Kinetics of Fluorescent Tripartite Complex Formation in Spinach and Broccoli.** (A) Individual kinetic traces and colored fits of K+ association, assayed by fluorescence. Each trace represents a single KCl concentration. All traces performed with constant RNA and DFHBI concentrations (see Methods). (B) Fits of each kinetic trace in (A) yield a $K_{obs}$ apparent rate constant. When each $K_{obs}$ is plotted against its respective KCl concentration and fit, individual kinetic components can be extrapolated; $k_{on}$ and $k_{off}$ determined from the slope and y-intercept, respectively. Despite their sequence and structural similarities, 20 F incorporation of Spinach reduces $k_{off}$ with negligible effect of $k_{on}$, whereas the same modification in Broccoli both reduces $k_{on}$ and increases $k_{off}$, explaining affinity differences observed in Fig. 1D. (inset) Explicit values determined from fits in (B) with the same color coordination. Red asterisk denotes high KD value for Broccoli 2'Fpy which may be unreliable with this particular technique. KCl additions too close to an apparent KD yield non-linear $K_{obs}$ values.

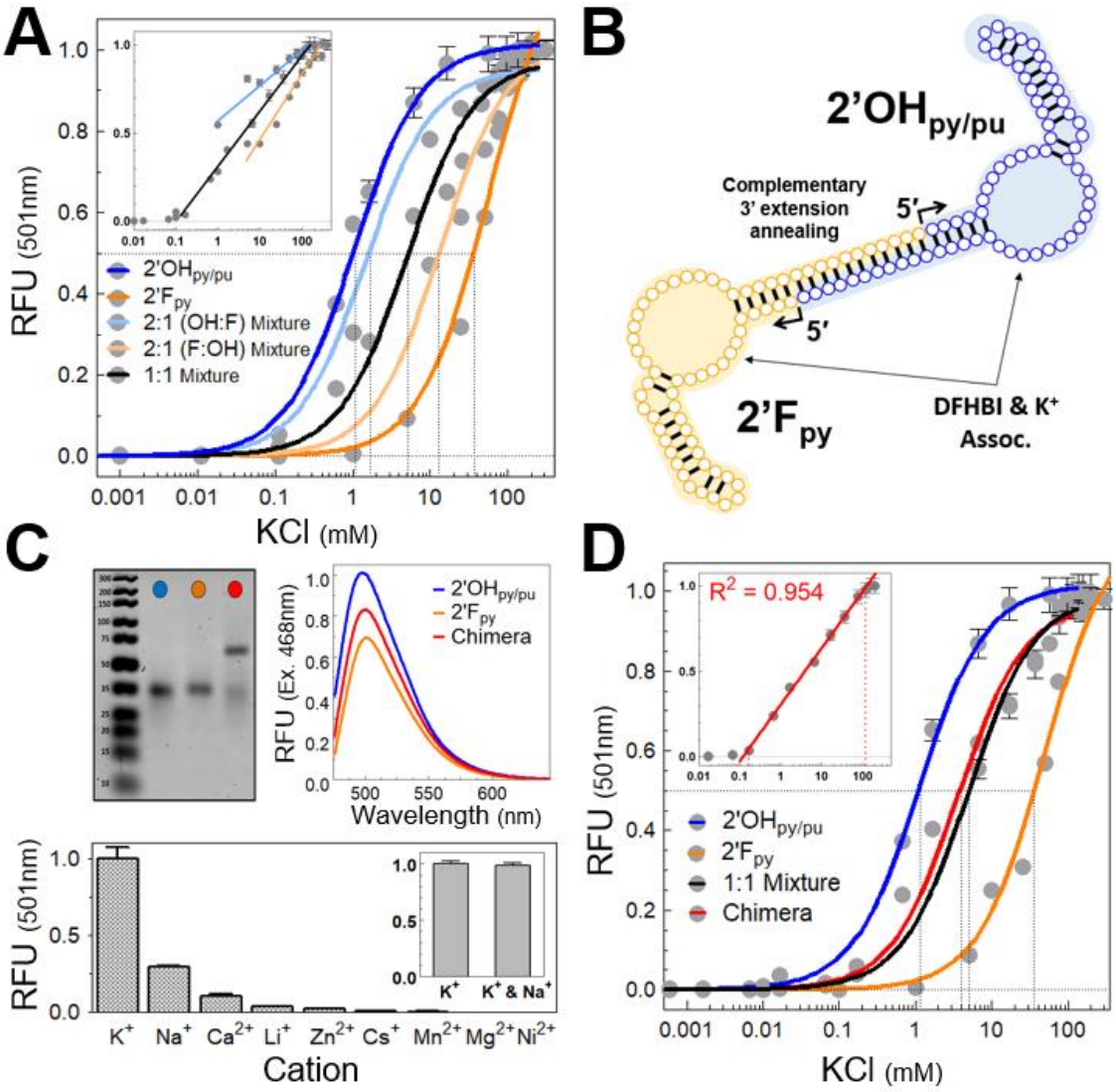



**Figure 4-3** (Legend on next page)

**Figure 4-3 Broccoli RNA Chimera Construction.** (A) fluorescence-based KCl titration of various ratios of monomeric Broccoli constructs, showing predictable curve shifting based on relative 20 chemistry composition. (*inset*) Linear regression of KCl concentrations was performed, containing the most points while retaining an $R^2$>0.90, of which the 1:1 ratio displays widest detection range. Inset axes identical to outer graph. (B) 2D schematic of chimera design, showing WC stems, G-quadruplex motif and complementary 30 extensions (see Methods for sequences). Chimera contains two G-quadruplex domains forming two independent tripartite complexes. Overhang sequences in Materials & Methods (C) (*upper left*) Agarose gel EMSA shows mass shift of chimera compared to the two monomeric chemistries. (*upper right*) Comparative fluorescence spectra of chimera and monomeric chemistries. Monomers at 0.5mM and chimera at 0.25mM, to account for (2) G-quadruplex motifs per chimera. Gel lane colors in EMSA correspond to fluorescence signal colors. (*lower*) Chimera fluorescence in the presence of various mono- and divalent cations (all chloride salts) normalized to KCl signal. Each cation present at 50mM in buffer (Methods). (*lower; inset*) Chimera fluorescence in the presence of 50mM $K^+$ is unchanged with or without the presence of competing 50mM $Na^+$. (D) Fluorescence-based KCl titrations of 2'OHpy/pu, 2'Fpy, 1: 1 mixture, and annealed chimera. Both the 1: 1 mixture and the annealed chimera split the difference between two monomeric chemistries with respect to apparent KD and broader detection range. Difference between 1: 1 mixture and chimera is negligible, confirming 30 extension and annealing does not alter K+ association and tripartite complex formation for individual chimeric subunits. (*inset*) Linear regression shows the chimera to have a linear range of detection from ~150μM to ~100mM. Inset axes identical to outset.

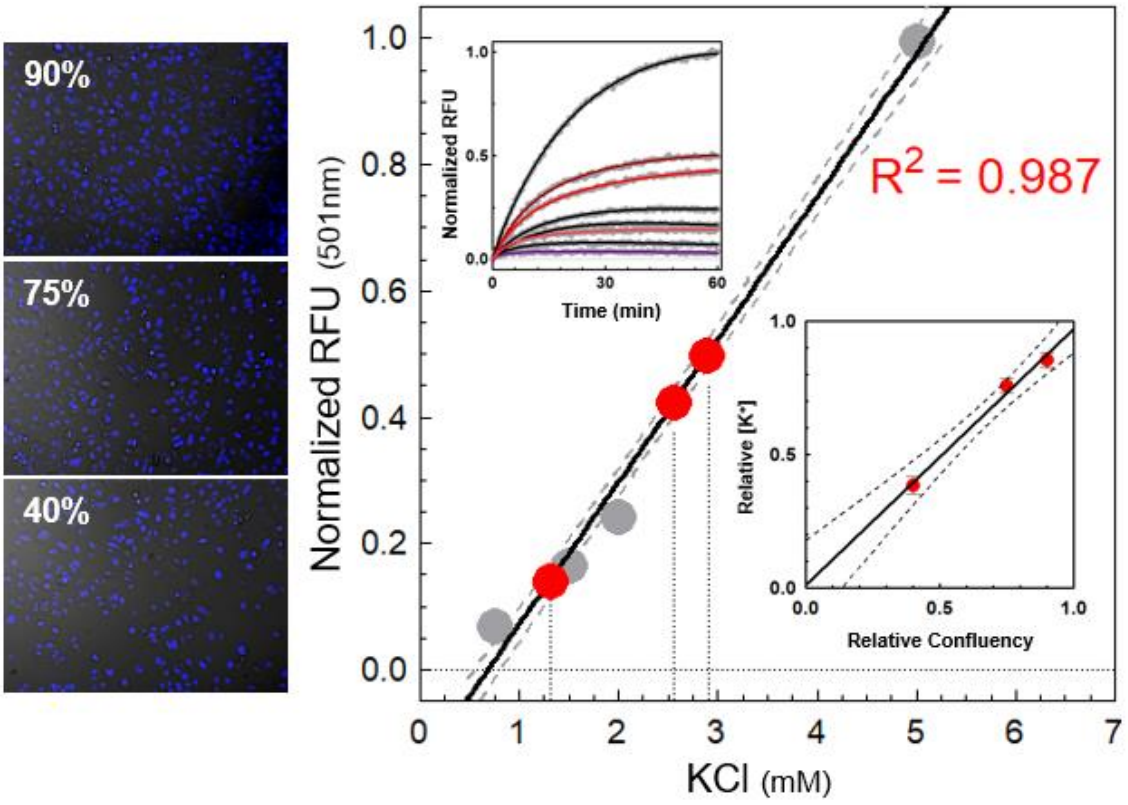



**Figure 4-4** ***In Vitro*** **Detection of $K^+$ Flux by Hypotonic Lysis.** HEK293T cells were cultured to ~40, 75, 90% confluence and media replaced with KCl free lysis buffer (see Methods). (*upper inset*) Raw fluorescence signal gain was collected (red fits) next to KCl positive standards (grey fits) and negative controls (purple fit). Unknown K+ concentrations (red dots) were interpolated from a standard curve (grey dots) generated from nonlinear regression of positive controls. (*lower inset*) Approximate cell confluence and interpolated $K^+$ concentrations were directly proportional over the range tested.

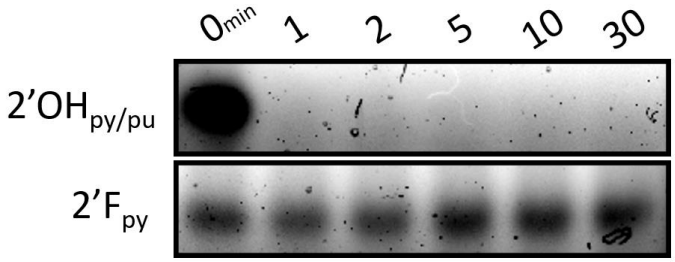



**Figure 4-S1 RNase stability of Broccoli RNAs.** Each construct was incubated with purified RNase A for given time before being quenched with EDTA and loading dye. Broccoli transcribed with 100% 2'OH rNTPs (2'OHpy/pu) was completely degraded within 1min of RNase A exposure, whereas the incorporation of 2'F pyrimidine NTPs resulted in no visible degradation for more than 30minutes.

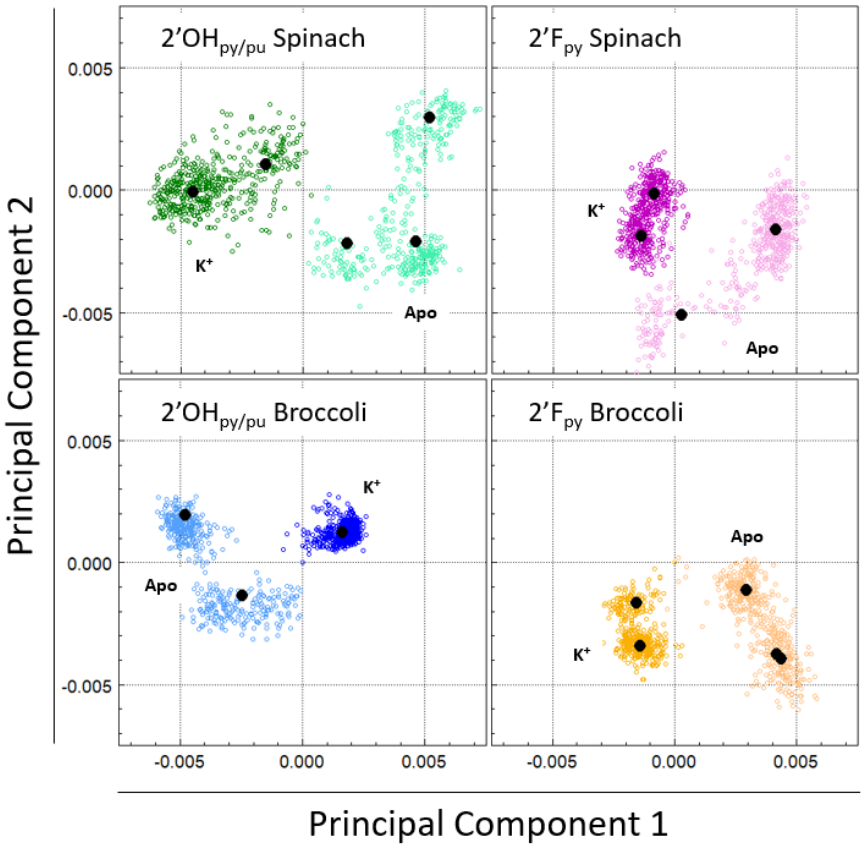



**Figure 4-S2 Principal Component Analysis of Molecular Dynamic Simulations.** 250ns Molecular dynamic simulations were performed to observe structural differences between 2'chemistries in Spinach and Broccoli over time. Principal component analysis of G-quadruplex motifs under 3 variable conditions are presented (RNA identity, 2'chemistry, and K+ coordination). Each colored dot corresponds to one G-quadruplex structure, sampled every 100ps; Black dots correspond to the structural average of each unique population.

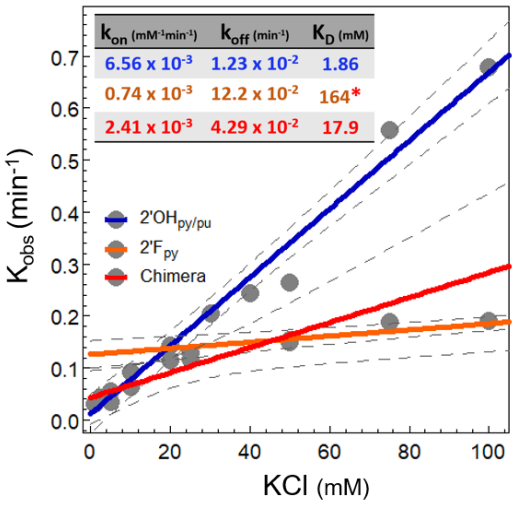



**Figure 4-S3 Chimera K+ Binding Kinetics**. Following the identical protocol described for 2'OHpy/pu and 2'Fpy, 0.25µM Chimera was incubated with various KCl concentrations, and tripartite complex formation was observed by fluorescence. Fluorescent traces of individual KCl concentrations were fit to acquire $K_{obs}$ values. These $K_{obs}$ are plotted above against their respective KCl concentrations, overlaid with identical data for 2'OHpy/pu and 2'Fpy monomers adapted from figure 2. The Chimera behaves as a mixture of the two component monomers, as expected.

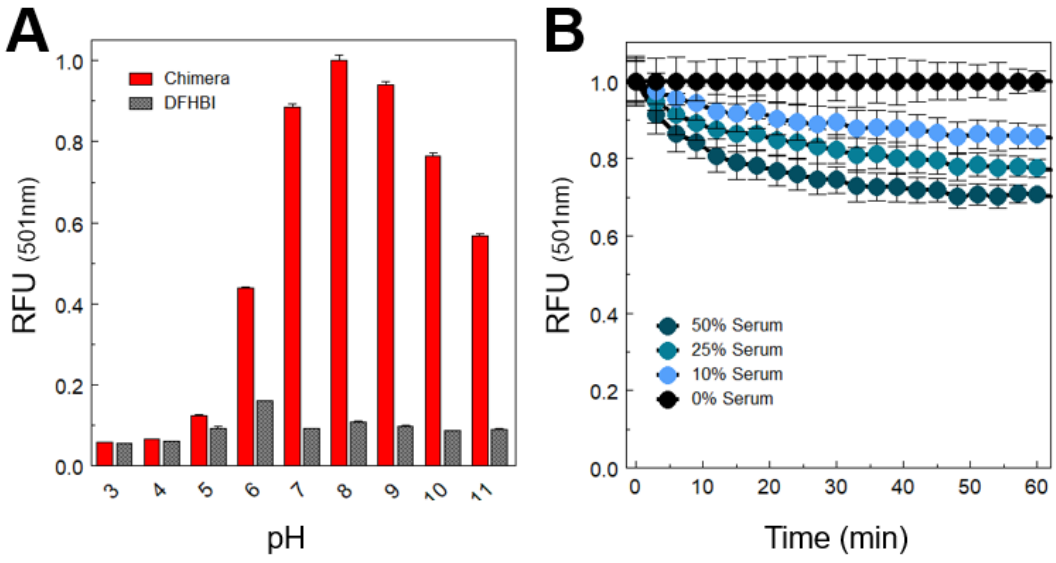



**Figure 4-S4 pH and Serum Chemical Stability of Broccoli Chimera.** (A) pH-dependent fluorescence of Chimera shows construct displays maximal complex fluorescence at pH 8. (B) Chimera fluorescence was observed over time in the presences of increasing cell culture serum concentrations, 10, 25, and 50% by volume. Buffer and KCl concentrations held constant for each serum concentration. Serum was a 1:1 mixture of Bovine Growth Serum and Fetal Calf Serum.

| Name | Chemistry | $K_D$ (mM) | Functional Range (mM) | λ Fluorescent Emission (nm) |
|---|---|---|---|---|
| NK1 | Small Molecule | 200 | 10-50 | 572 |
| KS6 | Small Molecule | — | 30-500 | 540 |
| TAC | Small Molecule | — | — | 540 |
| TAC-Red | Small Molecule | — | 0-50 | 574 |
| TAC-Chrimson | Small Molecule | — | 20-60 | 600 |
| PBFI | Small Molecule | 8 | — | 500 |
| CDC222 | Small Molecule | 15-34 | 2.5-6.5 | 445 |
| GEPII (family) | Protein | 0.1-20 | 0.01-1000 | 550 |
| KIRIN1 | Protein | 1.67 | 0.1-10 | 530 |
| GINKO | Protein | 0.4 | 3-100 | 514 |
| PSO-1 | DNA | — | — | 581 |
| PSO-2 | DNA | 7.3 | 0.5-10 | 480 |
| TBA | DNA | — | 100-200 | 583 |
| Broccoli Chimera | RNA (Modified) | 8-18 | 0.15-100 | 501 |

**Table 4-S1 Comparative Traits of Published Fluorescent $K^+$ Sensors.** Values as reported in stated supplemental references.

# 5

# Summary and Conclusions

## 5.1 Summary of Results

The cumulative body of presented research is centered on a single pyrimidine nucleotide chemical modification, incorporated into the sequence of two highly related fluorescent RNA aptamers, producing differential and unexpected changes in cation affinity, kinetics, and resulting fluorescent emission. The 2'fluorination of pyrimidine nucleotides was initially explored in the aptamer Spinach as proof-of-principle for its capacity to enhance RNA chemical stability in the presence of nuclease and $Pb^{2+}$ for *in vitro* sensing application (Figure 2-1 and Figure 2-2).[300] The analyte for the only published Spinach-based fluorescent cation biosensing research prior to the body of work presented herein.[202] Interestingly, the 2'F pyrimidine modification resulted in accelerated cation association kinetics and tripartite complex formation, increasing the accuracy of the Spinach sensor on short time scales, enhancing sensor function and utility (Figure 2-3). Spinach and Broccoli RNA aptamers were observed to have >90% sequence identity in their critical G-Quadruplex motif (Figure 3-2).[317] However, differences between the two

RNAs in critical base pairing interactions proximal to the G-Quadruplex appear to have substantial influence on the stability and dynamics of G-Quadruplex residues, resulting in disparate cation association and tripartite complex stability (Figure 3-1). The interesting results with Spinach and $Pb^{2+}$ lead to the questions of if, how, and to what degree the same modification would alter the properties and biophysical behavior of the closely related Broccoli aptamer.[338] $K^+$ titrations of Spinach and Broccoli revealed differences of $K^+$ affinity between the two aptamers with native chemistry which are exaggerated when chemically modified with 2'F pyrimidines (Figure 4-1). The combined upper and lower limits of $K^+$-dependent fluorescence of native and modified Broccoli span roughly three orders of magnitude (~0.15-100mM). $K^+$ titration in a mixture of the two Broccoli construct chemistries produced a curve that spans the K+ concentration range of constructs in isolation, with a significantly broader range of linear detection covering human physiological *intra*- and *inter*- cellular $K^+$ concentrations (Figure 4-3). Adding 3' complementary extensions and annealing these individual constructs into a single stoichiometric unit produced an mixed chemistry chimera with an identical $K^+$ titration curve as the mixed monomers, which validated the Chimera as a fluorescent K+ sensor that can detect physiologically relevant $K^+$ concentration changes in real time (Figure 4-3 and Figure 4-4).

## 5.2 Continued Development of the Broccoli Chimera $K^+$ Sensor

### *5.2.1 Covalent Circularization*

While the Broccoli Chimera $K^+$ sensor contains 2'F pyrimidine nucleotides and shows considerably enhanced chemical stability in the presence of RNases, the Chimera

is currently generated only by base pair annealing of 3' complementary extensions. This approach of connecting two RNAs is quick, simple, and specific, however covalent circularization of the chimera would enhance sensor reliability, raise the construct's melting temperature, abrogate vulnerability to *exo*nucleases, and add a sense of polished completeness. Unfortunately, circularization is not as simple as adding RNA ligase to the annealing reaction. T7 RNA polymerase incorporates trailing heterogeneous 3' extensions up to ten residues long in run-off transcription reactions, making ligations that depend on perfect overhang complementarity impossible without prior processing.[339,340] A solution is needed that produces defined and homogenous 3' termini, ideally reliant on post-transcriptional enzymatic manipulation. Two solutions based on ribonuclease cleavage have been widely discussed, one by RNase H1 and the other by self-cleaving ribozymes.[341,342] The preferred option is the self-cleaving ribozyme and involves the 3' fusion (and co-transcription) of ribozyme sequence, which folds and cleaves at a defined residue upon transcription.[342–344] This approach has been successfully performed by others using the hepatitis delta virus ribozyme (HDVr).[345] HDVr is preferred for its small size, catalytic activity in otherwise denaturing conditions (up to 5M urea or 18M formamide), and fast activity with several divalent cations ($Ca^{2+}$, $Mg^{2+}$, $Mn^{2+}$, and $Sr^{2+}$).[346,347] However, HDVr has not been confirmed to retain catalytic activity with incorporated nucleotide modifications as other, less ideal, self-cleaving ribozymes such as Hammerhead have.[117,348] As long as nucleobase identity directly 5' of the cleavage site is purine (which is tolerable), HDVr should theoretically retain activity when transcribed with 2'F pyrimidines (however, structural analysis reveals the critical general acid C75 in

HDV to favor a 2'*endo* pucker, which may not be allowed if C75 is 2'fluorinated).[349–352] Should 2'$F_{py}$ HDVr prove inactive or impractical, native chemistry HDVr or alternate ribozyme can be added in *trans*, however the substrate RNA will remain fluorinated at pyrimidine nucleotides which must be acceptable for the protocol to work.[353] While theoretically simple, reaction and recovery efficiencies of cleavage, phosphorylation, annealing, and *then* RNA ligation will add up. It may be found that, despite the high costs of RNA synthesis (particularly so with modified nucleotides), synthesizing appropriately phosphorylated Broccoli constructs for ligation is worth the price.

### *5.2.2 Detection of Cellular K+ Flux*

Ideally the Broccoli Chimera will prove to be as effective for visualizing $K^+$ flux in living, dynamic cellular systems as it is *in vitro*. The first step to achieve this is to successfully deliver the Chimera to a subcellular location of interest. Modified nucleotide incorporation prevents transcription by endogenous transcription machinery, as discussed (section 1.3.5), therefore the Chimera must be added exogenously. Depending on the subcellular location, this could range from challenging to extremely difficult. The simplest proof-of-principle for visualizing $K^+$ flux is to label the plasma membrane in a low KCl solution, then induce membrane permeabilization by mechanical or chemical means and observe $K^+$ efflux by Chimera fluorescence. To this end Broccoli Chimera was successfully anchored to the plasma membrane of adherent cultured cells and visualized by fluorescence microscopy, generally following recently established protocols of fluorescent cell labelling (Figure 5-1).[354–357] Briefly, the native chemistry 2'$OH_{py/pu}$ Broccoli constituent RNA was transcribed *in vitro* with 5-azido-PEG4-uridine triphosphate in a 1:12

molar ratio with standard UTP, yielding RNA with an average of one azido functional group per transcript. This RNA was annealed to 2'$F_{py}$ Broccoli, after which the Chimera was coupled to Cholesterol-$PEG_{2K}$-DBCO (dibenzocyclooctyne) via azide-alkyne strained cycloaddition CLICK chemistry.[358] The final construct was then added to adherent HEK293T cells, washed, and visualized in high KCl solution by fluorescent microscopy (with apparent stability for at least 30minutes post-addition). In addition to simply verifying that the construct is truly imbedded in the lipid bilayer, several basic questions must be answered before this construct can be used for any cellular $K^+$ quantification: (i) does azido-UTP incorporation alter the $K^+$ titration profile of the chimera? If so, is the new range of detection biologically relevant? (ii) What is the construct half-life in cell suspension present with RNAses? Is the half-life commensurate with the inherent kinetic limitations of Chimera-$K^+$ association? (iii) What is an appropriate positive control probe (a crown ether or alternative biologic polymer $K^+$ sensor) which can be similarly conjugated to cholesterol and anchored? These data are clearly *very* preliminary. However, simply seeing some degree of cell labelling is a satisfying and necessary first step.

## 5.3 Questions and Future Directions

### *5.3.1 Alternative Modifications in Spinach Family Aptamers*

The entirety of this body of work was performed utilizing a combination of unmodified RNA and 2'fluorinated pyrimidine nucleotides. While there remains substantial room for additional exploration with variations of the 2'F modification theme, there are hundreds of alternative chemical modifications that can be directly

incorporated into the native sequence of Spinach and Broccoli RNAs. Some may not yield any appreciable change in aptamer function, and many are likely to completely diminish it. However, it is reasonable to presume at least a few will yield something interesting or valuable. 2'fluorination is a particularly bio-compatible modification, with steric and electrostatic properties quite similar to natural ribonucleotide chemistry (section 1.3.4). However even 2'fluorination is not universally accepted; even other members of the Spinach such as Corn do not function when transcribed with 2'F pyrimidines (*unpresented*). The difficulty then is deciding on which particular modifications to pursue, and how to screen those modified constructs for function with reasonable ease and speed. Fluorescent aptamers such as Spinach and Broccoli are ideal models for scaled exploration of chemical modifications. Preparing 96- or 384-well microplates for multiplexed fluorescent assays that vary with RNA modification, or fluorophore and cation identity is not prohibitively expensive or unrealistically time consuming.

One subset of nucleotide modifications that may be particularly interesting to explore in the Spinach aptamer family are inherently fluorescent nucleoside analogs, especially 2-aminopurine (2AP) (adenine analog) and 6-methylisoxanthopterin (6MI) (guanine analog), which are minimally invasive as they do not rely on bulky aromatic functional groups and maintain consistent base pairing preference of native nucleotides (A-T and G-C). Both analogs display high fluorescence quantum yields (0.68 and 0.7 for 2AP and 6MI, respectively), that vary with local nucleotide environment making them conducive for studying nucleic acid hybridization, folding, stability, and intermolecular interactions.[359–361] 6MI can engage in G-Quadruplex formation, resulting in a quenched

and red shifted fluorescent emission that can be used to assay motif formation and stability.[362] While only two independent studies have been published, 6MI incorporation appears to have quite variable effects on G-Quadruplex formation and stability.[363,364] 6MI has been observed to destabilize certain sequences while stabilizing—and greatly accelerating $K^+$ association—in alternate structures, in a cation-dependent manner. Incorporating 2AP proximal to the G-Quadruplex motif, or at inter-plane positions, may allow for direct fluorescent observation of cation-association and G-Quadruplex folding. Additionally, 2AP and 6MI have spectral overlap conducive for FRET pairing, which could enhance signal to noise and open up additional nuanced G-Quadruplex tuning (Ex/Em of 2AP and 6MI are 315/360nm and 370/430nm, respectively). Should these or other fluorescent nucleotide analogs prove viable for observing cation association and G-Quadruplex in Spinach RNAs, their capacity for *in vitro* cation sensing would be greatly expanded.

### *5.3.2 The Other Half of the Tripartite Complex: Fluorophore-RNA Interactions*

As discussed, the tripartite complex of Spinach family aptamers involves the proper association of folded RNA, cation, and fluorophore (section 1.6.5). This body of work has focused exclusively on the cation association half, leaving fluorophore docking unexplored. Valuable biochemistry and resulting photophysical properties of the DFHBI-Spinach interaction have been elucidated.[161] However the study was just one of several fluorophores bound to one of several RNAs. Within the context of the work presented here, one has to wonder if and how chemical modifications of RNA affect DFHBI (and fluorophores) affinity, association kinetics, or photophysical properties such as quantum

yield. Differential behavior of DFHBI with native chemistry or 2'$F_{py}$ modified Broccoli RNA doesn't affect the ultimate functionality of the Broccoli Chimera as a sensor, as maximal DFHBI excitation/emission in both chemistries is 468/501nm, respectively, producing two individual emissions that coalesce into a single inseparable readout. However, this may not be the case with all HBI-derivative fluorophores that could be used, or all alternatively explored nucleotide modifications in future sensor development. A wider exploration of Spinach and Broccoli with other HBI-derivative fluorophores may yield a more favorable fluorescent combination for sensing $K^{+}$, $Pb^{2+}$, or $Sr^{2+}$, even within the same paradigm of exclusively modified 2'F pyrimidines.

### *5.3.3 The Implications of Base Interactions Proximal to G-Quadruplex Structures*

The body of work presented here justifies the continued exploration of modified nucleotide chemistry incorporation in G-Quadruplex structures for the tuning of cation interaction preferences, rates, and affinities. Just as interesting, however--and technically simpler to investigate--is the prospect of *sequence-base* tuning of G-Quadruplex structures that is suggested by the subtle differences in cation binding and thermal stability displayed by the native chemistry 2'$OH_{py/pu}$ Spinach and Broccoli constructs. Sequence analysis revealed the two RNAs are highly conserved and share a nearly identical G-Quadruplex motif structure that differs by a single outwardly oriented residue (Figure 4-2 and Figure 4-3). Despite this, their respective tripartite complexes display substantial differences in thermal stability, and cation affinity and kinetics (Figure 4-1). Directly proximal to their shared G-Quadruplex motifs Spinach and Broccoli subtly differ

in duplex base pairing composition. Within the same four pair duplex stretch, Spinach engages A-A and U-U whereas the same loci are C-G and A-U in Broccoli (Figure 4-3). Stronger WC interactions, particularly the closer G-C, directly proximal to the nearly identical G-Quadruplexes are proposed to cinch up the motif with differential stabilities, restricting the flexibility of G-Quadruplex residues in both Apo and cation-coordinated states. Switching two base pairs alters $K^+$ apparent affinity ~5-fold and tripartite complex thermal stability ~5°C. The base identity changes are quite modest, yet produced dramatic effects. While it is difficult to predict how other combinations of targeted mutations may affect Broccoli or Spinach behavior, it would not be difficult to engineer and test many of these combinations. The variable tolerance of proximal sequence provides substantial room for rational engineering.

The G-Quadruplex-proximal differences between Spinach and Broccoli exist on one side of the motif, however the alternate side may be just as influential on tripartite complex stability if not more so. The opposing cluster of residues has an additional role in stabilizing the DFHBI binding pocket, possibly providing any mutation a disproportionately large influence on tripartite complex stability. Intra-G-Quadruplex motif residues (i.e., inter-plane residues within the motif) are also of interest for functional tuning, particularly the conserved C11-G53 (Broccoli numbering). Mutations that produce a G-U wobble or a weaker A-U may destabilize Broccoli sufficiently to produce Spinach-like tripartite complex behavior. Alternatively, replacing conserved G-U, U-U, or A-U base pairings on both sides of the G-Quadruplex with G-C where possible may stabilize constructs substantially, increasing tripartite complex stability and effective

sensing range at lower K+ concentrations. Should this conceptual approach engender the creation of a flight of functionally distinct Broccoli constructs that more narrowly sample a wider K+ concentration range, the 2'F chemical modification used to make the Broccoli Chimera can perhaps be abandoned allowing for direct cellular transcription for *in cellulo* applications. Taking inspiration from *Filonov et at* (2014), it should be possible to construct a single native chemistry construct with multiple Broccoli G-Quadruplexes, each with varying proximal residue identity, producing a single coalesced fluorescent signal comparable to the mixed chemistry Broccoli Chimera.[165]

Mutational tuning additionally opens up avenues for reducing the overall size of both constructs. The G-Quadruplex motif of Broccoli and Spinach aptamers are off center, being closer to the hairpin in Spinach and closer to the termini in Broccoli (Figure 4-3). An optimal hybrid structure that combines the shorter termini-duplex of Broccoli and the shorter hairpin-duplex of Spinach can possibly reduce the overall size of a hybrid structure. This would be a boon for endogenous RNA labelling and visualization of subcellular trafficking (section 1.6), as well as *in vitro* sensing applications discussed.

The prospect of proximal duplex manipulation has additional interesting implications for model G-Quadruplex structures such as TBA or a minimized human telomere (hTelo), which are effectively composed of 100% quadruplex stacks. PSO-py and other TBA $K^+$ sensors from Takenaka and colleagues have quite narrow ranges of $K^+$ detection, each within a single order of magnitude.[240–242] Based on work with Spinach family tripartite complex formation, one would expect extending the duplex off of TBA termini to decelerate both $K_{on}$ and $K_{off}$, making the overall trend in $K^+$ affinity difficult to

predict. Favorable duplex formation should induce partial G-Quadruplex folding even in the absence of $K^+$, likely slowing $K^+$ association by steric occlusion. Upon eventual association however, favorable WC duplex should dramatically stabilize the now-favorable G-Quadruplex structure, slowing dissociation as well. It is difficult to imagine the addition of differential terminal duplex to TBA can be used to *both* increase and decrease sensitivity, however either change has potentially valuable applications.

The same logic tuning synthetic G-Quadruplex cation sensors by proximal duplex engineering raises an important question about the general utility of small aptamers and truncated telomere structures as models of biologically relevant G-Quadruplex behavior generally. It is well established that nucleotide chemistry (DNA or RNA), cation identity, inter-plane loop size, and overall construct size (i.e., number of planes) affect physiological G-Quadruplex topology, folding, and stability.[226,365–368] There has been little investigation, however, into the role of proximal sequence on G-Quadruplex formation or stability. Two studies have been published that intentionally approach this question, however they share the same caveat. *Ren et al* (2002) extended the 5' terminus of a minimized hTelo sequence (d(GGGTTA)$_3$GGG)) with a poly-dT$_{12}$ run.[369] This construct was then annealing to an exogenous poly-dA$_{12}$ sequence prior to biophysical analyses of CD and UV-absorbance (295nm) thermal melting, MD simulations, and differential scanning calorimetry (DSC). From these data they concluded there to be "little or no crosstalk" between folded G-Quadruplex and the flanking duplex poly A:T. *Konig et al* (2013) found the opposite to be true.[370] Using a similar scheme of 5' sequence extension and exogenous reverse complement addition (of mixed GC or AT rich content), they observed

a distance-dependent trend of exaggerated mismatch duplex destabilization by hTelo G-Quadruplex (and i-motif) formation. The criticism of both of these studies is the use of exogenous reverse complement oligonucleotides to create a proximal duplex on one side of the G-Quadruplex. How does this replicate a biological system? In their discussion, *Konig et al* (2013) states,

> "As many putative G-quadruplex-forming and IM [i-Motif] sequences are surrounded by stretches of duplex DNA, these results are of particular interest and could help to address whether these higher-order substructures are likely to form in a physiological context." (Pg. 7458)

Despite this recognition, the authors fail to note that telomeric G-Quadruplex exists within extremely long stacks and is flanked on one end by duplex DNA, or that alternate physiological G-Quadruplexes (such as c-MYC) exist within nucleic acid strands that extend from both termini of the G-Quadruplex motif.[210]

The issue of proximal duplex can be inferred from the dramatic differences in cation kinetics reported for optimized small G-Quadruplex structures (such as TBA, minimalist hTelo, and the like), and those observed for G-Quadruplexes that exist within extended, duplexed nucleic acid such as native chemistry Spinach and Broccoli aptamers. It is farcical to draw firm conclusions about natural G-Quadruplex behavior from experiments on chemically modified, synthetic 2'$F_{py}$ Spinach or Broccoli aptamers using $K^+$, let alone $Pb^{2+}$. However, the magnitude of difference in cation association between TBA or hTelo and Spinach family aptamers is indicative of the issues present when using any of the three as models of natural *in situ* G-Quadruplex behavior. Every single

published report that uses one of these model G-Quadruplexes reports association kinetics faster than Spinach or Broccoli. Not every publication can be uniformly compared, as some report $K_{obs}$, some $K_{on}$, with differential use of $Pb^{2+}$, $K^{+}$, and $Na^{+}$ cations. The trends, however, are consistent. Model G-Quadruplexes TBA and minimized hTelo associate with cations faster than Spinach and Broccoli, ranging from a modest 4-fold to a surprising 40,000-fold.[227,228,371–374] G-Quadruplex topologies display differential stabilities and folding kinetics.[367,375] However, no report exists showing anything close to a 40,000-fold difference attributable just to topology. This difference in kinetics between Spinach and TBA aptamers is likely the result of the proximal duplex structure on both sides of the Spinach G-Quadruplex, which partially pre-folds the motif, slowing cation association by steric occlusion. The G-Quadruplex structure of TBA is unrestricted in this sense, and can fold and unfold in a true two-state manner following cation addition or sequestration. Between Spinach and TBA or minimalist hTelo, biologically relevant structures such as the c-Myc or KRAS *intra*-promoter G-Quadruplexes appear to be more contextually reminiscent of Spinach. It seems likely that biophysical investigations performed looking at such traits as cation association and cooperativity, thermodynamic and kinetic stabilities, folding intermediates, or inter-topological conversion using TBA as a model do little to inform biologically relevant G-Quadruplex behavior, and instead thoroughly describe the conditional properties of TBA alone.

Researchers utilize small model G-Quadruplexes for two reasons (which are really two sides of the same reason): (i) other members of the field use them; (ii) their small size and high percentage of G-Quadruplex guanine composition allow them to be directly

observed by DSC, CD, and UV-absorbance without interfering signal from additional non-G-Quadruplex nucleic acid sequence or structure; in short, ease of experimental execution dictates the model, not biological relevance.[231–233] For this second reason, studying G-Quadruplex *in situ* is difficult (debate still exists whether G-Quadruplexes exist at all *in vivo).*[376] Spinach is not necessarily a good G-Quadruplex model. Nor are TBA or minimalist hTelo necessarily bad G-Quadruplex models. The dramatic differences in cation association between them simply suggests the potential importance of proximal nucleic structure, be that duplex or something else, that has not been addressed--in either experimentation or discussion--in the use of small, optimized, independent G-Quadruplex structures for modelling natural G-Quadruplex biophysics *in situ.*

## 5.4 Conclusion

This body of work has provided: (i) the first (and second) instance of inducing new function in an existing aptamer structure by chemical modification of nonterminal nucleotides (without a modification of primary sequence); (ii) a sequence analysis of the relation between the highly related Spinach family aptamers and the first 3D structural model of Broccoli from which rational biochemical mechanism and structural manipulation can be proposed; (iii) the first generation of an aptamer chimera of mixed nucleotide chemistries, engineering via a universal approach applicable to numerous aptamer paradigms; (iv) the first RNA $K^+$ biosensor combined with the widest range of fluorescent $K^+$ detection of any nucleic acid cation sensor yet published, covering nearly three orders of magnitude. This work stimulates questions to be pursued specific to the

Broccoli Chimera and the Spinach aptamer family specifically, broadly on the influence of sequence identity and structural proximity in G-quadruplex-cation interactions and its relationship to G-Quadruplex behavior with implications for semi-rational tuning, and methodologically on the merits of structural recycling and functional reprogramming of aptamers as a conceptual alternative to *de novo* aptamer generation.

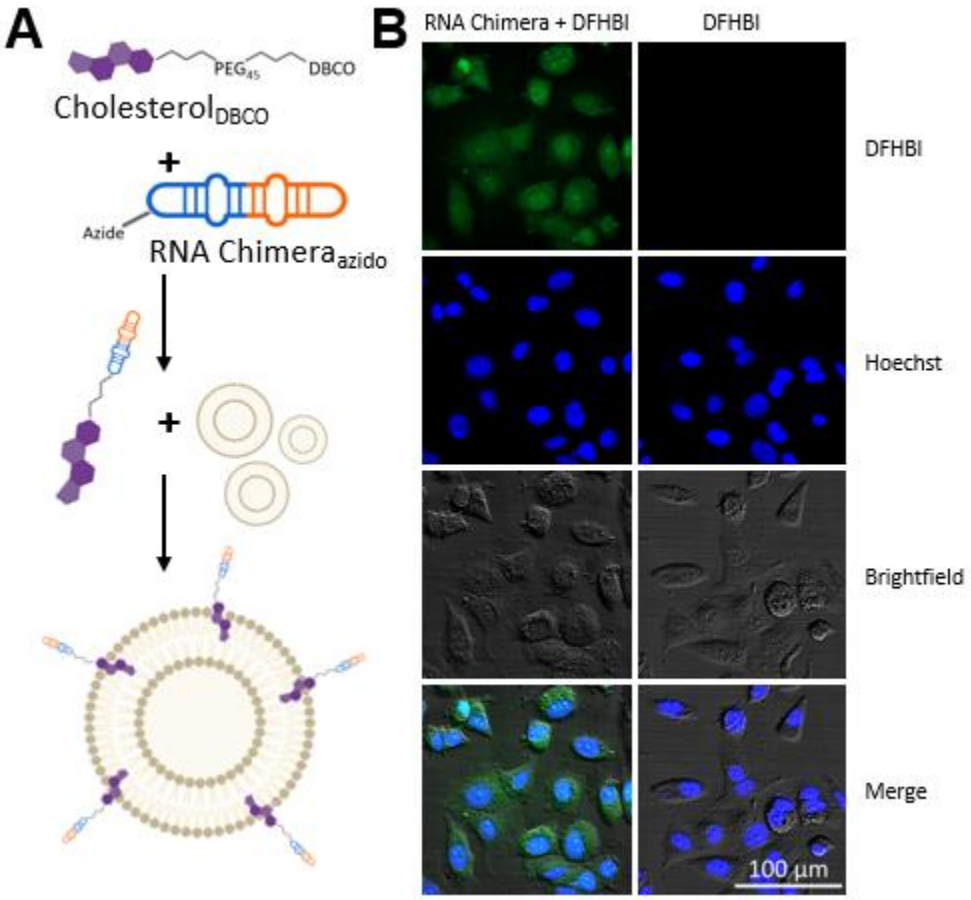



**Figure 5-1 Preliminary Cell Labelling for Detection of Cellular $K^+$ Flux.** (A) Basic outline of lipid-linked Chimera for membrane anchoring. Commercially sourced cholesterol-$PEG_{2K}$-DBCO is CLICKed to Broccoli Chimera composed containing an additional 5-azido-PEG4-Uridine nucleotide by standard azide-alkyne strained cycloaddition in aqueous solution. The conjugated construct is added to cells for outer leaflet membrane anchoring. (B) Microscopy images of HEK293T adherent cell labelling with cholesterol-linked Chimera. Left panels are labelled with RNA, right are not, both are incubated with DFHBI fluorophore. 100um scale bar applies to all images.

# 6

# Perspective

## 6.1 Introduction

The conception of nucleic acid aptamers and their *in vitro* selection synergized the biochemical knowledge and biotechnological advancements in nucleic acid research from the preceding decade into a single technique. Immediately after the initial development of the Systematic Evolution of Ligands by EXponential enrichment (SELEX), researchers began designing customized derivations of the protocol.[3–6] Every variation attempts to favor selection of certain functions and biophysical characteristics, or to streamline the selection process generally making it more likely to be successful. To date—depending on one's definition of unique—hundreds to possibly thousands of unique SELEX derivations have been described, yielding an aptamer canon diverse in structure, function, and chemistry. Aptamers continue to hold value for their unique biophysical properties, in-house identification, and alluring blend of history and futurology, stimulating hypotheses on the origin of life while embodying the promises of the synthetic biology revolution.

SELEX exploits the defined and favorable conditions of *in vitro* biochemistry to expedite the evolution of complex structures, accelerating processes which took billions

of years in a primordial soup down to days and weeks in a test tube. However, unlike natural molecular evolution which uncovers novel function through the mutability of extant species via conservative and spontaneous change, SELEX protocols consistently reach all the way back to the random sequence soup to initiate their aptamer selections. We refer to this *de novo* approach as "classical SELEX". Theoretically, a classical SELEX progenitor library contains every possible sequence (and thus every possible structural permutation) that can ever exist for a given sequence length. Additionally—and also theoretically—only and all favorable sequence-structures are enriched as a classical SELEX experiment progresses, a product of the carefully defined selection conditions and unbiased molecular biology reactions. Regrettably, these theoretical assumptions are rarely so clean, making classical SELEX experiments laborious and risky with unknowable durations. Despite these hurdles the diversity of function and application for which aptamers have been successfully exploited is enormous and the allure of SELEX as potent now as at its conception.

There is clearly a strong precedent and sound logic to perform a selection experiment following classical SELEX approaches. Yet *in vitro* evolution of protein, ribozyme engineering, and numerous natural evolutionary and biological examples suggest it may not be necessary that *all* selections against protein ligands begin with a completely random library. Is there a reliable way to exploit aptamer structures with proven stability and functionality? Has the aptamer structural lexicon of *in vitro* evolved species reached sufficient diversity to enable structural recycling following the logic of natural evolution? The answer to these questions must in large part be yes, and further,

embracing this alternative, though by no means mutually exclusive, approach to *de novo* selection will facilitate a rapid expansion of successful aptamer generation from a broader swath of research laboratories.

## 6.2 Structural Recycling and Targeted Epitope Re-selection

### *6.2.1 Classical SELEX*

Aptamer selection by classical SELEX begins with an extremely heterogeneous randomized progenitor library of nucleic acid sequences, within which exists an exceptionally small population that inherently possesses a desired biophysical trait. All SELEX paradigms aim to expand this small desirable population through iterative cycles of classical Darwinian selection—favorable selection followed by PCR replication—enriching sequence representation over time. Progenitor libraries typically contain degenerate lengths of >25 nucleotides, providing $>10^{15}$ unique aptamer sequence-structures. Intuition might suggest ever increasing sequence diversity can only aid in aptamer discovery, however a balance must be struck between heterogeneity, concentration, volume, and efficiency of recovery at each step in every cycle. That balance is not easy nor is there a universal rule for determining its optimal condition. While plenty of successful selections have been published, it is generally accepted that classical SELEX fails to identify a genuine aptamer in 50-70% of attempts made.[40,377] Any conceptual advancement which increases the ease of success and thus the ubiquity of aptamer identification and utility in basic research is a scientific boon.

### *6.2.2 Progenitor Libraries & Heterogeneity*

The foundational and often repeated power of SELEX is the immeasurably vast sampling of structural space. Theoretically, every sequence and structure that can possibly exist for a given polynucleotide length is present in equal proportions, competing under the same unbiased selective pressure, resulting in enrichment based purely on thermodynamic favorability. However, representative equality breaks down when longer sequences are employed to sample more complex sequence space. For scale, a 1mM solution of a 70nt randomized sequence must be $10^{22}$ liters (roughly ten times the volume of water on earth) to contain just one copy of each possible species. At such lengths, an infinitesimal fraction of the available sequence space is actually being sampled in a laboratory setting. The inherent complications of size and scale at the onset of selection are exasperated by additional factors that contribute to unselected and thus detrimental homogeneity. These include nonrandom library synthesis, off-target interactions, non-targeted compositional drift, and biased amplification of small or differentially stable structures during transcription and PCR.[14,93–95] It is clear that substantial sequence bias exists during classical SELEX, even under idealized conditions. Some bias may need to be ignored, some can be accounted for and mitigated, and some could and should be strategized.

One way of strategizing a bias is to look at natural functional RNA as a guide. While we cannot observe the evolutionary progression of past RNA sequences, analysis of contemporary RNAs with wide functional and phylogenetic divergences reveals substantial and universal sequence biases.[378] This may suggest the available sequence space has been widely explored and particular compositional biases are (or were)

universally adaptive. Conversely, and more likely, substantial sequence space remains unexplored, and a surprisingly narrow swath is sufficient to provide the functional diversity we observe in contemporary life.[21] Either way it may be superfluous to even attempt to sample much of the sequence space in an *in vitro* selection, with total randomness—where even possible—unnecessarily burdensome. One logical conclusion from this is to incorporate the compositional biases—or further yet, complete subdomains—of extant RNAs into an otherwise random *in vitro* selection library, restricting heterogeneous populations to favorable, though still novel, sequence space.

### 6.2.3 S*trategic Bias & Directed Evolution*

The earliest attempt of motif doping in aptamer selection for protein binding was the intentional incorporation of a fixed stem-loop sequence in a selection against GTP.[379] The constant stem-loop sequence was flanked by two 26nt random runs, hypothesized to bias the representation of structurally complex aptamers in the progenitor library. The selection was successful with the engineered pool heavily favored over the completely random control pool that was simultaneously selected. Structural determination validated the engineered stem-loop to be integral to GTP binding.[380] An alternative approach employing a position-specific pyrimidine/purine compositional bias favoring stem-loop enrichment has been employed during library synthesis, similarly revealing structural doping to be beneficial for selecting structurally complex aptamers (ruff 2010).[381] G-Quadruplex motifs can similarly be enriched during library synthesis. While more structurally complex, G-Quadruplex enrichment paradigms are actually simpler requiring only a guanine bias.[382] Though a slightly more hands-on approach allows for

topological tuning of G-Quadruplex motifs, tipping equilibriums towards parallel, antiparallel, and mixed G-Quadruplex orientations (McManus 2013).[383]

The previous examples utilized sequence enrichment to aid selection of affinity binding aptamers, which can be as small as 15nt and still obtain complex structure.[234] Natural ribozymes, on the other hand, require at least ~40nt though often exceed 150nt (however, extremely small catalytic aptamers have been evolved *in vitro*).[14,22,33,34,384] *In silico* analysis of randomized sequence pools suggests long lengths are required to achieve, with any statistical likelihood, the structural diversity necessary for complex catalysis.[92,385] Additionally, natural structured RNAs are significantly more thermodynamically favorable than the average structure found in random sequence populations of equivalent length (or compared to their own scrambled sequence), further highlighting their rarity.[386,387] The hurdles of identifying large and complex structures *de novo* has been overcome by incorporating entire domains of natural ribozymes into a selection, a process called directed evolution. Within a constant scaffold of evolutionarily vetted structure, partially randomized catalytic core domains can be readily evolved for novel catalytic function.[388–394]

### *6.2.4 Aptamers & Riboswitches*

Current aptamer selection protocols necessitate the *de novo* identification of *both* an appropriate primary sequence *and* a stable structure. A selection approach that exploits the stability of a vetted and characterized aptamer by reprogramming it to selectively bind an alternate ligand needs only identify a favorable primary sequence at select positions. Such a dimensional reduction should simplify the selection process and

increase the likelihood of success. This logic of structural recycling by targeted epitope selection was successfully implemented once for the generation of novel riboswitches by Batey and colleagues in 2017.[395] Synthetic small-molecule binding aptamers display unpredictable efficacy when transferred to cytosolic conditions for *in vivo* use. To surmount this obstacle, Batey and team turned from *de novo in vitro* selection to exploiting evolutionarily validated structures with robust folding and a stable architecture. They attempted to reprogram several structurally conserved riboswitch classes by selectively randomizing the ligand-binding epitopes in 3WJ domains while holding the conserved stem and pseudoknot domains constant (Figure 6-1). Two riboswitches (c-di-GMP and guanine) and a ribozyme (hammerhead) were selectively randomized then re-selected against alternate ligands, 5HTP or L-tryptophan. Selectivity was successfully redirected, and structural analysis post-SELEX revealed the overall pseudoknot-stabilized 3WJ topology was retained.[395]

Riboswitches are theorized to be some of the last vestiges of life preceding protein as the principle tool for cellular function, a theory otherwise known as the RNA World.[1,2,396] Consistent with this, all known riboswitch classes bind only small-molecule metabolites—co-enzymes, nucleotides and several derivatives, and three amino acids—and a handful of ions.[181] Riboswitches generally regulate gene expression through ligand-dependent allostery, sequestering their small-molecule ligands in internalized binding pockets. Internalization necessitates that most residues be inwardly oriented, which is obvious upon quick inspection of a riboswitch structure where the vast majority of residues are engaged in intramolecular helical duplex or tertiary pseudoknot contacts

(Figure 6-1). The resulting solvent accessible surface is predominantly helical grooves which are regularly structured, highly negative, and provide minimal sequence-specific chemical identity from which to confer binding specificity. An RNA world would not have contained a protein-binding riboswitch by definition, and if one was evolved during the transition to the current DNA/protein world, we have no evidence for it. Similarly, reengineering contemporary riboswitches *in vitro* to bind protein is difficult to imagine. Therefore, to utilize the concepts of structural recycling in identifying protein-binding aptamers, we must turn to an alternate source of evolutionarily vetted structures.

Compared to riboswitches (Figure 6-2.A), *in vitro* evolved aptamers are generally much smaller, display higher ligand-independent stability, and utilize fewer critical residues ($R_L$) to confer selectivity to their respective ligands (Figure 6-2.B).[11] These likely result from three factors: (i) as opposed to riboswitches, protein-binding aptamers interact through outwardly juxtaposed surface interactions; (ii) folded proteins contain more chemi-structural information than small-molecules, allowing aptamers to outsource necessary chemical information of interaction to their protein ligands; and (iii) *in vitro* selection tends to bias selection of smaller aptamers, which exist with higher copy number in early cycles and retain fidelity through selection cycles, a phenomena referred to as the "tyranny of the small".[14]

An extreme example of the comparatively simple 1° and 2° structural contribution from an aptamer can be observed in the structure of "Thrombin binding aptamer" (TBA) in complex with its ligand, α-thrombin.[96,236] The minimalistic 15nt TBA, composed exclusively of guanine and thymine, utilizes eight residues to generate its core G-

Quadruplex structure and four more residues in simple loops to achieve a low nanomolar-affinity interaction with α-thrombin (Figure 6-3.A). This trend of nucleic acid minimalism can be additionally observed in reciprocal protein evolution. Piccirilli and colleagues have engineered antigen-binding fragments (Fabs) against short RNA sequences with simple hairpin structures, for pull-down assays and chaperoned crystallization.[397] The interaction between the RNA-Fab pair requires minimal complexity from the RNA, just 5 nucleotides in a simple hairpin, as interaction complexity is heavily supplemented by the protein constituent (Figure 6-3.B).[297]

At the time of this writing, the structurally characterized, protein-binding aptome consists of roughly 50 aptamer-protein complexes, which drops to 30 unique structures when redundant crystallizations are removed. Despite the modesty of this data set, very obvious structural trends can be observed (Figure 6-2.C). By far, the most common structural element in protein-binding aptamers is the classical nucleic acid helical duplex, which is found in >80% of determined structures (with >50% composed *entirely* of the motif). While nucleobases in duplex are inward facing and engaged in intramolecular structural contacts, interspersed mismatch non-WC base pairing in aptamers produce unique bulge domains that act as the sites of ligand interaction. These bulges outwardly juxtapose specific residues or form cavities for amino acid side chain engulfment, forming highly specific binding interfaces. Given the number of independent angles of torsional rotation per residue available to the nucleotide backbone (five as opposed to two in peptides), even aptamers with a predominately helical structure can form exquisitely specific binding interfaces against protein ligands, with mean shape complementary

indices that can outperform antibodies.[13] The ubiquity of duplex motifs in crystallized aptamers is initially surprising considering the monumental structural heterogeneity that SELEX is advertised to sample. However, perhaps a bias of duplex structure should be expected, duplex is stable and amenable to short sequences, and protein-binding aptamers are often derived from libraries <40nt long which restricts their capacity for higher order complexity.[92]

Batey and colleagues observed recycled guanine and c-di-GMP riboswitch scaffolds retained ligand-dependent structural stability after being reprogrammed to bind alternative molecules, strongly suggesting that is an inherent property of the 3WJ-pseudoknot scaffold of those riboswitch classes and independent of ligand identity.[395] Based on this, along with knowledge of the general nature of ligand binding in riboswitches and aptamers as discussed above, it appears that $R_L$ in riboswitches exist as a subclass of structural residues, located within a specific domain. This is in contrast to protein-binding aptamers wherein $R_L$ act as fully independent residues that are more dispersed throughout the primary sequence, both distinct from and much fewer than residues that form background structure. The separation of residues descriptively and spatially facilitates the manipulation of each independent of the other. In so doing, vetted aptamers evolved *in vitro* to bind proteins may be similarly—and possibly more—amenable to structural recycling than their riboswitch analogs, promoting the exploitation of their stability and expediting the discovery of novel aptamer constructs.

### *6.2.5 Targeted Epitope Re-selection*

To contextualize these notions of epitope re-selection within the context of protein-binding aptamers concretely, a simplified but explicit experimental approach is presented for the recycling of the aptamer "toggle-25t" (Figure 6-4). Structural studies of toggle-25t complexed with α-thrombin (PDB 3DD2) revealed ligand selectivity and sub-nanomolar affinity to be the result of just six residues (Figure 6-4.A).[398,399] These residues exist in two small clusters, juxtaposed out of the simple helical structure and devoid of intramolecular interaction. It is reasonable to suppose selective randomization of these residues individually will not perturb the overall helical structure of the aptamer but will completely abrogate the aptamer affinity for α-thrombin, similar to the selective 3WJ domain randomization by Batey and colleagues. A very basic experimental flow path is presented, with a hypothetical "6Ntoggle" library, containing six randomized N positions, constant scaffold positions, and primer sequences adapted from the original toggle-25t selection (Figure 6-4.B).[398] Reusing primer sequences further simplifies re-selection, minimizing PCR optimization and reducing the chance of primer sequence structural integration.

Several examples of selective randomization for the opposite purpose—to enhance nucleic acid affinity to its original ligand—have been performed successfully.[400–403] While the goal of these example cases opposes what we are proposing here, the logic of re-selecting a specific target epitope with a constant structure is consistent. In one particular example the affinity of *E. coli* glutamine tRNA for its native protein ligand, glutaminyl-tRNA synthetase (GlnRS), was optimized.[401] The resulting tRNA maintained residue identity at protein contact positions, but altered identity in the background

structure that increased overall 3◦ stability, elevating binding affinity by reducing the entropic cost of interaction. This approach may be a valuable second step to an epitope re-selection attempt like that described with 6Ntoggle. Targeted epitope re-selection against a novel ligand may yield a genuine aptamer, however constraints of the original sequence may reduce its efficacy against a new ligand. This is consistent with previous investigations wherein binding improvements are more dramatic after optimizing intramolecular contacts within the aptamer than by increasing intermolecular contacts with a ligand.[380,404] Re-selecting specific background residues *post*-epitope re-selection could facilitate relaxation of the newly selected $R_L$ residues into their least constrained positions, similar to an *in silico* energy minimization during model construction and MD simulation. This two-step re-selection could be used to great effect while holding true to the logic of recycling a validated aptamer structure.

### *6.2.6 Biological and Experimental Inspiration*

Natural evolution occurs through subtle and spontaneous change of extant structure, facilitating the gradual emergence of divergent function and increased molecular complexity. The conserved riboswitch motif *ykkC* is an interesting example of natural evolution in RNA.[405] A structurally conserved helical domain provides a scaffold upon which smaller variable and modular substrate domains confer ligand specificity.[406,407] Alignment of *ykkC* subclasses reveals how the singular conserved structure is tuned to recognize ligands of drastically different size, shape, charge, and binding location (Figure 6-5). More extreme conservation can be observed between adenine and guanine riboswitches, which use an identical fold to juxtapose a single ligand-

discriminating residue.[408] Targeted epitope re-selection is a semi-rational approach that exploits the successes of previous chemi-structural sampling, in a manner somewhere between natural evolution and *de novo* classical SELEX.

Epitope re-selection is similarly an established and defining feature of the adaptive immune system. Somatic V(D)J recombination in developing lymphocytes produces diverse and unique mature Immunoglobulin (Ig) and T-cell receptor (TCR) proteins with conserved structure but highly heterogeneous complementarity-determining regions (CDR).[140–142] V(D)J recombination produces populations of genetically distinct lymphoid cells which survive and replicate upon antigen binding, analogous to the Darwinian selection utilized in all SELEX derivations. Targeted epitope randomization through V(D)J recombination allows the immune system to exploit a single validated structure—the product of eons of evolutionary selection—for the discrimination of diverse antigens within a single generation.

The logic behind epitope re-selection and structural recycling has been applied to an alternative *in vitro* evolution technique to great effect. Phage display is the fundamental method for the *in vitro* evolution of protein structure, and similar to SELEX, numerous customized protocols have been designed.[9,10] The technique relies on targeted insertion of random sequences within specific epitopes of select viral capsid proteins, without perturbing the virion's capacity to infect a replicating host. Rational randomization of structural epitopes on a constant structural background allow for unguided evolution of large combinatorial peptide libraries.

## 6.3 Caveats & Considerations

There are two principle caveats to epitope reselection in small protein-binding aptamers that must be considered, which residues should be randomized and if there is sufficient heterogeneity in the number of those residue positions to evolve a new interaction. Some structural information is necessary for epitope reselection to be performed as described, and many aptamer-protein complexes--far more than the small collection that has been crystallized--have been probed by low resolution structural assays such as mutagenesis, hydrogen-deuterium exchange, UV-crosslinking, or the like. However, limiting epitope re-selection to aptamers that have been crystallized in complex with their original ligand may be the most fruitful option, as it provides the identity of residues to randomize while ensuring the structure being recycled is at least stable enough to be crystallized (often in the hands of structural biologists who did not perform the original selection and characterization). Using crystallized aptamers exclusively does exclude the vast majority of known aptamer sequences. One must also wonder what structural bias is present in the current canon, and what—if any—structural motifs are currently invisible and thus un-exploitable. Continued structural investigations will expand the aptamer lexicon and amenable structural motifs.

Figure 6-4 presents the hypothetical reprogramming of 6Ntoggle by the randomization of six interfacial residues responsible for direct ligand interaction. However, complete randomization of those six residues provides just $4^6$ (4,096) unique interfacial combinations, a paltry number compared to the $>4^{25}$ combinations often associated with aptamer selection. The question must be asked, does this approach provide enough sequence heterogeneity to successfully reprogram an aptamer structure

to bind an alternate protein by re-selection? Batey and colleagues were allowed the heterogeneity of >20 positions of randomness in the 3WJ ligand-binding domains of their riboswitch re-selection (Figure 6-1).[395] As discussed, protein-binding aptamers are smaller, providing fewer positions to randomize before the overall structure is lost and the very concept of structural recycling breaks down. Assuming no cross reactivity—an admittedly unrealistic assumption—a population of 6Ntoggle will only recognize 4,096 unique protein ligands. What are the odds a given protein-of-interest falls into this exceedingly narrow slice of the universal proteome? In this regard, the conceptual comparison to antibodies appears to similarly fall short, as the same degree of randomization in protein would provide ~15,000-fold more diversity with $20^6$ unique combinations (though while $20^n$ is theoretically possible, immunoglobulin CDRs display significant amino acid bias).[139,409] While the hurdle of homogeneity is problematic for epitope recycling, we can imagine several mitigating solutions.

One way of expanding the heterogeneity of recycled structures in new re-selection is to combine multiple epitope-randomized structures into a single compound library. As of the writing of this manuscript, there are eleven empirically resolved structures of native chemistry RNA aptamers complexed with protein ligands deposited in the PDB. These structures range in $R_L$ count from 3 to 9, combining for a representative combination of 526,144 unique structures if each were to be fully represented. This approach does, however, forego homogeneity of length typically constant during SELEX. The eleven aptamers vary in total length ($R_T$) from 20 to 59nt, with a maximum difference between any two constructs being 19nt and a minimum difference of just 1nt. Visual

inspection of PCR and transcription integrity by gel electrophoresis during a compound library selection will yield complex gel patterns. This may not be a deal breaker, but is a major consideration that could make troubleshooting a fickle selection quite difficult. Additionally, a compound library's heterogeneity is additive not multiplicative, which greatly limits its capacity to resolve the limited heterogeneity raised above. While 526,144 is already a 128-fold increase over 6Ntoggle alone, the heterogeneity is quite low for *in vitro* selection. It is, however, readily expandable with time. Every structurally resolved aptamer of compatible chemistry can be prepared ($R_L$ randomized and identical primer sites extended) and added to the compound library, increasing the library's value with every new PDB-deposited structure.

One of the most attractive traits of aptamers is their ease of chemical modification, which can directly replace specific residues during a structural recycling experiment.[300,338] However their use for expanding heterogeneity in an epitope re-selection experiment is laborious, as the mechanism of nucleobase fidelity in polymerases restricts chemical exploration to a single modification *per* nucleobase *per* library. This results in the same degree of heterogeneity ($4^6$ combinations per library in the case of 6Ntoggle), as exploring multiple modifications of the same nucleotide requires multiple independent selections in parallel each with its own nucleotide pool. Modified nucleotides expand the sampled chemi-structural space, but like a compound library are additive not multiplicative. Additionally, enzymatic incorporation places a modified nucleotide at every template position, including in the background structure which might

be detrimental to the very structural integrity a recycling approach is intending to preserve.

Increasing the heterogeneity while retaining original structure by inserting residues in loops between critical structural elements may be possible in certain cases. For toggle-25t, an insertion of 4 residues in the hairpin cluster may be modest enough to retain helical structure of the aptamer, yet increase the population diversity of reselection from $4^6$ to $4^{10}$. Similarly, and perhaps more effectively, residue positions can be inserted into the inter-plane loops of G-Quadruplex aptamers, drawing inspiration from the biased libraries discussed.[381,383] G-Quadruplexes are common in protein-binding aptamers (Figure 6-2.C), and for those that bind their ligands at inter-plane loops (such as TBA, Figure 6-3.A), residue insertion may be highly effective for epitope re-selection. Figure 6-6 shows hypothetical library designs that utilize various two-residue loop insertions to expand heterogeneity in TBA recycling, which can be increased from $4^4$ to possibly $4^{13}$ combinations while retaining the core G-Quadruplex motif.

Perhaps the most promising approach for maximizing heterogeneity is an expanded genetic code, which is, on the other hand, multiplicative and directly compatible with targeted epitope reselection. “Hachimoji” (Japanese for “eight letters”) is a recently developed eight-base code containing the familiar AT/UGC alongside two additional synthetic purine-pyrimidine pairs (Figure 6-7).[410] The resulting expanded WC interactions of dA:dT, dG:dC, dP:dZ, and dB:dS can be replicated (and transcribed into a Hachimoji RNA) using polymerases evolved *in vitro* that accept the synthetic nucleotides.[97–99,410] In the 6Ntoggle example, one could readily synthesize a DNA

template that restricts dP dZ dB and dS to only the randomized interfacial positions, expanding the heterogeneity of the reselected epitope from $4^6$ to $8^6$ without any background structural alteration. Use of expanded code in aptamer selection will incentivize the genesis of modified variants of these synthetic nucleobase pairs, further expanding the chemi-structural sampling they provide. Expanded genetic codes are of great interest in the fields of abiogenesis and synthetic biology, wherein they are critical pieces of a large and complex puzzle with years still in the works. Minimalist SELEX, on the other hand, requires only the code and compatible polymerases.[411–413] A convenient expanded-code SELEX kit containing all of the appropriate reagents is not available currently, but continued interest should drive these tools to ubiquity soon.

It is unlikely that all aptamers will be equally amenable, or even capable, of ligand re-selection as described. Such aptamer-dependent conditionality that will likely result can be gleaned from an examination of one particular riboswitch-ligand complex.[414,415] The *T. petrophila* fluoride riboswitch is composed of complex high-order tertiary contacts, stabilized by proper fluoride coordination within a water-cation binding pocket (Figure 6-8). Interestingly, this critical water-cation complex is itself coordinated exclusively by backbone phosphates of the riboswitch (Figure 6-8 *inset*). In other words, no single nucleobase interaction is directly involved in fluoride binding, and thus by the very metric presented in this document (Figure 6-2), there are no critical residues in this structure that confer selectivity to fluoride binding. Clearly this cannot be the case. In fact, the opposite—that *every* nucleotide is critical for proper overall backbone juxtaposition—is likely closer to the truth. Does this observation extend to every aptamer-ligand structure?

How influential are background structural residues for the juxtaposition of $R_L$ in conferring ligand selectivity? How much variance in the capacity to be reprogrammed exists between aptamer structures? Ultimately, as exemplified by the fluoride riboswitch, the capacity to reprogram the selectivity of a ligand interface will be highly aptamer-specific, and trial and error with these considerations in mind would be a prerequisite for success.

## *6.4 In silico* Structural Prediction

Structural determination provides invaluable nuanced information and, following the above discussion, opens numerous biochemical doors of optimization and manipulation. While preliminary publications of novel aptamers rarely contain empirical structural determination, many researchers attempt to discern at least some structural information. Considering the small size and minimal tertiary contacts of aptamers, and incompatible modified chemistries often used in their selection, experimental low-resolution structural studies such as 2'-hydroxyl acylation analyzed by primer extension (SHAPE), nuclease-protection, or $Pb^{2+}$-dependent degradation assays are not easily undertaken.[263,416,417] When faced with these limitations, computational modelling and structural prediction are commonly turned to. The current workflow of computational prediction and *in silico* experimentation of aptamer structure begins after a complete SELEX experiment has run its course and a sequence or family of sequences has been identified. From here 2° structural prediction software is employed and tested by various structural assays or mutagenesis. Some available programs produced a complete 3° prediction, from which docking studies, molecular dynamic simulations and aptamer optimization can be performed.[312,313,316,418,419] This general experimental flow has been

followed numerous times. Nucleic acid structural modelling, particularly for RNA, has made great strides in the previous two decades. With this, the number of available software packages and web servers for executing $2^o/3^o$ structural prediction is ever growing. A proper discussion of the nuanced differences of each program and their contributions is well beyond the scope of this manuscript, however several informative reviews are worth further reading.[420,421]

The current mainstay of computational structural prediction of nucleic acid or protein is based on two approaches. The first relies on deterministic biophysical principles governing nucleic acid fold and generates models with the assumption the lowest free energy structure is most likely correct.[422] The prediction can only be as accurate as the biophysical parameters that define its construction. Which is made extremely difficult by the conformational flexibility, relatively weak tertiary interactions, and strong dependency on cation identity and concentration that nucleic acid structures often display. The second predictive approach assembles unknown structures based on relatedness to the primary sequences of previously solved structures.[423] This approach requires a wide diversity of templates with determined structures, and further it assumes closely related sequences will yield closely related structures. This is an assumption not as reliable for nucleic acids as for proteins.[20] The sequence-structure relationship (also referred to as genotype-phenotype relation) of nucleic acids remains very convoluted. Whereas >99.9% of random ssRNA sequence will fold stably, only 2% of random peptide sequence will do so.[20] While the absolute structural diversity of that 2% can adopt is greater than the 99.9% of RNA, the sequence dependency is much higher—in other

words, highly related protein sequences produce fewer, more similar structures than highly related RNA sequences. These results suggest RNA to be a better equipped evolutionary vector than protein for sampling subtler diversity in phenotypic space from substantially larger genotypic space. This supports both the RNA World hypothesis and the merits of aptamer selection by SELEX. Unfortunately, it also makes accurate structural prediction of nucleic acids quite difficult.

Structural biology is on the precipice of a third approach, which in time will surely supersede alternative methods as the dominant tool for all biophysical questions which require structural information. The winners of the critical assessment of protein structural prediction 2020 (CASP-14) were announced in November 2020, with Google's artificial intelligence offshoot DeepMind taking first place with their machine learning algorithm "AlphaFold2".[424] The same group won CASP-13 in 2018 with Alphafold(1), however the success of AlphaFold2 is such a dramatic improvement as to be considered the beginning of a new revolution in structural biochemistry.[424–426] As with any machine learning algorithm, AlphaFold2 uses large sets of experimentally derived data--in this case the entirety of protein structures deposited in the PDB--to generate and train an algorithm for structural prediction. A machine learning algorithm such as AlphaFold2 does not require or directly provide any basic biophysical knowledge as to *why* a given sequence folds the way it does, just that a particular structure is likely based on previous training input of related sequence. While the chemical and biophysical principles governing protein folding remain to be completely elucidated, ubiquitous access to plug-and-play webservers that can quickly provide accurate three dimensional structures from nothing

but sequence is indeed a revolution for all biochemical research, from basic mechanism, to diagnostics, to drug development and therapeutic intervention.

Machine learning for the structural prediction of nucleic acids has similarly been an active area for several years.[427–429] Successful development of machine learning-based approaches for structural prediction bypasses the otherwise necessary elucidation of governing biophysical principles that determine nucleic acid structure. Contemporary prediction is complicated for the reasons discussed. However the principle reason may simply be the dearth of intellectual and financial investment in solving nucleic acid structures, much of which can likely be traced to the 40 year window wherein it was *known* that proteins had the unique structures which nucleic acids simply coded for. The differential investment is obvious from the numbers of respective structures deposited in the PDB (~170,000 for protein and ~3,500 for nucleic acid, at the time of this writing). The efficacy of machine learning algorithms are directly related to the depth and breadth of their training data. For nucleic acids, training data is the limiting factor. While deposition of nucleic acid structures is accelerating, it may be years yet before enough structures of sufficient diversity are solved to allow for a machine learning "AptaFold" or equivalent. This problem is further exaggerated for aptamer structures which contain modified nucleotides, which contain widely divergent and unique structural elements, and for which there are significantly fewer crystal structures.[13,143] Continued interest and investment in aptamer selection and structural characterization will result in increased power, speed, and most importantly, accuracy of computational tools for nucleic acid

structural prediction. This will further expand the availability of aptamer selection and utility, possibly leading to the aptamer holy grail: completely *in silico* SELEX.

## 6.5 The Democratization of Affinity Reagents

The ability to selectively bind a molecular target of interest is a foundational requirement of all biochemical research. To meet this need, researchers have generally turned to specialized biotechnology corporations which have spent decades engineering proprietary hybridomas for monoclonal antibody production and purification, targeting much of the human and common model animal proteomes. Antibodies are among the most valuable and commonly used biological research tools, with a projected market value of $5.3 billion by 2027 according to the market research firm Reports And Data. Four decades of entrenchment as *the* affinity reagent has lulled researchers into generally accepting that antibodies bind their self-reported target ligands with infinite affinity and 100% selectivity, leading to a recent crisis in reproducibility.[430] Despite their perceived ubiquity, many academic laboratories work in model systems wherein the commercial availability of an antibody with even questionable reliability is an unknown luxury, with outsourced custom antibody generation remaining prohibitively expensive. While antibodies will continue to be essential tools, they will remain biologically-derived immune system components being utilized in ways not naturally intended. It is becoming increasingly clear the batch variability and unreliable selectivity are reoccurring problems, and increased stringency or, better yet, effective synthetic alternatives are needed.

Against this backdrop aptamers have made promising but quiet alternatives for three decades. Aptamers are small, stable, inexpensive, reversibly denaturable, and chemically modifiable nucleic acid structures, capable of affinities and selectivities greater than that of antibodies. Perhaps most attractive however are aptamers' comparatively simpler in-house genesis by SELEX, and the ease of their dissemination by sequence alone. Yet despite the alluring theoretical simplicity of classical SELEX, knowledge of aptamer existence, let alone their utilitarian potential as research reagents and clinical tools remains regrettably low.

Some of this can be attributed to what Geoffrey Baird has referred to as "the thrombin problem", referring to the fact that in the preceding 28 years since TBA was first selected, thousands of publications from as many research groups have been released that describe some twist on aptamer application or modification or G-Quadruplex property using TBA as a model structure.[431] TBA was *the* first DNA aptamer as well as the first to target a protein ligand without physiological nucleic acid interaction. The excitement of TBA was accentuated by its simplicity, just 15nt of guanine and thymine folded into a small G-Quadruplex (making it additionally the first G-Quadruplex containing aptamer.[234,236] These firsts were groundbreaking for aptamers and SELEX as concepts and for nucleic acid biochemistry more broadly, and numerous investigations of TBA structure, function, and manipulation were merited. However, the excitement of TBA led reputable journals to publish nearly any manuscript submitted that contained TBA. The nearly guaranteed publication paired with ease and cost of purchasing TBA (one can obtain 100nmol of a 15nt DNA oligonucleotide within a week for ~$10), incentivized

aptamer researchers to invest in “novel” ways of manipulating TBA instead of riskier endeavors such as performing SELEX against any other protein target. Similar trends of obsession can be observed following publications of other now popular models like anti-cocaine, ATP, or PDGF aptamers. Through these targets, aptamers have become victims of their own success.

While aptamers maintain all the same desirable traits that lead to their conception and early excitement, *de novo* aptamer identification is currently monopolized by specialized research groups. New aptamers against novel targets continue to be published regularly. However a second renaissance is needed, heralding new therapeutic aptamers, and more universally applicable aptamer-based research methodologies. Most importantly however, is an increased democratization of novel aptamer identification. Any technique that expands the breadth of interest or contracts both the time and risk of selection is a boon to the scientific community as a whole. With a continued interest in aptamer selection and commitment to expand the diversity of aptamer functional utility, it is reasonable to envision a world of biochemical research where novel in-house aptamer selection is as standard a practice as molecular cloning and protein purification.

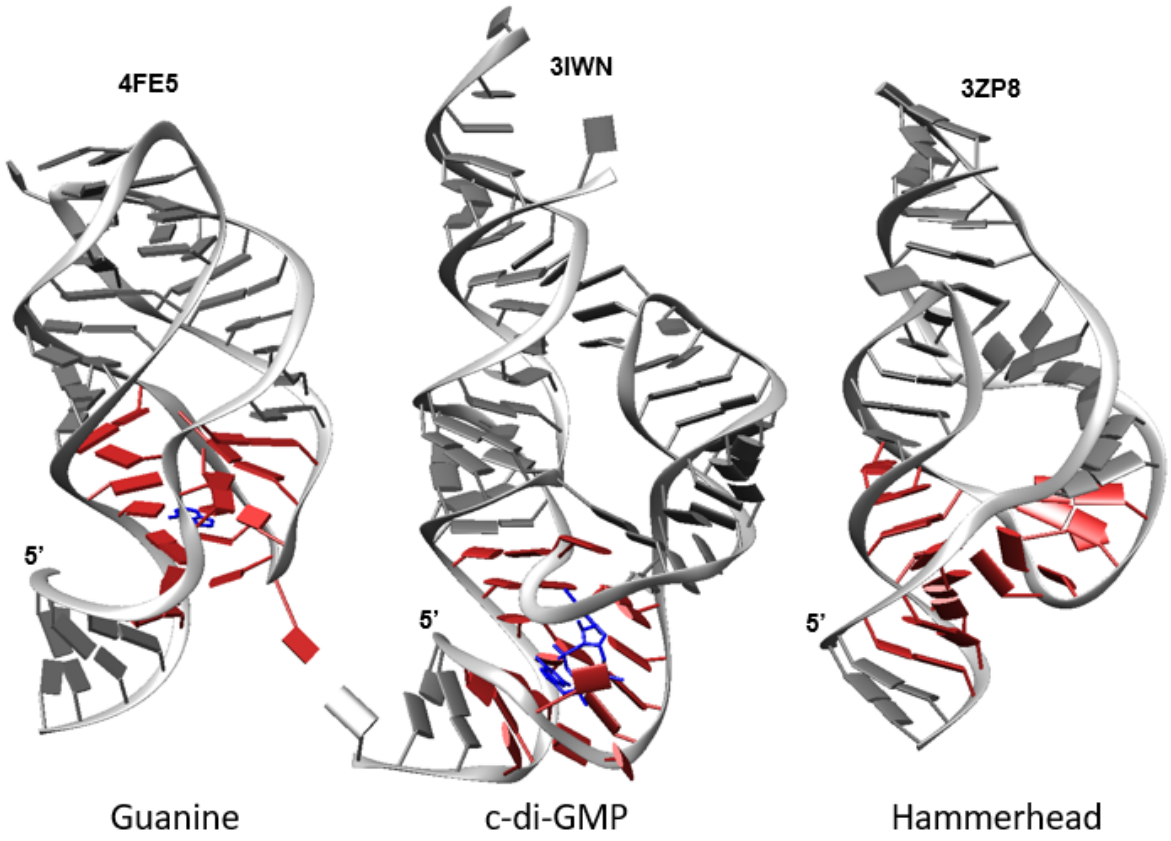



**Figure 6-1 Targeted Epitope Reselection of 3WJ Domains in Natural Ribozymes and Riboswitches.** Presented are the structures of the guanine and c-di-GMP riboswitches and Hammerhead ribozyme reprogrammed to bind 5HTP and L-tryptophan by *Porter et al* 2017. Grey residues were held constant, providing the structural scaffold within which red residues (binding native ligands shown in blue) were randomized to reprogram ligand selectivity by *in vitro* re-selection. PDB IDs presented

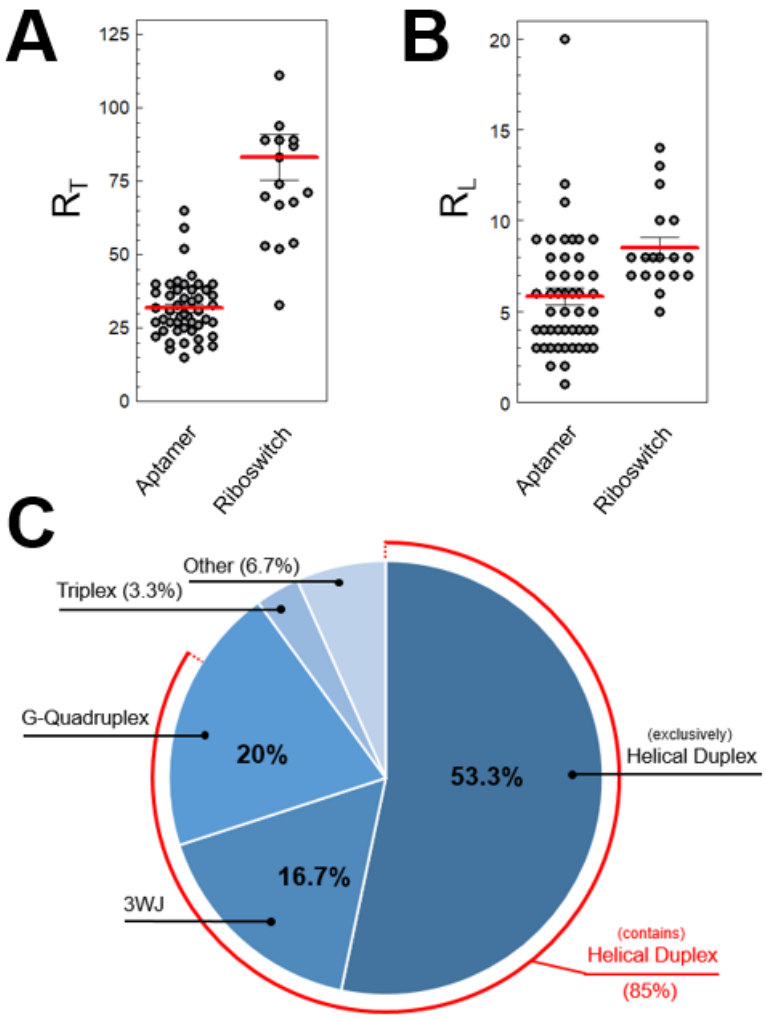



**Figure 6-2 Analysis of Nucleobase-Ligand Interactions of Structurally Determined Riboswitch/Aptamer-Ligand Complexes.** Critical ligand interacting residues ($R_L$) were compared to the total number of residues in the construct ($R_T$), for aptamers and riboswitches (excluding three solved structures for ion-binding riboswitches). (A) Analysis of $R_T$ showing that aptamers are generally smaller, and (B) exploit ~2/3 the number of residues required by riboswitches. (C) Analysis structural elements of PDB deposited protein-binding aptamer structures.

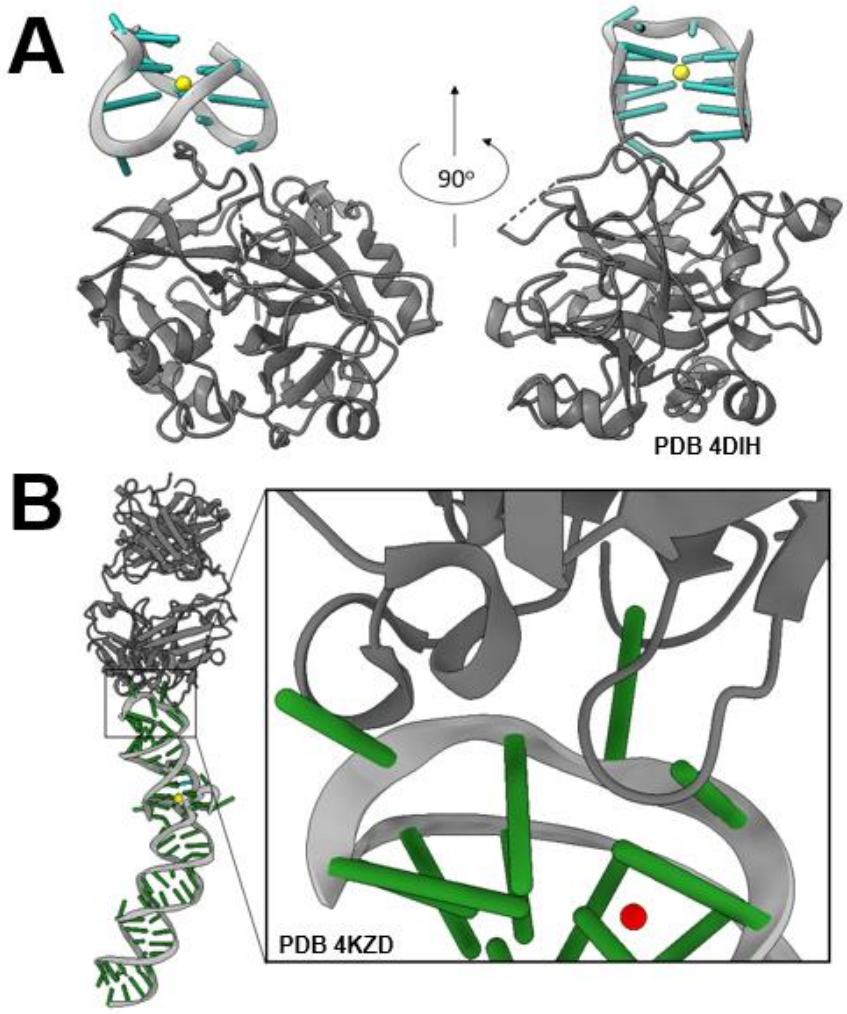



**Figure 6-3 Structural visualization of Minimalist Aptamer-Protein Interactions.** (A) The small 15nt TBA (cyan) contacts a-Thrombin (grey) with 4 thymine residues under the bottom plane of its G-quadruplex motif (PDB 4DIH). (B) "Spinach" aptamer (green) containing a 5nt recognition sequence crystallized in complex with an RNA-binding Fab (grey; PDB 4KZD). The two cases show how minimal aptamer residues are sufficient to confer high affinity interactions with protein ligands.

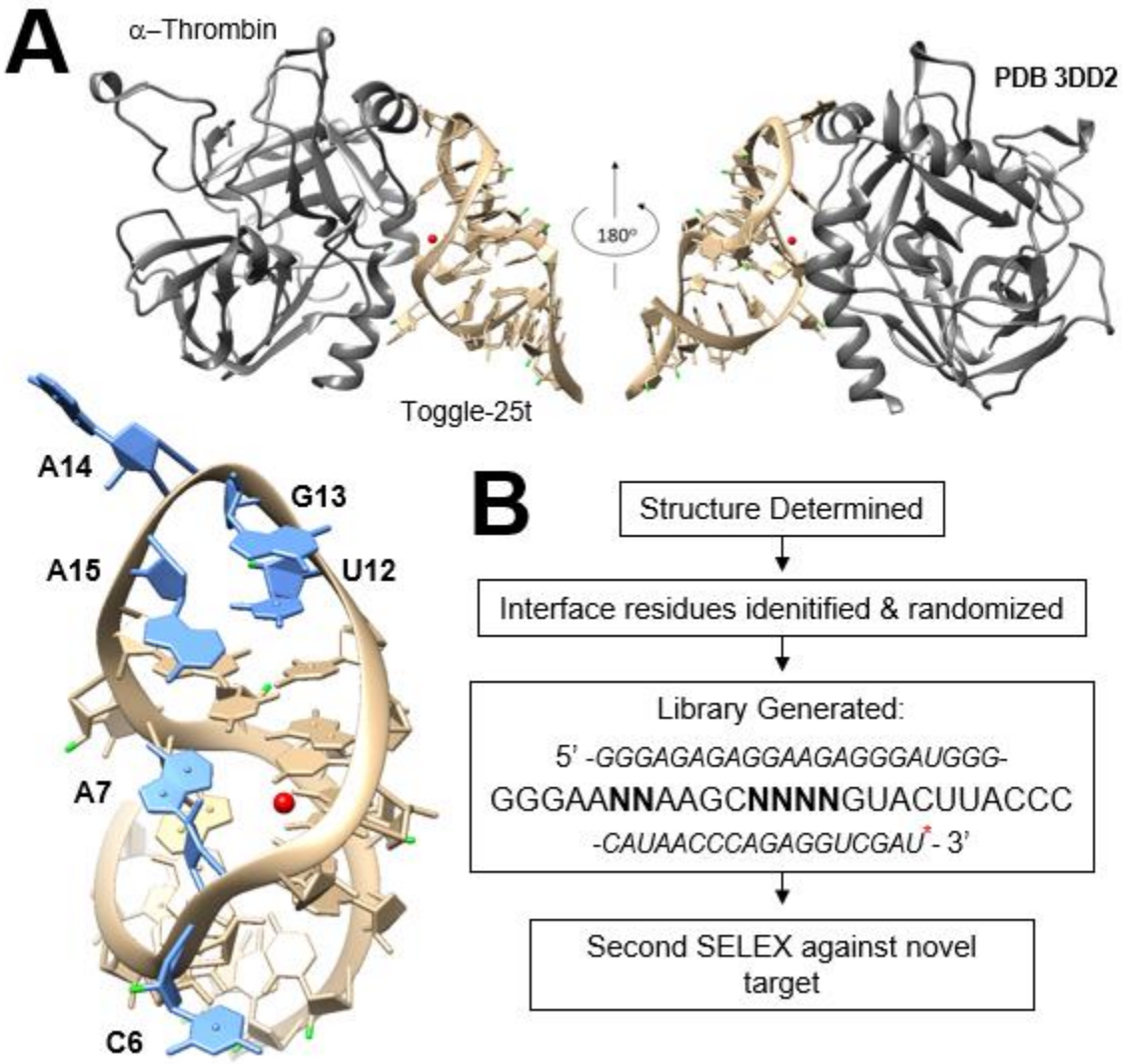



**Figure 6-4 Example Experimental Outline of Targeted Epitope Re-Selection.** (A) Structure of "toggle-25t" aptamer in complex with α-thrombin. The aptamer-ligand interface reveals only six residues in two small clusters contribute selectivity to thrombin, yielding $K_D$ ~ 0.5nM. (B) The $R_L$ nucleobases in toggle-25t are better visualized without thrombin, labelled and colored in blue. (C) A basic experimental flow path is presented. The red asterisk denotes primer docking sequences from the original selection of toggle-25t (White *et al,* 2001). This scheme prevents unnecessary PCR optimization, and mitigates the chance of primer sites becoming structurally integrated into a final aptamer.

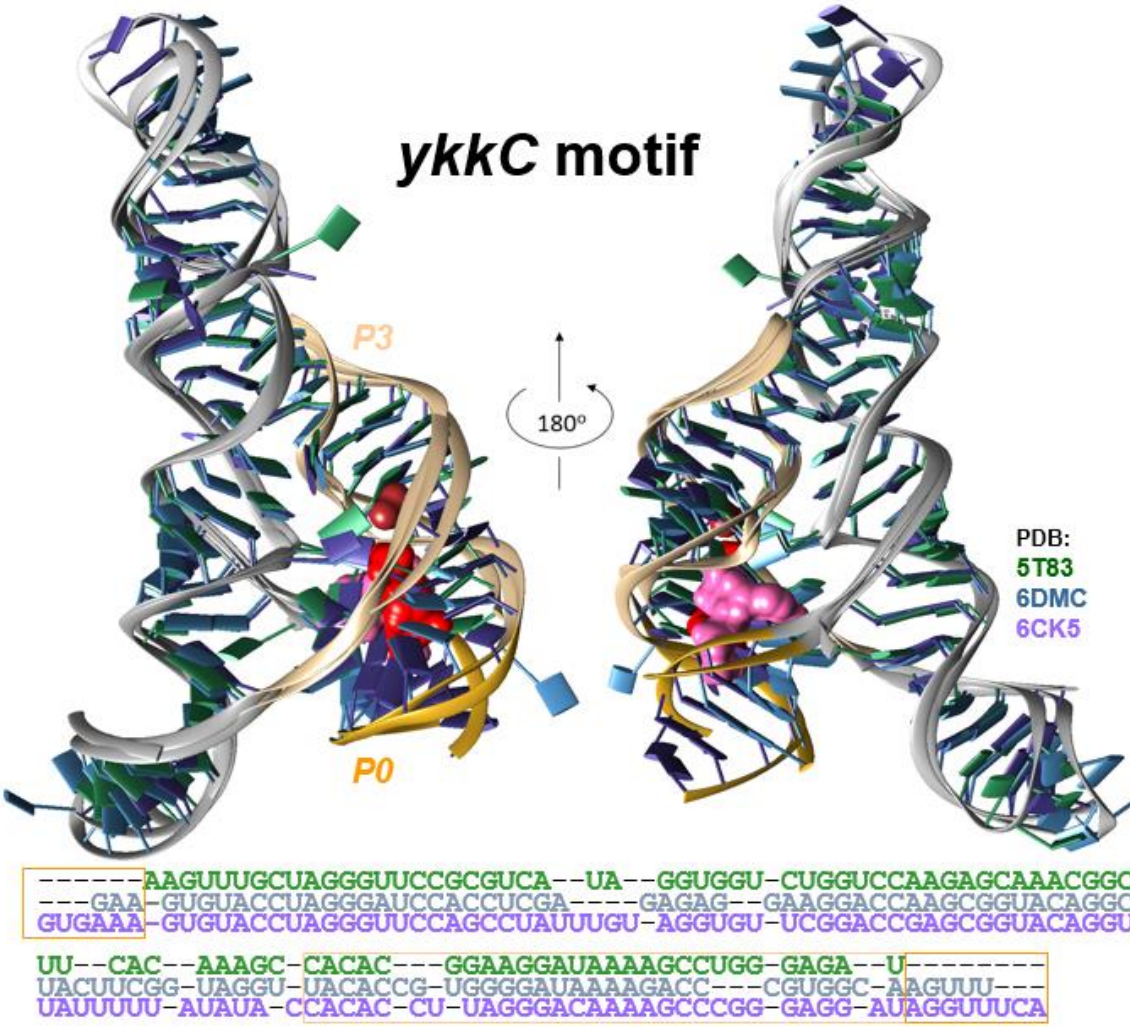



**Figure 6-5 Analysis of *ykkC* Riboswitch Motif Subclasses**. Despite dramatic chemical differences in their respective ligands, high degree of structural conservation is observed between alignments of guanidine (5T83), phosphoribosyl pyrophosphate (6CK5), and guanosine tetraphosphate (6DMC) riboswitches. The modular ligand binding domain, composed of helices P3 and P0, displays the widest sequence divergence. Nomenclature of the P3/P0 domain borrowed from *Knappenberger et al* 2020.

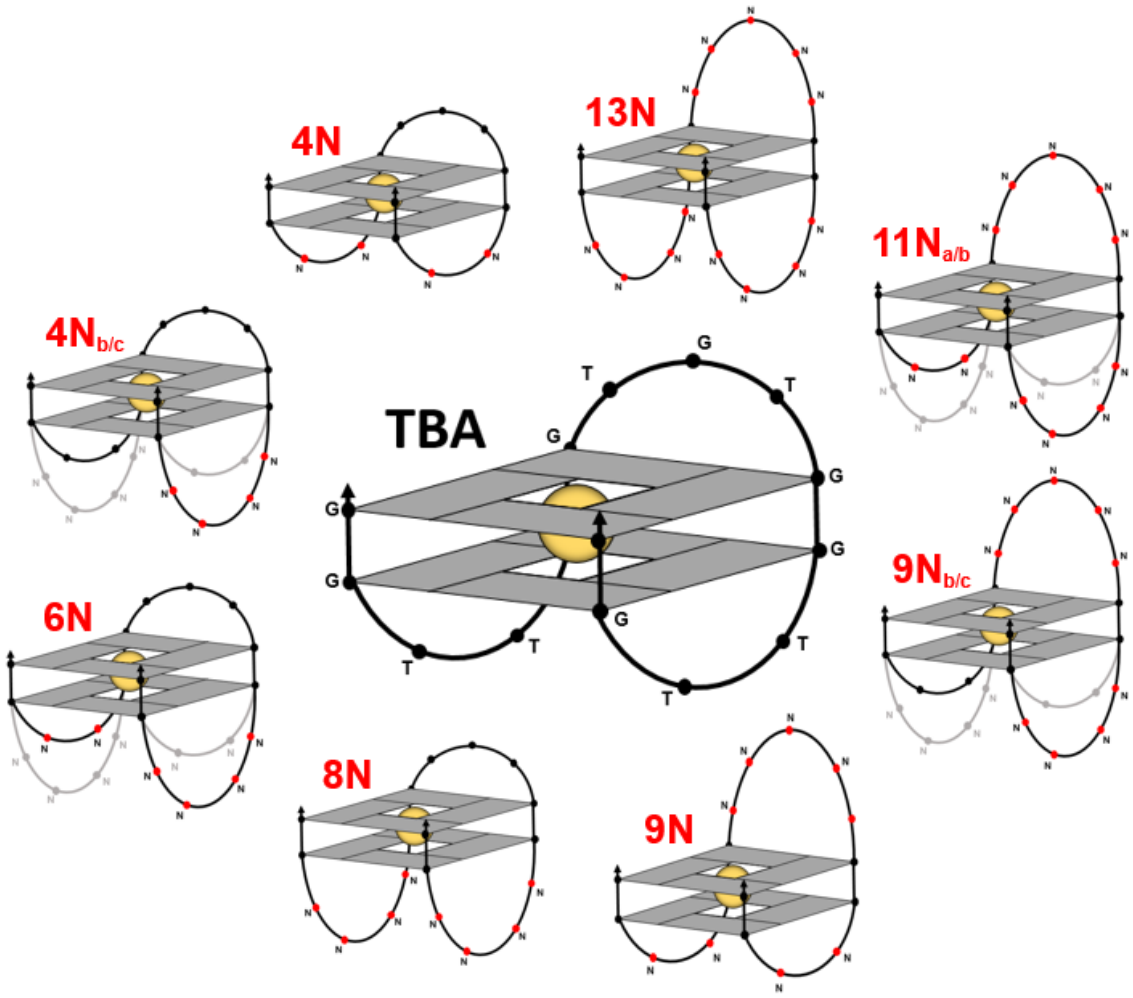



**Figure 6-6 Library Design Considerations for Targeted Epitope Re-Selection of TBA**. TBA aptamer binds a-thrombin via four thymine residues below the G-quadruplex plane (Figure 6-3A). Randomizing these residues exclusively allows for reselection of the basic structure against an alternative protein ligand. To increase the heterogeneity of a selection, and thus expand the proteome of potential novel protein ligands, inter-plane loops can be lengthened as well as randomized. The 11 various possible libraries with 2nt loop insertions are presented. Red dots represent a randomized position, black dots are constant and identical to the original sequence of TBA (which is shown in the center structure). Libraries with randomization in all three loops may require biased nucleotide availability during synthesis which disfavors guanine. This would limit additional G-quadruplex plane formation and deviation from the validated TBA structural motif (Zhu 2012, McManis 2013).

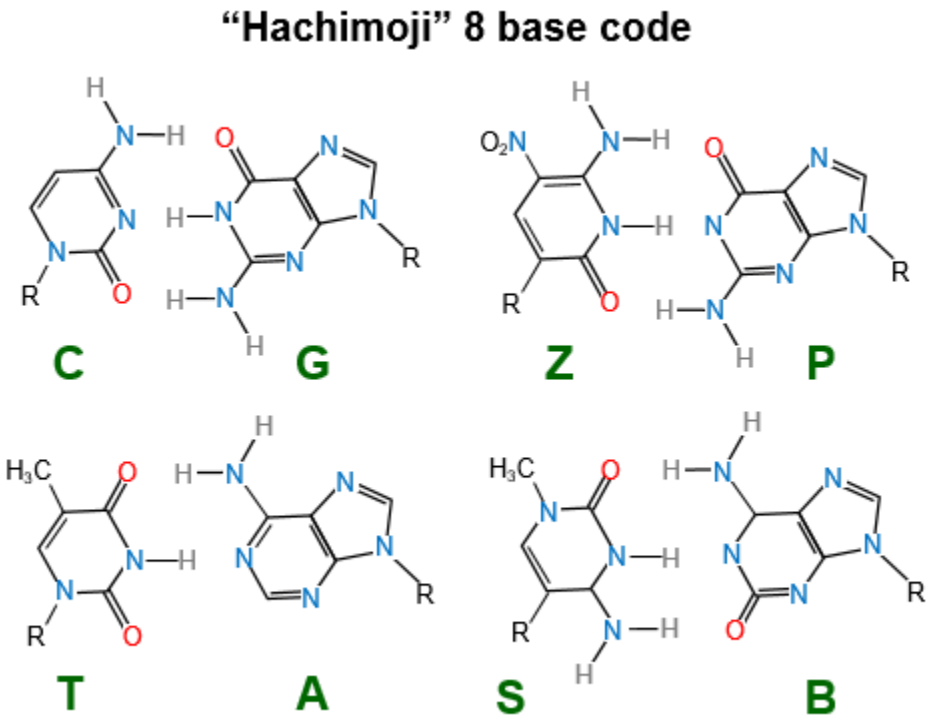



**Figure 6-7** **The Eight Hachimoji DNA and RNA Bases and their Respective Base-Pairing.**

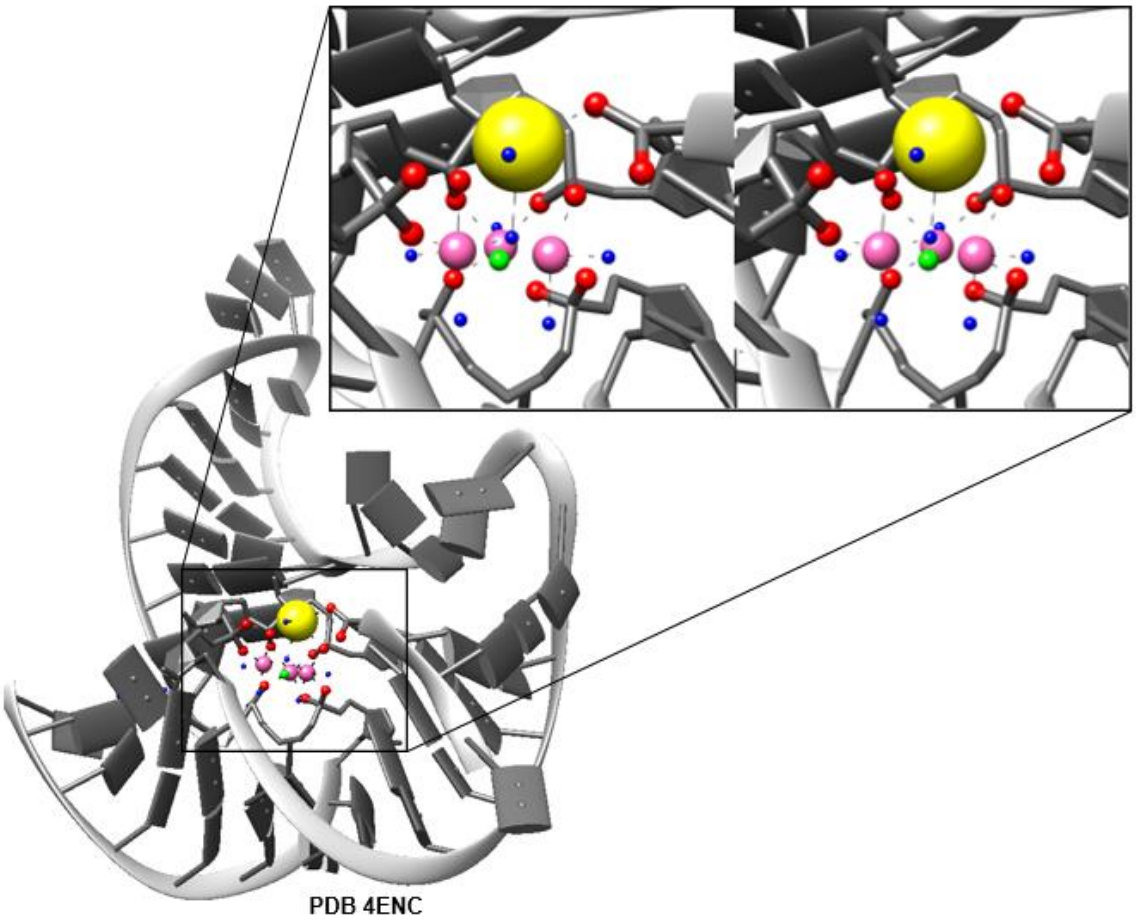



**Figure 6-8 Analysis of the Fluoride Riboswitch (PDB 4ENC).** A minimalist structure of the fluoride riboswitch of *T. petrophila* is shown. Inset stereo view shows the specific binding of a fluoride ion ligand (green) coordinated by three magnesium ions (pink), a single potassium (yellow), and 7 waters (blue). This metal/water coordination complex exclusively interacts with the riboswitch backbone (phosphate oxygen in red). The resulting complex is an interesting situation wherein no single residue confers ligand selectivity, yet many residue's specific juxtaposition is necessary to promote the metal/water complex binding pocket and fluoride selectivity.

# 7

# Formatted Publications

This chapter contains the research and discussion presented in Chapters 2–4, displayed with their final formatting as published in the journal *Chemical Communications.*

Citation numbers this chapter are self-contained and consistent with the original publications. Citation numbers in this chapter *do not* correspond to citation numbers in the preceding Chapters 1-6. References for Chapters 1-6 are contained in Chapter 8 of this dissertation.

ChemComm

COMMUNICATION

View Article Online
View Journal



Cite this: DOI: 10.1039/c9cc01697j

Received 28th February 2019,
Accepted 16th April 2019

DOI: 10.1039/c9cc01697j

rsc.li/chemcomm

# A ribose modification of Spinach aptamer accelerates lead(II) cation association *in vitro*†

Jonathan C. Savage,[a] Pushkar Shinde,[ab] Hans Peter Bächinger,[ac] Monika A. Davare[d] and Ujwal Shinde *[a]

**Spinach aptamer fluorescence requires formation of a tripartite complex composed of folded RNA, a GFP-like fluorophore, and selective cation coordination. 2′F pyrimidine modified Spinach has retained fluorescence, increased chemical stability, and accelerated cation association *via* increased G-quadruplex dynamics, thereby reducing readout time and enhancing Spinach utility for aqueous $Pb^{2+}$ detection.**

The fluorescent RNA aptamer "Spinach" contains two stacked G-quadruplex planes stabilized by cation binding (Fig. 1A).[1,2] Spinach was selected to bind 3,5-difluoro-4-hydroxybenzylidene imidazolinone (DFHBI), a synthetic derivative of green fluorescent protein chromophore, using the aptamer selection technique SELEX: systematic evolution of ligands by exponential enrichment.[3–5] Two structural studies revealed a unique non-parallel G-quadruplex that is coordinated with $K^+$ ions.[1,2] When folded, DFHBI docks on the G-quadruplex motif *via* nine hydrogen bonds that stabilize the *cis*-conformation, facilitating fluorescence.[2] Folded RNA, coordinated cation, and docked DFHBI are all essential to form the tripartite complex necessary for fluorescence.

Interactions between $Pb^{2+}$ and stacked G-quadruplex motifs are well documented.[6–8] One $Pb^{2+}$ atom docks between two successive G-quadruplex planes stabilizing the structure.[6–8] Physiologically, G-quadruplex are observed in coordination with $K^+$, however several other cations stabilize the structures.[6–8] $Pb^{2+}$ also interacts with non G-quadruplex motifs, facilitating catalysis of synthetic nucleic acids.[9,10] Specific and selective interactions between nucleic acids and metal cofactors is critical for DNA and RNA structure and catalysis in natural and synthetic structures. Both are of interest in understanding the evolution of the central dogma of molecular biology and the role of nucleic acids therein.[11,12]

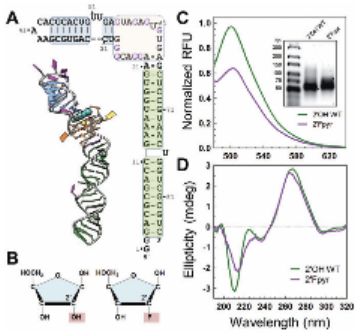



Fig. 1 (A) Spinach construct: cartoon of spinach. Flattened projection is color coordinated with 3D structure (adapted PDB 4kzd[2]): green and blue are stem loops; orange is G-quadruplex; cyan molecule is DFHBI; yellow is $Pb^{2+}$. (B) Ribose fluorination: modified ribose differ by a fluorine at C2. (C) Fluorescence: 2′OH WT and 2′Fpyr facilitate DFHBI Ex/Em at 468/501 nm. Gel inset depicts RNA integrity. (D) Circular dichroism: comparison of CD spectra showing structural similarity.

Spinach was proposed as a tool for fluorescent detection of soluble $Pb^{2+}$.[13] The convenience of nucleic acid sensors is attractive, yet the chemical instability of RNA—especially for environmental sample analyses—complicates its utility. RNA has an additional complication of $Pb^{2+}$-catalyzed degradation.[14] Difficulties of using RNA sequences as experimental tools *in vitro* has led to the use of modified nucleotides.[15,16] We hypothesized modified nucleotides may provide similar avenues for chemical improvement to Spinach for use in *in vitro* applications. Of the various available modifications, ribose C2 fluorination (2′F) is among the most convenient and economical.

Ribose with a fluorine substitution favors a 3′*endo* pucker, allowing for structural rigidity comparable to 2′OH RNA.[17] Yet, a 2′F modification abrogates nucleophilic attack on the phosphodiester

[a] *Department of Biochemistry & Molecular Biology, Oregon Health & Science University, 3181 SW Sam Jackson Park Rd, Portland, OR 97239, USA. E-mail: shindeu@ohsu.edu*
[b] *MSC181105, Emory University Main Campus, 1762 Clifton Rd, Atlanta, GA 30022, USA*
[c] *Research Department, Shriners Hospital, Portland, OR 97239, USA*
[d] *Papé Pediatric Research Institute, Division of Pediatric Hematology/Oncology, Department of Pediatrics, Oregon Health & Sciences University, Portland, OR, USA*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c9cc01697j

backbone at modified positions, increasing chemical stability compared to 2′OH RNA.[17] A mutant T7 RNA polymerase was used to transcribe "2′OH WT" or "2′Fpyr" Spinach, the latter transcribed with equimolar ratios of 2′OH purine and 2′F pyrimidine triphosphates (see ESI†).[17–19] 2′Fpyr contains a 2′F ribose on 34 residues (40%) of the sequence (Fig. 1A and B).

To examine whether 2′Fpyr retains properties similar to 2′OH WT, fluorescent and circular dichroism (CD) spectroscopies were used. Despite having 34 fewer potential donor hydrogen bonds, 2′Fpyr docks DFHBI and facilitates its fluorescence, albeit with loss of ~35% (Fig. 1C), and takes an overall structure similar to 2′OH WT (Fig. 1D). Though reduced sharpness and intensity at 210, 224, and 264 nm in the spectra suggests diminished G-quadruplex compaction.[13] Additionally, 2′Fpyr Spinach displayed similar cation selectivity as 2′OH WT (Fig. S1, ESI†). Together, the data provide functional validation of 2′Fpyr, showing it to be comparable to 2′OH WT.

We next investigated the relative changes in chemical and structural stability between the constructs. The ubiquity and stability of ribonucleases complicate the use of RNA as a probe in physiological or environmental samples. We therefore examined the comparative stabilities of the constructs in the presence of purified RNase A, wherein 2′OH WT is more than 99% degraded in 5 minutes, while 2′Fpyr is less than 50% degraded in 5 hours (Fig. 2A). $Pb^{2+}$ catalyzed RNA hydrolysis is well established, though the mechanism is unclear.[14,20] To test relative resistances to $Pb^{2+}$-catalyzed degradation, we exposed the two constructs to 500 μM $Pb^{2+}$ at 37 °C (Fig. 2B). 2′Fpyr displayed increased resistance to $Pb^{2+}$, degrading ~3 fold slower than 2′OH WT. The relative structural stability of the G-quadruplex was measured by decrease in fluorescence signal with temperature, which is proportional to reduced tripartite complex stability, as loss of RNA structure or binding of either ligand results in complete loss of fluorescence. 2′F chemistry reduces G-quadruplex melting by ~7 °C (Fig. 2C). While previous investigations suggest that electron withdrawing 2′F ribose modifications increase the stability of Watson–Crick bound RNA duplexes,[21,22] the specific constraints of 2′F on G-quadruplex structures have yet to be explicitly explored. Our data show that while 2′F enhances chemical stability of Spinach, the thermal stability of this G-quadruplex orientation is reduced.

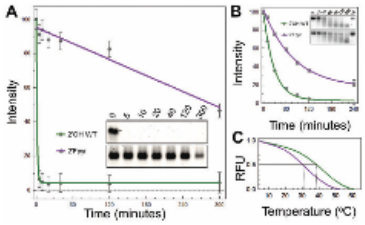



Fig. 2 (A) RNase A-dependent stability: gel inset shows time-dependent aptamer degradation of single experiment (minutes). Graph plotted from densitometry quantification of three replications. (B) $Pb^{2+}$-Dependent stability: gel inset shows time-dependent aptamer degradation of single experiment (minutes). Negative control is 240 minute incubation without $Pb^{2+}$. Graph plotted from densitometry quantification of intact aptamer band in three replications. (C) Thermal stability: melt curves are readouts of tripartite complex stability.

Preliminary fluorescence studies revealed a consistent increase in signal after $Pb^{2+}$ addition, even tens of minutes after initial scans. To explore this, we compared binding kinetics of the tripartite components required for fluorescent readout. First, the kinetics of DFHBI association were investigated by mixing 10× concentrated DFHBI with a solution of Spinach RNA folded with $Pb^{2+}$, with fluorescence immediately recorded (Fig. S2A, ESI†). The resulting kinetic fits are similar, with DFHBI saturating 2′OH WT and 2′Fpyr at similar rates, which—for 2′OH WT—agrees with previously published data.[23] Reciprocally, constructs were folded with DFHBI in an "Apo" state without $Pb^{2+}$, then mixed with 10× concentrated $Pb^{2+}$ and monitored similarly. 2′OH WT required more than 200 minutes to reach fluorescent saturation. In contrast, 2′Fpyr saturated within 40 minutes. $k_{obs}$ values were acquired from mono-exponential fits for a range of $Pb^{2+}$ concentrations (Fig. 3A). Third, RNA was folded without either DFHBI or $Pb^{2+}$, both were then added and fluorescence recorded. This resulted in a slow rate of fluorescent gain similar to the $Pb^{2+}$ addition (data not shown), suggesting cation binding is rate limiting for fluorescence. Linear regression of $k_{obs}$ values from $Pb^{2+}$ concentrations plotted against corresponding sums of $Pb^{2+}$ and RNA concentrations provides $k_{on}$ and $k_{off}$ (slope and $y$-intercept, respectively) (Fig. 3C).[23] We obtained kinetic values for 2′OH WT and 2′Fpyr, $k_{on} = (4.77 \times 10^2\ M^{-1}\ s^{-1})$, $k_{off} = (5.17 \times 10^{-3}\ s^{-1})$, and $k_{on} = (3.02 \times 10^3\ M^{-1}\ s^{-1})$, $k_{off} = (5.93 \times 10^{-2}\ s^{-1})$, respectively. These yield a calculated $K_D$ (termed $K_D^{Calc}$) equaling 10.8 μM, and 19.6 μM, respectively, consistent with previous work.[13]

Prior studies presenting 2′OH WT Spinach fluorescence have consistently reported experimental time-scales shorter than 15 minutes.[13] Our data suggests such an approach would produce artificially suppressed fluorescence readouts. To estimate the magnitude of this suppression we determined a series of theoretical $K_D$ (termed $K_D^{Thr}$) values for various time points within the kinetic fits (Fig. S2B, ESI†) (see Methods, ESI†). The various $K_D^{Thr}$ values were then compared to the $K_D^{Calc}$, with the difference between the two presented as a percent error of $K_D^{Calc}$. In 10 minutes, $K_D^{Thr}$ for 2′Fpyr was found to be within 2% of $K_D^{Calc}$. In the same time, 2′OH WT $K_D^{Thr}$ deviated from $K_D^{Calc}$ by 196%. 2′OH WT did not produce a $K_D^{Thr}$ within 10% error of its $K_D^{Calc}$ for 180 minutes. Accelerated

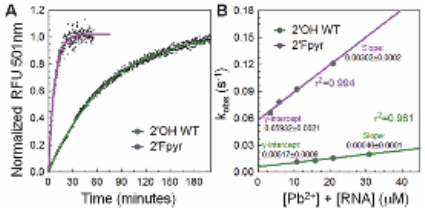



Fig. 3 (A) Binding kinetics of $Pb^{2+}$: fitted fluorescence profile of 20 μM $Pb^{2+}$ mixed with folded 0.5 μM Apo RNA. (B) $K_{obs}$ vs. [$Pb^{2+}$ + RNA]: $k_{obs}$ values of multiple $Pb^{2+}$ concentrations plotted against the sum of $Pb^{2+}$ and RNA in each experiment. Linear regression provides $k_{on}$ values (slope) and $k_{off}$ ($y$-intercept), from which $K_D$ is calculated.

association kinetics is advantageous for the detection of $Pb^{2+}$ in samples with unknown RNase capacity. More importantly, however, is the increased internal consistency of 2′Fpyr fluorescence over a shorter time scale. Despite the potential limits of RNA for $Pb^{2+}$ detection, fluorination produces interesting kinetic results with potential utility as of yet unexplored.

To understand reasons for the differences in kinetics of $Pb^{2+}$ binding between the two constructs, we employed molecular dynamic (MD) simulations. The structure of 2′OH WT Spinach (PDB 4kzd) was used to generate a model of 2′Fpyr (see ESI†).[2] The two constructs were subjected to a 10 nanosecond simulation using AMBER ff14 RNA force field in both Apo and $Pb^{2+}$-coordinated states.[24] When coordinated with $Pb^{2+}$ both 2′OH WT and 2′Fpyr showed similar conformational rigidity over the 10 ns simulation (Fig. 4). However, in the Apo state 2′Fpyr showed drastic local reduction in G-quadruplex compaction compared to its $Pb^{2+}$ bound state and the Apo state of the 2′OH WT (Fig. 4B, C and Fig. S3, ESI†). MD results agree with CD analysis (Fig. 1D), thermal stability (Fig. 2C), and the accelerated $Pb^{2+}$ binding kinetics (Fig. 3B and C). Molecular breathing of the G-quadruplex residues could facilitate faster association by reduced steric occlusion and alterations in localized electrostatics.

The 2′OH groups of RNA contribute to both entropic and enthalpic aspects of RNA stability. Electrostatic repulsion between 2′ and 3′ oxygens shifts the ribose pucker equilibrium in favor of 3′*endo* orientation, providing rigidity to intricate RNA architectures DNA cannot favorably maintain. The 2′OH is additionally capable of donating one and accepting two H-bonds, interacting with nucleobases and backbone phosphates in long range interactions promoting favorable non-Watson–Crick interactions not observed in DNA. Electronegativity of the hydroxyl contributes to the highly negative charge of RNA, enriching and complicating electrostatic interactions involved in folding and ligand binding. 2′F ribose similarly favors 3′*endo* pucker conformation enabling comparable entropic rigidity. Additionally, organic fluorines are weak hydrogen bond acceptors and may replace a 2′hyroxyl in limited capacity.[25–28] Yet 2′fluorine moieties cannot donate a hydrogen bond, and the increased electronegativity of fluorine may alter localized electrostatics in ways incompatible with certain secondary or tertiary interactions. Given these considerations, it is surprising a direct replacement of 2′ chemistry in Spinach retains secondary structure and fluorescence without sequence optimization.

The Spinach tripartite complex presents an interesting paradigm, bringing together concepts of folding, equilibrium dynamics, affinity, and cooperativity. Our experimental data—CD, thermal stability, kinetic association—and MD simulations suggest cation association drives G-quadruplex stability and is the rate limiting step of fluorescent complex formation. MD simulations of 2′OH WT suggest G-quadruplex instability is independent of cation presence, which is unexpected given our observations that Spinach does not fluoresce without a coordinating cation. This raises interesting questions about possible long-range interactions between a coordinated cation and DFHBI through the G-quadruplex plane, independent of G-quadruplex structural stability. While revealing, our data does not completely elucidate the relationship between cation coordination and DFHBI fluorescence and additional investigations are merited.

In this work 2′F is limited to pyrimidines, and not present on the guanine nucleotides in the G-quadruplex, and the retention of ~60% of 2′OH chemistry in modified 2′Fpyr leaves the construct vulnerable to spontaneous alkaline or $Pb^{2+}$ dependent hydrolysis at those positions. The dramatic increase in chemical stability upon just pyrimidine fluorination suggests complete modification of the construct with alternative 2′ chemistry at every position may further increase chemical stability of the structure. However, given differences in hydrogen bond capacity and electrostatic character of a completely modified RNA structure, it seems unlikely that total modification would be possible without sequence optimization.

Our data demonstrate residues proximal to G-quadruplex have significant influence on the properties of the final motif. Hence, it may be possible to rationally tune G-quadruplex properties and ligand binding by altering proximal residues with modified chemistries, or in combination with sequence alterations. This will be increasingly feasible with characterization of novel G-quadruplex conformations in future SELEX experiments. The exploitation of modified nucleotides opens up new avenues for the investigations of G-quadruplex biology, and further broadens the expansive toolbox of nucleic acid biosensors.[25]

To our knowledge, this work represents the first investigation of 2′F nucleotides in a synthetic fluorescent RNA structure. Since the initial publication of Spinach, SELEX experiments have produced additional fluorescent RNAs: optimized Spinach derivatives such Spinach2, iSpinach, and Broccoli, and independent structures Mango, and Corn.[29–34] Interestingly, all of these have converged on derivatives of this stacked G-quadruplex motif, highlighting its importance as an ideal orientation for fluorophore docking. Our studies suggest that G-quadruplex folding, fluorophore docking,

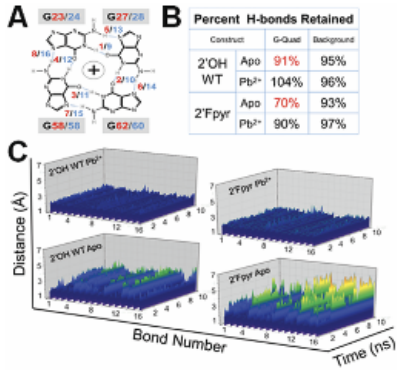



Fig. 4 (A) G-quadruplex H-bonds: schematic of 16 guanine H-bonds of the two G-quadruplex planes. Red and blue correspond to bottom and top planes, respectively. (B) Percent retention: H-bonds retained after 10 ns simulation as percentage of starting count. H-bonds extended beyond 2.5 Å in the simulation are considered broken. Background structure H-bonds calculated as the difference between total and G-quadruplex. (C) 3D visualization of distances: left plots are 2′OH WT, right are 2′Fpyr. Top depict RNA simulations in coordination with $Pb^{2+}$ and bottom are Apo.

altered kinetics, and affinity for cation association with modified nucleotides is a promising approach for functional modification of alternative RNA aptamers with stacked G-quadruplexes. Optimized Spinach derivatives share conserved sequences in the G-quadruplex core. Despite this, the constructs have varying ligand affinities, thermal stabilities, and DFHBI fluorescent maxima. The comparative properties of these related structures raise interesting questions about the long-range influence of Watson–Crick duplexes, G-quadruplex proximal residues, and altered electrostatics on the tripartite complex formation and fluorescence. This growing class of synthetic fluorescent aptamers remains underutilized as *in vitro* tools, and although it remains to be seen whether fluorination of these perturbs their folding, fluorescence, or dimerization, our findings with Spinach necessitate exploration. Increasing the chemical stability of these RNAs, while retaining their fluorescence for use *in vitro* through modified nucleotide chemistry, is a valuable expansion of their experimental utility.

While 2′Fpyr displays reduced thermal stability compared to 2′OH WT (Fig. 2C), fluorescent-based kinetic studies revealed $Pb^{2+}$ associates with 2′Fpyr seven fold faster than 2′OH WT (Fig. 3). Increased chemical stability allows for sensor utility over longer time-scales despite the observation that faster association kinetics enables $Pb^{2+}$ visualization on a shorter time frame, before RNA hydrolysis can become an issue. Accelerated kinetics of 2′Fpyr result in internally consistent fluorescent signal in 10 minutes, whereas 2′OH WT requires greater than 90 minutes before the kinetics are within 50% error of saturation conditions (Fig. S2B, ESI†). MD simulations suggest 2′F reduces compaction of the G-quadruplex in the absence of $Pb^{2+}$ (Fig. 4C and Fig. S3, ESI†). Increased movement of 2′Fpyr G-quadruplex Apo state explain experimentally observed changes in CD and fluorescence spectra, reduction in thermal stability, and accelerated binding kinetics with $Pb^{2+}$.

Accurate and rapid detection of heavy metal toxins is a public health necessity, with soluble lead of particular importance.[35–39] Classical approaches for lead quantification—Atomic Absorbance Spectrophotometry (AAS) and Inductively Coupled Plasma Mass Spectrometry (ICP-MS)—are associated with prohibitive costs.[40,41] Fluorescence is a comparatively simpler approach.[42,43] We present a improvement of Spinach for $Pb^{2+}$ detection. 2′Fpyr displays a dramatic increase in chemical stability in the presence of ribonuclease and $Pb^{2+}$, both of which pose a significant liability for RNA utility. We propose pyrimidine fluorination is a chemical improvement of Spinach for *in vitro* $Pb^{2+}$ detection, with 2′Fpyr Spinach displaying increased chemical stability and internally consistent fluorescent signal over shorter time scales.

We would like to thank Dr David Farrens for the use of his spectroscopic equipment, and Drs Phil Yates and Omar Davulcu for their thoughtful discussion. This research was supported by the Medical Research Foundation of Oregon (GBIMO0231A; U. S.), the St. Baldrick's Foundation (GPEDI1059A; M. A. D.) and the NIH (5T32AI007472-24; J. C. S).

## Conflicts of interest

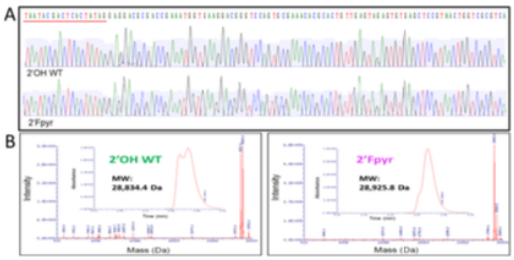
The authors of this manuscript have no conflicts to declare.

## References

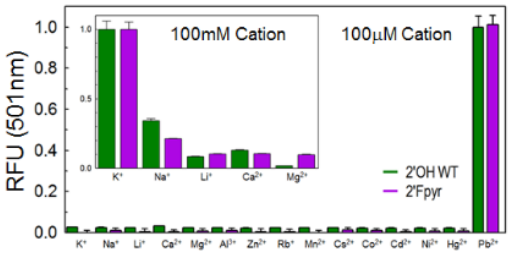
1 K. D. Warner, M. C. Chen, W. Song, R. L. Strack, A. Thorn, S. R. Jaffrey and A. R. Ferré-D'Amaré, *Nat. Struct. Mol. Biol.*, 2014, **21**, 658–663.
2 H. Huang, N. B. Suslov, N.-S. Li, S. A. Shelke, M. E. Evans, Y. Koldobskaya, P. A. Rice and J. A. Piccirilli, *Nat. Chem. Biol.*, 2014, **10**, 686–691.
3 J. S. Paige, K. Y. Wu and S. R. Jaffrey, *Science*, 2011, **333**, 642–646.
4 A. D. Ellington and J. W. Szostak, *Nature*, 1990, **346**, 818–822.
5 C. Tuerk and L. Gold, *Science*, 1990, **249**, 505–510.
6 W. Zhou, R. Saran and J. Liu, *Chem. Rev.*, 2017, **117**, 8272–8325.
7 D. Bhattacharyya, G. Mirihana Arachchilage and S. Basu, *Front. Chem.*, 2016, **4**, 38.
8 T. Li, E. Wang and S. Dong, *J. Am. Chem. Soc.*, 2009, **131**, 15082–15083.
9 R. R. Breaker and G. F. Joyce, *Chem. Biol.*, 1994, **1**, 223–229.
10 J. Li and Y. Lu, *J. Am. Chem. Soc.*, 2000, **122**, 10466–10467.
11 R. R. Breaker and G. F. Joyce, *Chem. Biol.*, 2014, **21**, 1059–1065.
12 G. F. Joyce, *Cold Spring Harbor Symp. Quant. Biol.*, 2009, **74**, 17–23.
13 S. DasGupta, S. A. Shelke, N. Li and J. A. Piccirilli, *Chem. Commun.*, 2015, **51**, 9034–9037.
14 W. R. Farkas, *Biochim. Biophys. Acta*, 1968, **155**, 401–409; S. Ni, H. Yao, L. Wang, J. Lu, F. Jiang, A. Lu and G. Zhang, *Int. J. Mol. Sci.*, 2017, **18**, 1683–1704.
15 S. Ni, H. Yao, L. Wang, J. Lu, F. Jiang, A. Lu and G. Zhang, *Int. J. Mol. Sci.*, 2017, **18**, 1683–1704.
16 F. Lipi, S. Chen, M. Chakravarthy, S. Rakesh and R. N. Veedu, *RNA Biol.*, 2016, **13**, 1232–1245.
17 A. D. Keefe, S. Pai and A. Ellington, *Nat. Rev. Drug Discovery*, 2010, **9**, 537–550.
18 R. Sousa and R. Padilla, *EMBO J.*, 1995, **14**, 4609–4621.
19 R. Padilla and R. Sousa, *Nucleic Acids Res.*, 2002, **30**, e138.
20 M. Lindell, M. Brännvall, E. G. H. Wagner and L. A. Kirsebom, *RNA*, 2005, **11**, 1348–1354.
21 P. S. Pallan, E. M. Greene, P. A. Jicman, R. K. Pandey, M. Manoharan, E. Rozners and M. Egli, *Nucleic Acids Res.*, 2011, **39**, 3482–3495.
22 A. Patra, M. Paolillo, K. Charisse, M. Manoharan, E. Rozners and M. Egli, *Angew. Chem., Int. Ed.*, 2012, **51**, 11863–11866.
23 K. Y. Han, B. J. Leslie, J. Fei, J. Zhang and T. Ha, *J. Am. Chem. Soc.*, 2013, **135**, 19033–19038.
24 D. Tan, S. Piana, R. M. Dirks and D. E. Shaw, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, E1346–1355.
25 Y. Du and S. Dong, *Anal. Chem.*, 2016, **89**, 189–215.
26 J. Sponer, G. Bussi, M. Krepl, P. Banas, S. Bottaro, R. A. Cunha, A. Gil-Ley, G. Pinamonti, S. Poblete, P. Jurecka, N. G. Walter and M. Otyepka, *Chem. Rev.*, 2018, **118**, 4177–4338.
27 C. Dalvit, C. Invernizzi and A. Vulpetti, *Chem. – Eur. J.*, 2014, **20**, 11058–11068.
28 C. Dalvit and A. Vulpetti, *Chem. – Eur. J.*, 2016, **22**, 7592–7601.
29 R. L. Strack, M. D. Disney and S. R. Jaffrey, *Nat. Methods*, 2013, **12**, 1219–1224.
30 A. Autour, E. Westhof and M. Ryckelynck, *Nucleic Acids Res.*, 2016, **6**, 2491–2500.
31 P. Fernandez-Millan, A. Autour, E. Ennifar, E. Westhof and M. Ryckelynck, *RNA*, 2017, **23**, 1788–1795.
32 G. S. Filonov, J. D. Moon, N. Svensen and S. R. Jaffrey, *J. Am. Chem. Soc.*, 2014, **136**, 16299–16308.
33 E. V. Dolgosheina, S. C. Y. Jeng, S. S. S. Panchapakesan, R. Cojocaru, P. S. K. Chen, P. D. Wilson, N. Hawkins, P. A. Wiggins and P. J. Unrau, *ACS Chem. Biol.*, 2014, **9**, 2412–2420.
34 K. D. Warner, L. Sjekloća, W. Song, G. S. Filonov, S. R. Jaffrey and A. R. Ferré-D'Amaré, *Nat. Chem. Biol.*, 2017, **13**, 1195–1201.
35 P. B. Tchounwou, C. G. Yedjou, A. K. Patlolla and D. J. Sutton, *EXS*, 2012, **101**, 133–164.
36 A. L. Wani, A. Ara and J. A. Usmani, *Interdiscip. Toxicol.*, 2015, **8**, 55–64.
37 H. Needleman, *Annu. Rev. Med.*, 2004, **55**, 209–222.
38 D. Bellinger, A. Leviton, C. Waternaux, H. Needleman and M. Rabinowitz, *N. Engl. J. Med.*, 1987, **316**, 1037–1043.
39 H. L. Needleman, A. Schell, D. Bellinger, A. Leviton and E. N. Allred, *N. Engl. J. Med.*, 1990, **322**, 83–88.
40 S. Selander and K. Cramér, *Br. J. Ind. Med.*, 1968, **25**, 209–213.
41 R. S. Hour, *Acc. Chem. Res.*, 1994, **27**, 333–339.
42 C. T. Chen and W. P. Huang, *J. Am. Chem. Soc.*, 2002, **124**, 6246–6247.
43 Q. He, E. W. Miller, A. P. Wong and C. J. Chang, *J. Am. Chem. Soc.*, 2006, **128**, 9316–9317.

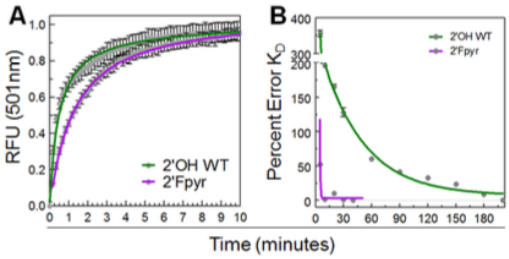
# Supplemental Material



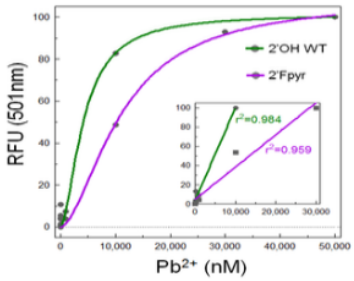

**Supplemental S1: A) Sequence Identity:** cDNA of Spinach constructs were sequenced by Sanger sequencing validating integrity of the transcripts generated by T7 RNA polymerase variants and modified nucleotides (T7 promoter sequenced is underlined in red). **B) LC-MS:** Spinach Constructs analysed by Liquid Chromatography—Mass Spectrometry to validate sample identity. 2'OH WT sample composed of two species differing by a single 5' terminal guanosine, likely due to T7 RNA polymerase slipping. Mass shift between 2'OH WT and 2'Fpyr constructs corresponds to fluorination of pyrimidine residues.



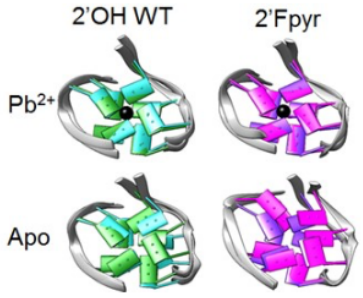

**Supplemental S2: Cation Selectivity:** Spinach constructs were folded in the presence of DFHBI and various cations. Both constructs show similar fluorescent yield with each cation, showing the strongest preference for $Pb^{2+}$ in the micromolar range. *Inset*: physiologically relevant cations were also investigated in 100mM concentrations. Both constructs show similar preferences. All cation solutions prepared from $Cl^-$ salts for consistent background.




**Supplemental S3: A) Binding Kinetics of DFHBI:** Fluorescent gain after 20μM DFHBI mixing with folded 0.5μM RNA pre-incubated with $Pb^{2+}$. Colored lines are monoexponential fits. **B) Percent error of fitted $K_D$ vs. Time.** $K_D$ values determined ($K_D^{Thr}$) at specified time windows 5, 10, 20, 30, 40 (2'Fpyr and 2'OH WT), 60, 90 120, 150, 180, 200 min (2'OH WT only). Deviation from calculated $K_D$ ($K_D^{Calc}$; calculated at saturation for 2'Fpyr and 2'OH WT, 40 and 200min, respectively) at each time point is presented as percent error of $K_D^{Calc}$, and is calculated as ($|K_D^{Thr} - K_D^{Calc}|/K_D^{Calc}$)*100. Lines are monoexponential fits.




**Supplemental S4: Detection Limit:** $Pb^{2+}$ was detectable by fluorescence down to 10nM by both Spinach constructs. *Inset*: Linear regression of $Pb^{2+}$ detection by both constructs for $Pb^{2+}$ concentrations between 50nM and 30μM.




**Supplemental S5: Structural Comparison of G-quadruplex at 0 and 10ns:** Images correspond to 2'OH WT and 2'Fpyr both in coordination with $Pb^{2+}$ and in Apo states. Green and purple correspond to 0ns, and cyan and magenta correspond to 10ns.

## Methods

***In vitro* Transcription:** RNA transcripts were prepared using linear dsDNA PCR templates, using the forward strand sequence (T7 promoter is underlined): 5'—GCGCGCGAATTCTAATACGACTCACTATAGGAGGACGCGACCGAAATGGTGAAGGACGGGTCCAGTGCGAAACACGCACTGTTGAGTAGAGTGTGAGCTCCGTAACTGGTCGCGTC—3'. Reactions assembled at room temperature, in the following order, to final concentrations of: water, 1X T7 RNAP transcription buffer (NEB), 2.5-5% DMSO, 24mM $MgCl_2$, 4mM each rNTP (2'OH purines and 2'OH pyrmidines from NEB, 2'F pyrimidines from Trilink Biotechnologies), 15-25ng/μl dsDNA, 5mM DTT, 1U/μl murine RNase inhibitor (NEB), 2.5mU/μl yeast inorganic pyrophosphatase

(NEB), 5U/μl WT T7 RNA polymerase (NEB) or 1.25U/μl mutant T7 R&D polymerase (Lucigen, Inc). RNA was transcribed at 42°C for ≥4 hours, and observed on a 4% agarose TBE gel. Target transcript band was purified by gel excision and electro-elution in dialysis tubing, and dialyzed against ≥5,000X volume of 10mM Tris (pH 8.0) for ≥2 hours at room temp. RNA was cleaned with three successive mixes with *acidified* phenol:chloroform:isoamyl alcohol (Ambion), then precipitated in 0.3M sodium acetate (pH 5.2) with either ≥50% isopropanol or ≥75% ethanol at -20°C for ≥2 hours. RNA was pelleted by centrifugation (16.1k *x g* for ≥30 minutes at 4°C), washed with ice-cold 70% ethanol and centrifuged again (16.1k *x g* for ≥10 minutes at 4°C). RNA pellets dried at 37°C and suspended in RNase free 20mM Tris (pH 8.0), quantified by UV-absorbance at 260nm or fluorescence (Qubit, Invitrogen), and stored at -80°C. Based on current reagent prices and our average recovery efficiency at a low scale of transcription, we calculate 2'Fpyr modified RNA to be ~2-4 fold more expensive per unit mass than 2'OH unmodified. Unless otherwise stated, all RNA experimentation was performed in a standard base buffer solution of 20mM Tris (pH 8.0), 5mM $MgCl_2$, 20μM DFHBI. Prior to experimental use of frozen samples, RNA was folded following a basic protocol: Dilution in buffer then successive thermocycler incubations at 85°C (20sec), 50°C (1min), addition of DFHBI, 37°C (1min), 25°C (≥30min).

***RT.PCR:*** RNA constructs were reverse transcribed and amplified using NEB OneTaq OneStep RT.PCR kit following manufacturer protocol. PCR cycling conditions as follows: 48°C (30min), 95°C (1min), 25cyles at 95°C (20sec) 63°C (30sec) 68°C (30sec), 68°C (10min), 4°C. RT.PCR amplicons were gel purified as described above and ethanol precipitated, before ligation into a TOPO.pca2.1 TA sequencing vector (Invitrogen) and used for transformation of DH5α E. coli. Sanger sequencing was performed commercially by Eurofins Scientific.

**Liquid Chromatography—Mass Spectrometry:** LC-MS experimentation and analysis performed commercially by Novatia, LCC, using a Dionex UltiMate HPLC and Thermo Finnigan LTQ mass spectrometer.

**Circular Dichroism Spectroscopy:** Measurements were performed on an Aviv 202 CD Spectrometer at 20°C. RNA samples were folded and prepared similarly to fluorescence based experiments. 5μM RNA samples were scanned in 1mm quartz CD cuvette from 320 to 185nm by 0.25nm steps, with a 2 second integrated read per step. Data presented as average of three individual scans. Control scans of DFHBI in buffer subtracted from RNA spectra. Values plotted in GraphPad Prism 5.

**Chemical Stability Measurements:** *RNase-dependent:* 0.5μM RNA was folded as described in standard spinach buffer without lead. RNA was incubated with purified RNase A (Qiagen) at a final concentration of 1.25mU/μl at 25°C. 10μl samples were removed from pooled incubation at designated time points quenched with 10mM EDTA, and 2X DNA loading dye (Thermo Scientific). Samples were run on a 3% agarose TBE gel, stained with Sybr Gold nucleic acid dye. Band intensities were quantified by densitometry with Image Lab software and plotted in GraphPad Prism 5. *$Pb^{2+}$-dependent:* 0.5μM RNA was folded as described in standard Spinach buffer solutions, and incubated with or without 500μM $Pb^{2+}$ for 4 hours at 37°C. 10μl samples were removed from pooled incubation at designated time points, quenched with 10mM EDTA, and frozen. Samples were run on a 3% agarose TBE gel, stained with Sybr Gold nucleic acid dye. Band intensities were quantified by densitometry with Image Lab software, and plotted in Graph Pad 5.

**Thermal Stability Measurements:** Thermal melting was observed using an ABI StepOne Real Time PCR Thermocycler. 0.5μM RNA was folded as described in 20mM Tris (pH 8.0), 5mM $MgCl_2$, 50μM $PbCl_2$, 20μM DFHBI, 1% DMSO, and loaded in 20μl volumes in triplicate wells in an ABI brand 96 well QPCR plate. Fluorescence was measured in the Sybr Green Ex/Em channel from 4°C to 70°C at a 2% gradient. Control scans of DFHBI in buffer subtracted from RNA spectra. Fluorescent values averaged and plotted in GraphPad Prism 5.

**Fluorescence Measurements:** All fluorescent measurements were performed on a Photon Technologies International QM-1 steady state fluorescent spectrophotometer. Samples were prepared at stated concentrations and excited at 468nm, at 20°C held constant by a circulating water bath, in a 2mm quartz fluorescence cuvette. Kinetic experimentation was performed by rapid, in-cuvette mixing of folded Apo-state RNA in buffer and 10% vol. of 10X $PbCl_2$ in water. Emission at 501nm was read every 29.5 seconds with a 0.5 sec shutter exposure. $K_D^{Calc}$ values were determined using $k_{on}$ and $k_{off}$ determined under steady state fluorescence conditions, at 200min and 40min for 2'OH WT and 2'Fpyr, respectively. $K_D^{Thr}$ values were determined by excluding data beyond stated time points, and refitting to generate new theoretical $k_{obs}$ at those time points. Control scans of DFHBI in buffer subtracted from RNA spectra. Fluorescence data was plotted and fit in GraphPad Prism 5.

**Modelling and Molecular Dynamic Simulations:** Spinach aptamer (PDB 4kzd) without co-crystalized Fab was used as the template for modelling and MD simulations. 2'Fpyr construct was modelled by direct chemical replacement of the C2 hydroxyl group with fluorine on every pyrimidine ribose using YASARA *Structure* modelling software. A revised AMBER ff14 RNA force field was used for energy minimizations, using an explicit solvent TIP3P water model. Nucleotide bases were frozen and the backbone was subject to an energy minimization, after which the backbone was frozen and the bases were minimized. The structure was then subject to a final, full-structure minimization with no constraints. Potassium was replaced with a lead atom, in both constructs, and full-structure energy minimizations were performed with no constraints. MD simulations were run for 10ns, with snapshots taken every 100ps, using the YASARA md_run macro with explicit solvent conditions at 298K with a physiological intracellular NaCl concentration. Hydrogen bonding atoms present in the bases of the starting aptamer structure were identified and the distances between them were catalogued for every snapshot. A distance over 2.5Å was defined as a broken hydrogen bond. YASARA hydrogen bond interpolation was then used to count the number of hydrogen bonds present within the aptamer over time, within the G-Quadruplex, duplex sections, and the transition between the duplex and the G-Quadruplex core. These values were obtained for every snapshot. Hydrogen bond values were plotted in Sigma Plot, and structural images presented with Chimera.

ChemComm

ROYAL SOCIETY OF CHEMISTRY

COMMUNICATION

View Article Online
View Journal



Cite this: DOI: 10.1039/c9cc09375c

Received 3rd December 2019,
Accepted 29th January 2020

DOI: 10.1039/c9cc09375c

rsc.li/chemcomm

# Subtle sequence variations alter tripartite complex kinetics and G-quadruplex dynamics in RNA aptamer Broccoli†

Jonathan C. Savage,[a] Monika A. Davare[b] and Ujwal Shinde *[a]

**Though extensively utilized, the fluorescent RNA aptamer Broccoli is poorly characterized with an unknown structure. Spectroscopic and kinetic investigations of tripartite complex formation reveal surprising differences between Broccoli and Spinach aptamers despite extreme sequence conservation. Our studies highlight how subtle sequence variations impart functional consequences of G-quadruplex–cation interactions in RNA.**

The RNA aptamer "Spinach", selected *in vitro* by SELEX,[1–4] docks to DFHBI (3,5-difluoro-4-hydroxybenzylidene imidazolinone) and induces DFHBI's fluorescence through a unique, cation-coordinated, two-planed G-quadruplex motif, that coalesces into a tripartite complex of RNA + DFHBI + Potassium, or "RDK".[5–7] Spinach has been minimized in successive selections producing a family of smaller and generally more stable iterations including Spinach 2, Spinach mini, baby Spinach, and iSpinach.[5,6,8–11] "Broccoli", one of the smallest DHFBI-binding RNAs, has been extensively utilized *in vitro* and *in vivo* but its structure remains unknown.[11–13] Here we investigate the relationship between kinetics and structure of RDK formation in Broccoli and Spinach. The two aptamers display notable differences in kinetics of complex formation, yet they share extensive sequence identity and secondary structure, as observed by sequence alignment, and Circular Dichroism (CD) and fluorescence spectroscopies. To understand our experimental observations we built a homology model of Broccoli, analysed its structure, and subjected it to molecular dynamic (MD) simulations. Our data suggest that both aptamers adopt the same unique G-Quad motif, but with altered stability imparted by subtle sequence differences, resulting in differential RDK formation and altered complex stability.

Broccoli and Spinach RNAs display comparable fluorescence under high concentrations of KCl and DFHBI (100 mM and 20 μM, respectively), with both showing maximal excitation and emission at 468 and 501 nm, respectively (Fig. 1A). The fluorescent complex of Broccoli has an increased thermal stability over Spinach, as assayed by fluorescence (Fig. 1B), which may be explained by the higher affinity for $K^+$ Broccoli displays over Spinach (Fig. 1C). Interestingly, the kinetics of complex formation upon rapid mixing of concentrated KCl and DHFBI with RNA are notably different (Fig. 1D). Fig. 1E shows the linear regression of $K_{obs}$ values for RNA→RDK determined from the

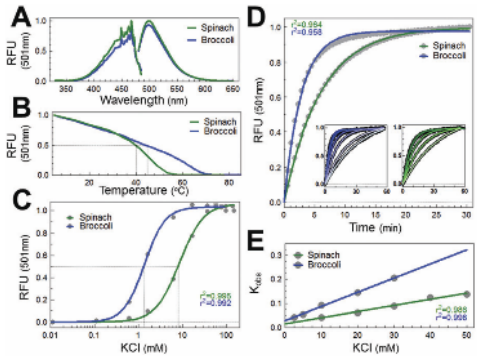



Fig. 1 (A) Fluorescence spectrum: fluorescence of DFHBI induced by Spinach and Broccoli show similar properties including Ex/Em wavelengths (468/501 nm). Constructs display >5% difference in intensity. (B) Thermal melting: RDK fluorescence as a function of temperature. Broccoli displays increased thermal stability ($T_m$ = 45 °C) compared to Spinach ($T_m$ = 40 °C). (C) $K^+$ titration: RNA incubated with 20 μM DFHBI and varying [KCl], and allowed to reach equilibrium. Fluorescent Em at 501 nm was recorded, and values fit for approximate affinity for $K^+$. Broccoli shows a ~10-fold increase in affinity for $K^+$. (D) Comparative kinetics: Apo folded Spinach or Broccoli was rapidly mixed with concentrated KCl and DFHBI in a cuvette and scanned for fluorescence upon RDK formation. Colored lines are fit. Inset: Identical approach performed with multiple concentrations of KCl, with fixed concentrations of RNA and DFHBI. All scans normalized and fit. (E) Comparative kinetics continued: $K_{obs}$ values for each KCl concentration from fits of (B) inset were plotted against respective KCl concentrations. Slope of the linear regression equals the $K_{on}$ for complex formation. Spinach and Broccoli $K_{on}$ = 0.0025 $M^{-1}$ $min^{-1}$ and 0.0059 $M^{-1}$ $min^{-1}$, respectively.

[a] Department of Chemical Physiology & Biochemistry, Oregon Health & Science University, 3181 SW Sam Jackson Park Rd, Portland, OR 97239, USA. E-mail: shindeu@ohsu.edu

[b] Papé Pediatric Research Institute, Division of Pediatric Hematology/Oncology, Department of Pediatrics, Oregon Health & Sciences University, Portland, OR, USA

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c9cc09375c

fit of five KCl concentrations with constant DFHBI (Fig. 1D; inset). The slope of the regression provides $K^+$ association forming RDK, which is observed to be roughly 2.5-fold faster for Broccoli than Spinach (Fig. 1E).

Researchers have treated Spinach and Broccoli as unrelated constructs, as per their nomenclature and 2D structural representations. However, the fact they induce fluorescence of the same ligand with similar photophysical properties and cation dependency necessitates deeper analyses of their sequence and structure. Sequence alignments show the G-Quad motif within iSpinach and Spinach share >92% sequence identity (Fig. 2A and Fig. S1, ESI†). Only one residue differs in the G-Quad motif ($U_{17}$ and $A_{27}$ in iSpinach and Spinach, respectively). The superimposed X-ray structures of the two have a room mean square deviation (RMSD) of 0.63 Å within the G-Quad motif (Fig. 2B and Fig. S2, S3, ESI†). Broccoli also shares 91% and 95% sequence identity within the Spinach and iSpinach G-Quad motifs, respectively, suggesting Broccoli also adopts the same unique structure (Fig. 2A and Fig. S2, ESI†). To experimentally compare Broccoli and Spinach, CD spectroscopy was utilized.[14–17] The CD spectra reveal the two RNAs adopt nearly identical secondary structures, validating the similarity suggested by the sequence alignment (Fig. 2A, C and Fig. S1, ESI†). Furthermore, the ensemble of peaks at 295, 260, and 240 nm provide experimental evidence of a G-Quad motif, independent of Spinach as a crystallized reference.[5,6,8,17] Together, the alignment and spectroscopic analysis strongly suggest Broccoli contains a G-Quad, in an orientation likely identical to Spinach.

Given the similarities based on sequence alignments and spectroscopic analyses, it is surprising that the kinetics observed in RDK complex formation between Broccoli and Spinach are different (Fig. 1D and E). In pursuit of a mechanistic explanation of these differences a homology model of Broccoli was generated (see Methods). The template, iSpinach, was chosen for four reasons: (1) Broccoli is more similar to iSpinach than Spinach in sequence length and identity (Fig. 2A and Fig. S1, ESI†); (2) a crystal structure of iSpinach exists, which (3) is nearly identical to Spinach (Fig. S3, ESI†); yet, (4) the reported thermal stability and $K^+$ affnity of iSpinach are more similar to values of Broccoli than Spinach (Fig. 1).[9]

Considering their sequence conservation, it is unsurprising the Broccoli model is superimposable upon the crystal structures of Spinach and iSpinach with RMSDs of their aligned G-Quad motifs <1.0 Å (Fig. 2B and Fig. S2, S3, ESI†). The G-Quad motif of Spinach and Broccoli is flanked by two A-form Watson–Crick (WC) duplex helices, one ending in both 5′ and 3′ termini and the other ending in a hairpin loop (Fig. 3A). Structural analysis of Broccoli and Spinach reveal three base pair interactions located in the termini duplex that may, in part, explain the observed difference in thermal

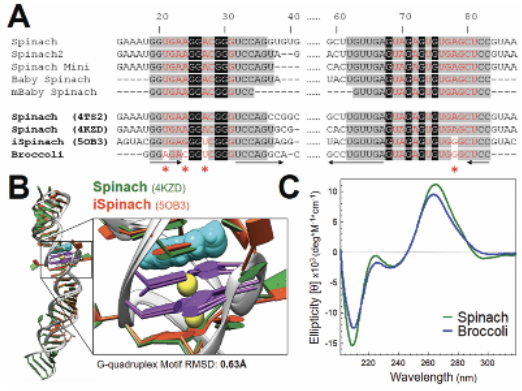



Fig. 2 (A) Sequence alignment: comparison of Spinach family members and Broccoli. Color coding: red letters/grey background are G-Quad motif and transition residues; white letters/black background are G-Quad guanine residues; black lettering/grey background are conserved residues; arrows indicated complementary regions. Bottom four bold sequences correspond to crystalized Spinach sequences and Broccoli. Red asterisks denote positions of difference in the G-Quad and transition motifs between Spinach and Broccoli. (B) Spinach *vs.* iSpinach: the crystal structures of Spinach (4KZD) and iSpinach (5OB3) are aligned. RMSD of the alignment at the G-Quad motif show considerable structural similarity, suggesting iSpinach to be a satisfactory template for homology modelling Broccoli. G-Quad guanine bases are purple, DFHBIs are cyan, and $K^+$ ions are yellow. (C) Circular dichroism: Spinach and Broccoli show similar spectral signature by CD, suggesting Broccoli contains the G-Quad motif shared by all other Spinach derivatives. Signal intensities at wavelengths 295, 260, and 240 nm suggest a G-Quad independent of Spinach as reference structure.

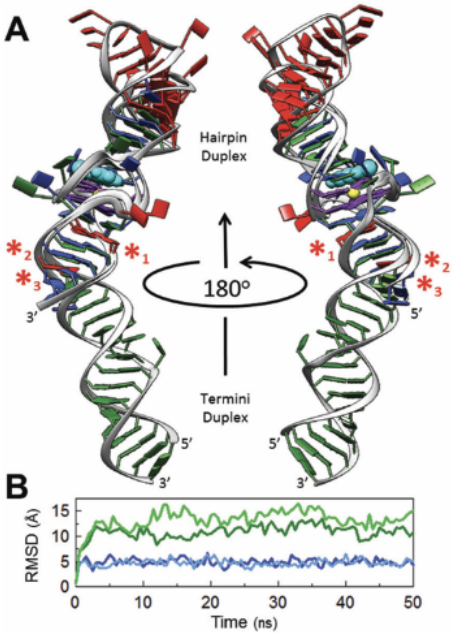



Fig. 3 (A) Broccoli and Spinach comparison: Broccoli and Spinach (4KZD) structures are aligned, in blue and green, respectively. G-Quad guanines in purple, DFHBI in cyan, $K^+$ ions in yellow, non-conserved residues in red. *1 is $C_7$–$G_{54}$ in Broccoli and $A_{21}$–$A_{64}$ in Spinach; *2 is $A_4$–$U_{57}$ in Broccoli; *3 is $U_{18}$–$U_{67}$ in Spinach. (C) Time course of MD simulation: Spinach and Broccoli subjected to 50 ns simulations. Broccoli displays increased conformational rigidity over the simulation window. Light and dark hues are Apo and potassium bound, respectively.

stability and cation affinity presented in Fig. 1. The 3 differing residues around the G-Quad in Broccoli result in two canonical WC duplex base pairs with increased hydrogen bonding compared with Spinach ($A_4$, $C_7$, $G_{54}$ and $U_{18}$, $A_{21}$, $A_{64}$, in Broccoli and Spinach, respectively; Fig. 3A). This suggests that, while substantially shorter than Spinach, the termini duplex of Broccoli is likely more stable. The effect of alternate base interactions may be the most dramatic at the first base-pairing downstream of the G-Quad motif in the termini duplex ($A_{21}$–$A_{64}$ in Spinach and $C_7$–$G_{54}$ in Broccoli). $C_7$–$G_{54}$ in Broccoli compresses the transition region between G-Quad and termini duplex in Broccoli, zipping up the duplex immediately adjacent to the G-Quad motif with a CG WC pair, increasing its stability with a minimum predicted enthalpic contribution of $-5.8$ kcal mol$^{-1}$ at that position.[18]

To generate hypotheses and rationalize observed differences in kinetics of $K^+$ association between Broccoli and Spinach, 50 ns MD simulations were performed in YASARA using a modified Amber14ff RNA force field[19] (Fig. 3B). Broadly, MD simulations with KCl show Broccoli is more stable than Spinach (Fig. 3B). Yet both display an increase in the conformational flexibility of their G-Quad motifs in an Apo state without $K^+$, Broccoli more so than Spinach (Fig. 4A). Interestingly, movement is predominately just two H-Bonds from one guanine residue, Broccoli G49 and Spinach G59. The dynamics of $G_{49}$ in Broccoli are drastically increased over the movement of equivalent $G_{59}$ in Spinach, with Broccoli $G_{49}$ flipping out of the G-Quad plane at $\sim$5 ns and remaining open for the duration of the Apo simulation (Fig. 4A and B). $G_{49}$ flipping is particularly striking when the absolute distance of its movement is compared to the remaining G-Quad H-bonds of both structures (Fig. 4B and C). In the Apo state, the total G-Quad motif of Spinach and Broccoli is rigid (Fig. 4C; top). The 7 G-Quad guanines without $G_{49}$ of Broccoli display subtly increased conformational rigidity (Fig. 4C; middle), and the majority of the G-Quad movement observed for Broccoli in Fig. 4B may be attributed to $G_{49}$ (Fig. 4C; bottom). The remaining residues, particularly $G_{17}$ and $G_{46}$, provide the site for DFHBI docking, and the subtly increased stability of these residues in Broccoli (Fig. 4C; middle) may explain its increased Apo fluorescence (Fig. 4D). Together it appears $G_{49}$ acts as a $K^+$ "gatekeeper", resulting in faster $K^+$ association and increased DFHBI fluorescence in the Apo state (Fig. 1D and 4D).

Despite the differences between Broccoli and Spinach derivatives suggested by their nomenclature and presented 2D structures, the two aptamers share much in common. They both promote the fluorescence of the same ligand with similar photophysics in a $K^+$-dependent manner, in the same concentration range. Sequence analysis reveals the two share >90% sequence in the G-Quad motif suggesting structural similarity, which is experimentally corroborated by CD. Despite these similarities the two display differences in stability and kinetics of RDK formation.

Our model allows for investigations into the relationship between sequence, structure, and function in Broccoli and Spinach, and provides the means for *in silico* rationale of experimental observations (Fig. 1). Two phenomena have been uncovered that may work synergistically. The first is canonical WC base pairing in the termini duplex stem of Broccoli that cinches the strands closer to the G-Quad motif through G–C hydrogen bonding, which may restrict

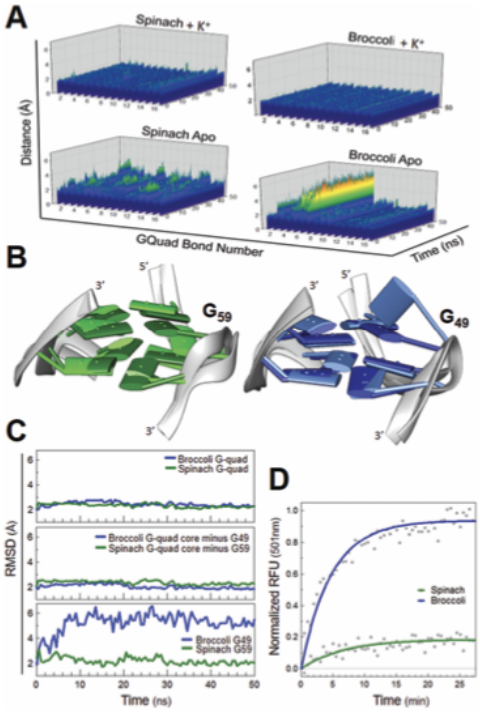



Fig. 4 (A) Specific G-Quad H-bond dynamics in simulations: the 16 H-bonds from the 8 guanine residues of the stacked G-Quad guanines are plotted as a function of stretch distance and time, in Spinach and Broccoli with and without $K^+$ ("Apo"). Broccoli shows more base movement than Spinach in the Apo state, however the majority of the movement is restricted to $G_{49}$. The $G_{49}$ "gatekeeper" can explain the accelerated $K^+$ association kinetics observed in Fig. 1 by providing an opening for $K^+$ docking. (B) Visualization of Apo state MD simulations: the G-Quad guanine residues of Spinach (left; green) and Broccoli (right; blue) are presented. $T = 0$ and $T = 50$ ns of each are aligned (dark shades are $T_0$, light shades are $T_{50}$). The most drastic movement for both structures corresponds to the same guanine ($G_{59}$ of Spinach and $G_{49}$ of Broccoli), with $G_{49}$ in Broccoli the most pronounced. (C) Cumulative RMSD of G-Quad guanines: stability of the G-Quad guanines without potassium was looked at *via* RMSD over the MD simulation. Both RNAs display rigid G-Quad structures throughout the simulation (top). Broccoli appears to have a subtle increase in G-Quad stability when $G_{49}$ is removed, *i.e.* remaining 7 guanines (middle), whereas $G_{49}$ alone in Broccoli accounts for most or all of the net G-Quad dynamics. This contrasts with $G_{59}$ in Spinach which is static (bottom). (D) DFHBI association with Apo RNA: DFHBI was spiked into Apo RNA and fluorescence was observed. Although low fluorescence was observed for both, Broccoli displayed 5-fold higher signal than Spinach.

the conformational flexibility of the motif, increasing the RDK melting temperature as observed (Fig. 1B). The second is the movement of $K^+$ gatekeeper $G_{49}$ to an open state seen in the Apo MD simulations, which could explain the accelerated $K^+$ association kinetics (Fig. 1D and E) by increased accessibility to the inside of the G-Quad motif. This hypothesis is consistent with our recent work on lead association with 2′F pyrimidine modified Spinach aptamer.[20] Subtle perturbations in local environment—either due to nucleotide modification or minor

change of primary sequence—produce conserved G-Quad motifs with altered dynamics and kinetic stabilities resulting in changes in cation association rates. These data posit further questions on the capacity for rationally tuning of cation selectivity into related RNA structures for modified function which is yet to be explored. Broadly, our work is further complemented by another recent publication,[13] wherein a single pyrimidine residue in Broccoli tunes the fluorescent emission of fluorophore DFHO by up to 20 nm. Together these demonstrate how subtle differences in chemistry and sequence can produce changes in biophysical properties of RNA, independent of structure. These observations are likely to have broad implications for a wade array of G-quad structures.

G-Quads are critical in gene regulation, metabolite sensing, and disease ranging from viral infection to Alzheimer's to cancer, and are attractive modalities in synthetic biology and the development of chemical tools.[21–29] Despite their importance in RNA biology, the majority of *in vitro* characterizations of G-Quads rely on telomeric DNA; G-Quads are inherently weak in their absorbance and fluorescence,[30–32] but telomere sequences are almost exclusively composed of G-Quad stacks, magnifying their signal strength with increasing sequence length. While this makes telomeres a spectroscopically convenient G-Quad model, telomeric DNA does not provide comparable insight into equivalent RNA sequence.[33] Spinach family RNAs are predominately used for subcellular visualization of larger complexes,[11,34] yet their utility as models for G-Quad and RNA-cation investigations are heretofore unrealized. This RNA family shares a unique combination of model features: (i) structure-dependent fluorescence by a small ligand with high affinity and specificity; (ii) two distinct cation cofactors of differing valences and hydration states ($K^+$ and $Mg^{2+}$); and (iii) three distinct structural motifs (Watson–Crick duplex, hairpin, and G-Quad). These features provide a system for RNA investigations from various functional and chemical perspectives within a constant background for direct experimental comparisons.

Our results provide mechanistic explanations for the increased thermal stability and accelerated rate of tripartite complex formation of Broccoli compared to Spinach. These investigations reveal how subtle sequence differences result in canonical WC base pairings, which stabilize the G-Quad motif and increase dynamics of gate-keeper $G_{49}$, promoting cation association. This work utilizes DFHBI fluorescence to investigate RDK tripartite complex formation; however, these investigations posit new questions about component interactions which require alternative, fluorescence-independent biophysical approaches for direct detection of cation hydration, order of association, cooperativity, or folding on sub-second timescales. Such investigations will provide proof-of-principle for biochemical investigations of RNA using this model system, a more complete picture of Broccoli-cation complex biophysics, and strengthen our understanding of the relationship between sequence, structure, and function in RNA.

The Authors thank Dr Dave Farrens for the use of his spectroscopic equipment. This work was supported by the NIH (J. C. S) and the Oregon MRF (U. S.).

## Conflicts of interest

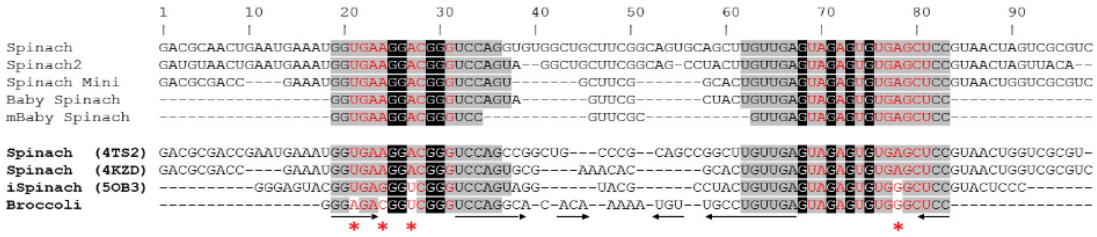
The authors of this manuscript have no conflicts to declare.

## References

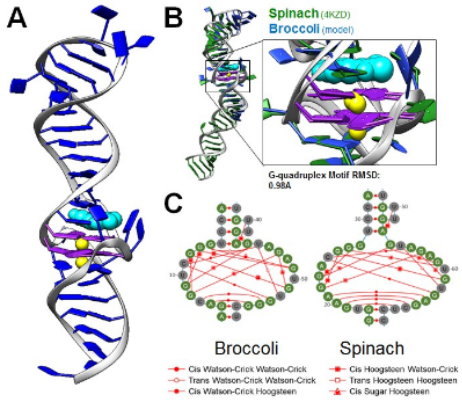
1 C. Tuerk and L. Gold, *Science*, 1990, **249**, 505–510.
2 A. D. Ellington and J. W. Szostak, *Nature*, 1990, **346**, 818–822.
3 D. L. Robertson and G. F. Joyce, *Nature*, 1990, **344**, 467–468.
4 J. S. Paige, K. Y. Wu and S. R. Jaffrey, *Science*, 2011, **333**, 642–646.
5 K. D. Warner, M. C. Chen, W. Song, R. L. Strack, A. Thorn, S. R. Jaffrey and A. R. Ferré-D'Amaré, *Nat. Struct. Mol. Biol.*, 2014, **21**, 658–663.
6 H. Huang, N. B. Suslov, N.-S. Li, S. A. Shelke, M. E. Evans, Y. Koldobskaya, P. A. Rice and J. A. Piccirilli, *Nat. Chem. Biol.*, 2014, **10**, 686–691.
7 K. Y. Han, B. J. Leslie, J. Fei, J. Zhang and T. Ha, *J. Am. Chem. Soc.*, 2013, **135**, 19033–19038.
8 P. Fernandez-Millan, A. Autour, E. Ennifar, E. Westhof and M. Ryckelynck, *RNA*, 2017, **23**, 1788–1795.
9 A. Autour, E. Westhof and M. Ryckelynck, *Nucleic Acids Res.*, 2016, **44**, 2491–2500.
10 R. L. Strack, M. D. Disney and S. R. Jaffrey, *Nat. Methods*, 2013, **10**, 1219–1224.
11 M. Okuda, D. Fourmy and S. Yoshizawa, *Nucleic Acids Res.*, 2017, **45**, 1404–1415.
12 G. S. Filonov, J. D. Moon, N. Svensen and S. R. Jaffrey, *J. Am. Chem. Soc.*, 2014, **136**, 16299–16308.
13 G. S. Filonov, W. Song and S. R. Jaffrey, *Biochemistry*, 2019, **58**, 1560–1564.
14 W. Zhou, R. Saran and J. Liu, *Chem. Rev.*, 2017, **117**, 8272–8325.
15 S. DasGupta, S. A. Shelke, N. Li and J. A. Piccirilli, *Chem. Commun.*, 2015, **51**, 9034–9037.
16 M. Vorlíčková, I. Kejnovská, J. Sagi, D. Renčiuk, K. Bednářová, J. Motlová and J. Kypr, *Methods*, 2012, **57**, 64–75.
17 R. Del Villar-Guerra, J. O. Trent and J. B. Chaires, *Angew. Chem., Int. Ed.*, 2018, **57**, 7171–7175.
18 F. A. P. Vendeix, A. M. Munoz and P. F. Agris, *RNA*, 2009, **15**, 2278–2287.
19 D. Tan, S. Piana, R. M. Dirks and D. E. Shaw, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, E1346–E1355.
20 J. C. Savage, P. Shinde, H. P. Bächinger, M. A. Davare and U. Shinde, *Chem. Commun.*, 2019, **55**, 5882–5885.
21 M. M. Fay, S. M. Lyons and P. Ivanov, *J. Mol. Biol.*, 2017, **429**, 2127–2147.
22 S. Y. Yang, P. Lejault, S. Chevrier, R. Boidot, A. G. Robertson, J. M. Y. Wong and D. Monchaud, *Nat. Commun.*, 2018, **9**, 4730.
23 R. Simone, P. Fratta, S. Neidle, G. N. Parkinson and A. M. Isaacs, *FEBS Lett.*, 2015, **589**, 1653–1668.
24 P. Kharel, S. Balaratnam, N. Beals and S. Basu, *Wiley Interdiscip. Rev.: RNA*, 2020, **11**, e1568.
25 A. Bugaut and S. Balasubramanian, *Nucleic Acids Res.*, 2012, **40**, 4727–4741.
26 M. J. Morris, Y. Negishi, C. Pazsint, J. D. Schonhoft and S. Basu, *J. Am. Chem. Soc.*, 2010, **132**, 17831–17839.
27 E. Crenshaw, B. P. Leung, C. K. Kwok, M. Sharoni, K. Olson, N. P. Sebastian, S. Ansaloni, R. Schweitzer-Stenner, M. R. Akins, P. C. Bevilacqua and A. J. Saunders, *PLoS One*, 2015, **10**, e0143160.
28 A. Arora and B. Suess, *RNA Biol.*, 2011, **8**, 802–805.
29 T. Simonsson, P. Pecinka and M. Kubista, *Nucleic Acids Res.*, 1998, **26**, 1167–1172.
30 J. L. Mergny, A. T. Phan and L. Lacroix, *FEBS Lett.*, 1998, **435**, 74–78.
31 T. Majerová, T. Streckerová, L. Bednárová and E. A. Curtis, *Biochemistry*, 2018, **57**, 4052–4062.
32 N. T. Dao, R. Haselsberger, M. E. Michel-Beyerle and A. T. Phan, *FEBS Lett.*, 2011, **585**, 3969–3977.
33 A. Arora and S. Maiti, *J. Phys. Chem. B*, 2009, **113**, 10515–10520.
34 D. Guet, L. T. Burns, S. Maji, J. Boulanger, P. Hersen, S. R. Wente, J. Salamero and C. Dargemont, *Nat. Commun.*, 2015, **6**, 8882.

# Supplemental Material

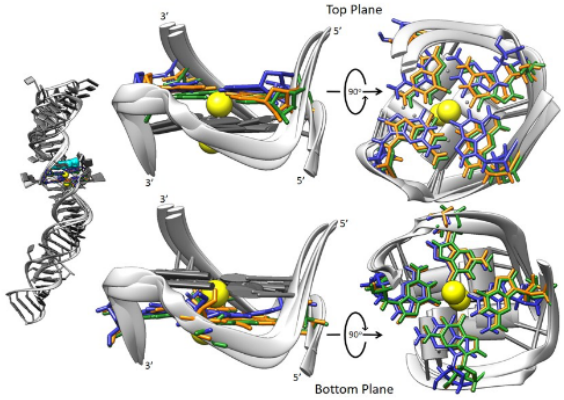
## Figures




**Supplemental 1: Full sequence alignment:** Comparison of Spinach derivative family members and Broccoli. Color coding: Red letters/grey background are G-Quad core motif residues; White letters/Black background are G-Quad guanine residues; Black Lettering/Grey background are conserved non-G-Quad residues; arrows indicated complementary regions that point towards the hairpin. Bottom and bold sequences correspond to crystalized sequences (or in the case of Broccoli, modelled here).




**Supplemental 2: (A) Broccoli model:** Cartoon representation of the homology model of Broccoli; DFHBI, G-Quad guanine bases, and $K^+$ ions are colored cyan, cyan, and yellow, respectively **(B) Spinach v Broccoli comparison:** The crystal Structure of Spinach (4KZD) and model Broccoli are aligned. The two show significant structural similarity, as would be expected based on the similarity between Spinach and iSpinach. G-Quad bases are purple, DFHBIs are cyan, and $K^+$ ions are yellow. **(C) Residue interaction map:** The G-Quad core residues are presented in a 2D flattened projection, with corresponding nucleotide interactions.




**Supplemental 3: G-Quadruplex Structural Alignment:** Comparison of G-quadruplex planes of Spinach (4KZD), iSpinach (5OB3), and Broccoli model, colored green, orange, and blue, respectively. DFHBI colored in cyan, and $K^+$ ions in yellow. Top and bottom planes are individually colored against a grey neutral background for visual clarity.

## Methods

**In vitro Transcription and purification:** RNA transcripts were prepared using linear dsDNA PCR templates. Sequences of the forward DNA strand are as follows (T7 promoter is underlined, construct in bold font): Spinach 5'—GCGCGCGAATTCTAATACGACTCACTATA**GGAGGACGCGACCGAAATG GTGAAGGACGGGTCCAGTGCGAAACACGCACTGTTGAGTAGAGTGTG AGCTCCGTAACTGGTCGCGTC**—3'; and Broccoli 5'—GCGCGCGAATTCTAATACGACTCACTATA**GGGAGACGGTCGGGTCCAG GCACACAAAAATGTTGCCTGTTGAGTAGAGTGTGGGCTCC**—3'. Prior to transcription, PCR templates were gel purified, phenol:chloroform cleaned, and desalted, similar to RNA clean up ( described below). Transcription reactions were assembled at room temperature in the following order, to final concentrations of: water, 1X T7 RNAP transcription buffer (NEB), 24mM $MgCl_2$, 4mM

each rNTP (NEB), 15-25ng/μl dsDNA, 5mM DTT, 1U/μl murine RNase inhibitor (NEB), 2.5mU/μl Yeast Inorganic Pyrophosphatase (NEB), 5U/μl T7 RNA polymerase (NEB). RNA was transcribed at 42°C for 4-10 hours. Transcriptions were quenched with DNase I treatment, and observed on a 4% agarose sodium borate gel, stained with Sybr Gold nucleic acid dye. Target transcript bands were purified by gel excision and electro-eluted into 3.5kDa MWCO dialysis tubing. RNA was concentrated with a 3.5KDa MWCO concentrating spin column (Millipore) and cleaned with acidified phenol:chloroform:isoamyl alcohol (Ambion). RNA was then desalted into 5mM Tris (pH 8.0) using a Zeba 7kDa MWCO desalting spin column (Thermo), and quantified by UV-absorbance at 260nm and fluorescence (Qubit HS RNA kit; Invitrogen), and stored at -80°C. Prior to experimental use of frozen samples, RNA was folded following a basic protocol: Dilution into desired buffer, then successive thermocycler incubations at 85°C (20sec), 50°C (10sec), 37°C (10sec), addition of DFHBI, 4°C (≥30min).

**Fluorescence:** All fluorescent measurements were performed on a Photon Technologies International QM-1 steady state fluorescent spectrophotometer. All samples were read under the following solution conditions unless otherwise stated: 0.5μM RNA, 20mM Tris (pH 8.0), 100mM KCl, 5mM MgCl2, 10μM DFHBI. Samples were 20°C held constant by a circulating water bath, excited at 468nm in a 2mm quartz fluorescence cuvette. Kinetic experimentation was performed by rapid, in-cuvette mixing of folded Apo-state RNA in buffer and a 10% volume of 1M KCl in water. Emission was then read at 501nm every 29.75 seconds with a 0.25 second shutter exposure. KCl titrations were performed by folding RNA in the presence of various KCl concentrations in individual reaction tubes. Samples were then incubated on ice for 2hr to ensure equilibrium was reached. Fluorescent of each sample was individually scanned (468/501nm). Fluorescence data was plotted and fit in GraphPad Prism 5.

**Circular Dichroism:** Measurements were collected on an Aviv 202 CD Spectrometer at 20°C. RNA samples were assayed in 10mM Cacodylate (pH 7.4), 100mM KCl, 5mM $MgCl_2$. RNA was folded as described above. 135ng/μl RNA samples were scanned in 1mm quartz CD cuvette from 320 to 190nm by a 0.2nm step, with a 3 second integrated read per step. Data presented as average of three individual scans. Control scans of buffer were subtracted from spectra. Values plotted in GraphPad Prism 5.

**Thermal Stability:** Thermal melting was observed using an ABI StepOne Real Time PCR Thermocycler. 0.5μM RNA was folded as described in 20mM Tris (pH 8.0), 5mM $MgCl_2$, 100mM KCl, 20μM DFHBI. 20μl sample volumes were scanned in triplicate wells in an ABI brand 96 well white opaque QPCR plate. Fluorescence was measured in the Sybr Green Ex/Em channel from 4°C to 70°C at a 2% gradient. Control scans of DFHBI in buffer were subtracted from RNA spectra. Fluorescent values averaged and plotted in GraphPad Prism 5.

**Homology modelling:** Initial homology models for Broccoli were produced using automated 3D RNA structural prediction servers, RNA Composer[1], ModeRNA[2], and SimRNA[3]. While these servers provided 3D predictions consistent with their structural templates, a complete and fully automated predictive model for Broccoli was only partially generated. Hence, a local multiple sequence alignment (MSA) of the target RNA, Broccoli, with derivatives of Spinach, and multiple structural templates 4TS2[5], 4KZD[6], and 5OB3[8] was obtained using ClustalW, and edited to ensure optimal base complementarity indicated by arrows pointing towards the hairpin (Fig.2A). The G-Quad motif for Broccoli was generated from 5OB3 based on the MSA that established this region of the target RNA is almost identical within all available structural templates. An extensive fragment library search identified a loop of 16-nucleotides from a zinc finger RNA (1UN6, residues 13-28) as an optimal fragment to bridge the G-quad motif in Broccoli from residues 20-35[4]. The model was energy minimized using a modified Amber14ff RNA force field[16] in YASARA modelling suite, and the quality of the model was assessed using RNAssess[5], a computational server that discriminates between potentially correct and incorrect conformations by comparing RNA 3D models with the reference structures.

**MD simulation:** Generated Broccoli model and Spinach (4KZD) without co-crystalized Fab were used in MD simulations. A revised AMBER ff14 RNA force field[19] was used for energy minimizations, using an explicit solvent TIP3P water model. MD simulations were run for 50ns, with snapshots taken every 100ps, using the YASARA md_run macro with explicit solvent conditions at 298K with a physiological intracellular NaCl concentration. Hydrogen bonding atoms present in the bases of the starting aptamer structure were identified and the distances between them were catalogued for every snapshot. Hydrogen bond values were plotted in Sigma Plot, and structural images presented with Chimera.

## References:


1 M. Popenda, M. Szachniuk, M. Antczak, K. J. Purzycka, P. Lukasiak, N. Bartol, J. Blazewicz and R. W. Adamiak, *Nucleic Acids Res.*, 2012, **40**, e112.

2 M. Rother, K. Rother, T. Puton and J. M. Bujnicki, *Nucleic Acids Res.*, 2011, **39**, 4007–4022.

3 M. J. Boniecki, G. Lach, W. K. Dawson, K. Tomala, P. Lukasz, T. Soltysinski, K. M. Rother and J. M. Bujnicki, *Nucleic Acids Res.*, 2016, **44**, e63.

4 D. Lu, M. A. Searles and A. Klug, *Nature*, 2003, **426**, 96–100.

5 P. Lukasiak, M. Antczak, T. Ratajczak, M. Szachniuk, M. Popenda, R. W. Adamiak and J. Blazewicz, *Nucleic Acids Res.*, 2015, **43**, W502–W506.

ChemComm

COMMUNICATION

View Article Online
View Journal



Cite this: DOI: 10.1039/d0cc07042d

Received 22nd October 2020,
Accepted 17th December 2020

DOI: 10.1039/d0cc07042d

rsc.li/chemcomm

# A Broccoli aptamer chimera yields a fluorescent $K^+$ sensor spanning physiological concentrations†

Jonathan C. Savage,[a] Pushkar Shinde,[b] Yizhou Yao,[a] Monika A. Davare[c] and Ujwal Shinde *[a]

**The RNA aptamer Broccoli accepts 2′fluorinated (2′F) pyrimidine nucleotide incorporation without perturbation of structure or fluorescence in the presence of potassium and DFHBI. However, the modification decreases Broccoli's apparent affinity for $K^+$ >30-fold. A chimera of Broccoli RNAs with mixed chemistries displays linear fluorescent gain spanning physiological $K^+$ concentrations, yielding an effective RNA-based fluorescent $K^+$ sensor.**

The RNA aptamer "Broccoli" was identified by SELEX to be an improved version of the aptamer "Spinach", the first fluorescent RNA equivalent of green fluorescent protein.[1,2] Fluorescent emission is contingent upon the association of three individual components, forming a "tripartite complex". This complex is composed of (1) a unique RNA G-quadruplex fold stabilized by (2) selective cation coordination, which promotes (3) the docking of fluorophore DFHBI (3,5-difluoro-4-hydroxybenzylidene imidazolinone) or select derivatives (Fig. 1A).[3–5] Fluorescence of Broccoli or Spinach is absolutely contingent upon the proper association of all tripartite complex components. Though the structure of Broccoli has not been solved crystallographically, sequence alignment, spectroscopic studies, and computational modelling suggest Broccoli shares significant structural similarity with Spinach aptamers.[1,3,5–8]

While biological systems required billions of years to select riboswitch aptamers capable of sensing ions and metabolites, *in vitro* evolution experiments by SELEX allow for rapid identification of a wider range of aptamer structures and functions.[9–14] Incorporation of synthetic nucleotide derivatives further expand the chemo-structural space RNA can adopt outside the constraints of biological compatibility. Most published aptamers do not contain modified nucleotides as intracellular application is constrained. Because of this, use of modified nucleotides to functionally retune characterized aptamer structures for *in vitro* utility remains unexplored.

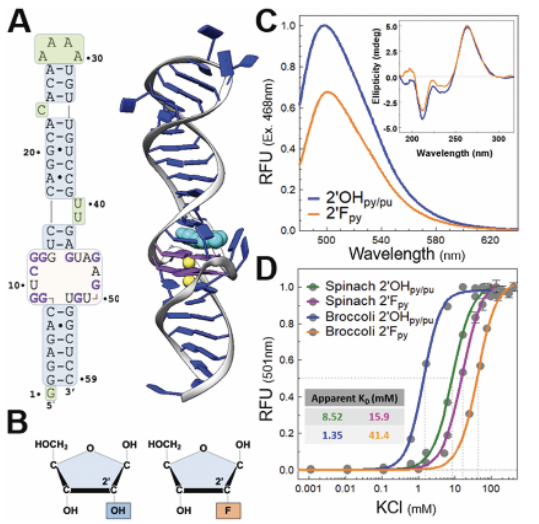



**Fig. 1** Broccoli accepts fluorinated pyrimidine nucleotides: (A) a 2D projection with sequence, and a 3D structural model of Broccoli RNA. In both representations, G-quadruplex guanines are purple, coordinating $K^+$ ions are yellow, and DFHBI is cyan. Broccoli structure adapted from Savage *et al.*, 2020.[5] (B) 2′hydroxylated and 2′fluorinated pyrimidines differ only by the functional group present at the ribose C2 position. (C) Fluorescence comparison between $2'OH_{py/pu}$ and $2'F_{py}$ Broccoli RNA with 20 μM DFHBI and 100 mM KCl. (C; inset) Circular dichroism (CD) spectra of $2'OH_{py/pu}$ and $2'F_{py}$ Broccoli show the two constructs share nearly identical secondary structures. (D) Fluorescence-based KCl titrations with RNAs provides an approximate $K_D$ for $K^+$ at the G-quadruplex. Despite the structural similarity between Spinach and Broccoli, the two RNAs display dramatic differences in $K^+$ association when $2'F_{py}$ nucleotides are incorporated, with a <2-fold difference between Spinach constructs, but >30-fold between Broccoli constructs. (inset) Apparent $K_D$ values for $K^+$ determined from fits.

[a] *Department of Chemical Physiology & Biochemistry, Oregon Health & Science University, 3181 SW Sam Jackson Park Rd, Portland, OR 97239, USA. E-mail: shindeu@ohsu.edu*

[b] *Department of Chemistry, Emory University, 1762 Clifton Rd, Atlanta, GA 30022, USA*

[c] *Papé Family Pediatric Research Institute, Division of Pediatric Hematology/Oncology, Department of Pediatrics, Oregon Health & Science University, 3181 SW Sam Jackson Park Rd, Portland, OR 97239, USA*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d0cc07042d

Previous work has demonstrated Spinach can be transcribed with 2′fluorinated (2′F) pyrimidine nucleotides without perturbing fluorescence or structure.[15] Though numerous modified nucleotide chemistries are available, 2′F is commonly used as it is inexpensive, accepted by certain RNA polymerases without loss of fidelity, increases nuclease resistance, and is well published.[15–19] Despite sequence similarities at the G-quadruplex with Spinach, Broccoli is smaller overall with increased thermal stability, making it more attractive for *in vitro* application (Fig. 1A).[1,5] Hence, we explored the use of modified chemistry in Broccoli aptamer and uncovered distinct biophysical properties open new avenues for *in vitro* biosensor applications.

Akin to Spinach, Broccoli transcribed with 2′F pyrimidines—without any change in primary nucleobase sequence—folds correctly, binds DFHBI and $K^+$, and displays dramatically increased chemical stability in the presence of RNase A (Fig. 1 and Fig. S1, ESI†). RNAs transcribed *in vitro* with 2′hydroxylated ($2'OH_{py/pu}$) or 2′fluorinated pyrimidines ($2'F_{py}$) (Fig. 1B) were purified and compared by fluorescence and circular dichroism (CD) spectroscopies (Fig. 1C and Methods). Fluorescence with KCl yielded spectra of $2'F_{py}$ that display no change in excitation and emission wavelengths, albeit with a 25% reduction in fluorescent intensity compared to $2'OH_{py/pu}$ (Fig. 1C). CD reveals both Broccoli constructs share similar secondary structures (Fig. 1C, inset), which are nearly superimposable on the spectrum of Spinach.[5,15]

Despite initial similarities, obvious differences between Broccoli and Spinach emerge at sub-saturating $K^+$ concentrations (Fig. 1D). Broccoli $2'OH_{py/pu}$ displays a higher affinity for $K^+$ than Spinach $2'OH_{py/pu}$.[1,5] However, Broccoli $2'F_{py}$ displays a lower affinity for $K^+$ than Spinach $2'F_{py}$ (Fig. 1D). The resulting apparent affinities between Spinach constructs differ <2-fold, but Broccoli constructs differ >30-fold. The subtle sequence variation between Broccoli and Spinach imparts significant biophysical differences in cation association and tripartite complex stability, which is exaggerated with 2′F pyrimidine incorporation.

To examine changes in affinity in finer detail, $K^+$ association kinetics were assayed as described (Fig. 2; Methods).[15,20] Kinetics were performed with each RNA construct in isolation (Fig. 2A). Fits of these data provide apparent association constants ($K_{obs}$), each containing discrete component binding ($k_{on}$) and unbinding ($k_{off}$) rates within it. To determine those components, $K_{obs}$ values are plotted against corresponding KCl concentrations. Linear regression provides $k_{on}$ (slope) and $k_{off}$ (*y*-intercept) (Fig. 2B). The results show pyrimidine 2′fluorination in Spinach increases $k_{off}$ with little effect on $k_{on}$, whereas the same modification in Broccoli both increases $k_{off}$ and reduces $k_{on}$. Individual kinetic values, and calculated dissociation constant $K_D$, roughly concur with $K^+$ affinities approximated in Fig. 1D (Fig. 2B inset).

To rationalize the kinetic observations, a structure of $2'F_{py}$ Broccoli was modeled and molecular dynamic (MD) simulations were performed to directly compare RNA constructs, 2′-chemistries, and $K^+$-coordination states (Methods).[5,15] Root mean square deviation (RMSD) analysis provides insightful structural information, but it does not allow for analysis of structural convergences that may occur during a simulation. To capture this information we performed a principal component analysis (PCA) of G-quad motifs from all time-points of each simulation (Fig. S2, ESI†). Projection plots of the

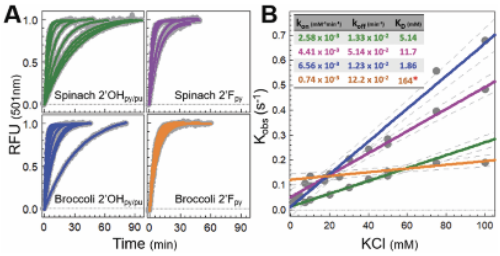



Fig. 2 Kinetics of fluorescent tripartite complex formation in Spinach and Broccoli: (A) individual kinetic traces and colored fits of $K^+$ association, assayed by fluorescence. Each trace represents a single KCl concentration. All traces performed with constant RNA and DFHBI concentrations (see Methods). (B) Fits of each kinetic trace in (A) yield a $K_{obs}$ apparent rate constant. When each $K_{obs}$ is plotted against its respective KCl concentration and fit, individual kinetic components can be extrapolated; $k_{on}$ and $k_{off}$ determined from the slope and *y*-intercept, respectively. Despite their sequence and structural similarities, 2′F incorporation of Spinach reduces $k_{off}$ with negligible effect of $k_{on}$, whereas the same modification in Broccoli both reduces $k_{on}$ and increases $k_{off}$, explaining affinity differences observed in Fig. 1D. (inset) Explicit values determined from fits in (B) with the same color coordination. Red asterisk denotes high $K_D$ value for Broccoli $2'F_{py}$ which may be unreliable with this particular technique. KCl additions too close to an apparent $K_D$ yield non-linear $K_{obs}$ values.

Spinach and Broccoli structures onto the subspace spanned by their respective first two principal components indicate the presence of several distinct conformational populations dependent upon the construct, chemistry, and coordination state. $2'OH_{py/pu}$ Broccoli samples a reduced set of distinct populations than $2'OH_{py/pu}$ Spinach, in both $K^+$-coordinated and Apo states, and is consistent with our experimental data that 2′OH Broccoli is more stable than $2'OH_{py/pu}$ Spinach. $2'F_{py}$ Spinach and Broccoli display the same population distribution when $K^+$-bound, yet $2'F_{py}$ Broccoli samples more distinct structural populations in its Apo state. We interpret this as evidence that $2'F_{py}$ Broccoli is less stable than $2'F_{py}$ Spinach, which is consistent with our experimental data (Fig. 1D and 2). This work corroborates to the Broccoli structural model and provides an additional approach for future *in silico* characterization of sequence-optimized constructs.

The results hitherto establish the same chemical modification in two similar RNAs does not yield effects proportional to their structural relatedness. While fluorescent and CD spectroscopies show Spinach and Broccoli retain their structural identity, binding kinetics and apparent affinities of $K^+$ are altered (Fig. 1C, D, 2 and Fig. S2, ESI†). We therefore explored how modified Broccoli could be exploited in a way Spinach cannot.

The combined detection limits of the two individual Broccoli constructs span physiological $K^+$ concentrations of ~0.15 to 100 mM, more than 20-fold the concentration Spinach constructs cover (Fig. 1D). We postulated a mixture of $2'OH_{py/pu}$ and $2'F_{py}$ Broccoli would produce a signal with a broad range of linear gain, equivalent to the combined upper and lower detection limits of the individuals. Such results could open avenues for development of RNA-based fluorescent $K^+$ sensors, tools with diagnostic potential in dysregulation of potassium-dependent homeostatic functions including blood pressure, muscular contraction, circadian rhythm, and others.[21]

Three ratios of Broccoli constructs were mixed—2 : 1, 1 : 1, 1 : 2—and KCl titrations were performed (Fig. 3A). The widest

detection range with a linear regression fit of $R^2 > 0.90$ was attributed to the 1 : 1 mixture (Fig. 3A inset). To engineer the mixture into a single construct which can be maintained under non-equilibrium conditions, complementary 3′ extensions with unfavorable self-annealing were designed and added to each Broccoli RNA (Fig. 3B, C, and Methods). The resulting construct contains two units of identical nucleobase sequence but is a hybrid of the two individual's chemistries, a chimera. The Chimera contains two distinct G-quadruplex motifs, each forming a separate tripartite complex with a respective $K^+$ ion and DFHBI pair. We postulated the annealed overhangs would not alter the behavior of the individual tripartite complexes, given the highly modular nature of RNA structural elements. As hypothesized, chimera formation does not change the $K^+$ titration profile or fluorescence intensity as compared to an un-annealed mixture, suggesting an absence of cooperativity in cation association after annealing (Fig. 3C(upper right) and D). The Chimera displays KCl association kinetics that roughly split the difference between $2'OH_{py/pu}$ and $2'F_{py}$ monomers (Fig. S3, ESI†). Similarly, as expected, the fluorescent properties of DFHBI in the $2'OH_{py/pu}$ and $2'F_{py}$ Broccoli subunits of the chimera remain unchanged with excitation and emission wavelengths at 468 and 501 nm, respectively. Thus both individual emissions coalesce into a single readout signal (Fig. 3C upper right). The Chimera produces maximal fluorescence at pH 8, and displays greater than 70% fluorescent signal intensity in up to 50% cell culture serum for over 60 minutes (Fig. S4, ESI†). Linear regression of KCl titration shows the Chimera to be suitable for visualizing $K^+$ over a wide concentration range, with signal changes observable as low as 70 μM (Fig. 3D). Additionally, the Chimera displays sufficient selectivity and $K^+$ preference over alternate cations, confirming its potential as a $K^+$ sensor (Fig. 3C).

The sensor was tested in an *in vitro* cell system to observe bulk $K^+$ flux upon hypotonic lysis of HEK293T cells (Fig. 4). Cultured cells at various confluence were lysed in a $K^+$-free hypotonic buffer containing DFHBI and Broccoli Chimera. Lysis-induced efflux of $K^+$ was observed (Fig. 4 upper inset) and interpolated from a standard curve (Fig. 4, and Methods). Changes in calculated $K^+$ concentrations are directly proportional to the observed cell density prior to hypotonic lysis, validating sensor applicability in an *in vitro* cell system (Fig. 4 upper inset).

In this work only select riboses were altered, not the nucleobase sequence. 2′F pyrimidines result in retained structure and fluorescence but dramatically alter cation association. Numerous alternately modified nucleotides are commercially available, and may prove more appropriate for certain applications.[22] Additionally, chemical synthesis—as opposed to transcription—allows for combining chemistries with site-specific incorporation.[23] The fluorescent property and small size of Broccoli's G-quadruplex makes it possible to test every combination of base and chemistry at each position. Based on these results we posit two avenues of aptamer modification—sequence and chemistry—can be used synergistically for finer adjustment of functional tuning for new applications. There are several examples of previously reported aptamer fusions.[24–26] However, these fusions are the result of connecting the primary sequences of two aptamers together to be transcribed as a single unit. There are advantages of this approach, though it restricts the

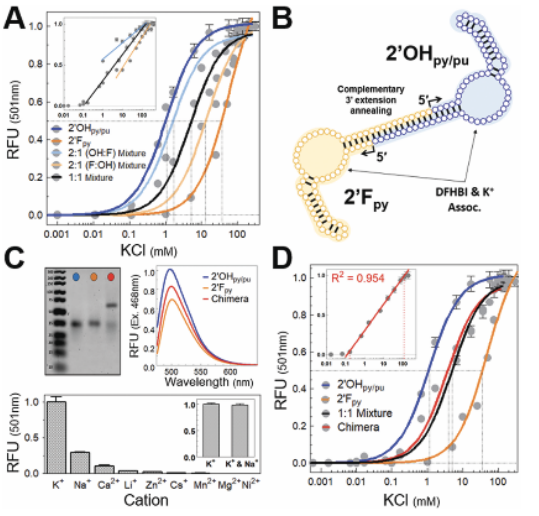



**Fig. 3** Broccoli RNA chimera construction: (A) fluorescence-based KCl titration of various ratios of monomeric Broccoli constructs, showing predictable curve shifting based on relative 2′chemistry composition. (inset) Linear regression of KCl concentrations was performed, containing the most points while retaining an $R^2 > 0.90$, of which the 1 : 1 ratio displays widest detection range. Inset axes identical to outer graph. (B) 2D schematic of chimera design, showing WC stems, G-quadruplex motif and complementary 3′extensions (see Methods for sequences). Chimera contains two G-quadruplex domains forming two independent tripartite complexes. Overhang sequences in ESI.† (C) (upper left) Agarose gel EMSA shows mass shift of chimera compared to the two monomeric chemistries. (upper right) Comparative fluorescence spectra of chimera and monomeric chemistries. Monomers at 0.5 μM and chimera at 0.25 μM, to account for (2) G-quadruplex motifs per chimera. Gel lane colors in EMSA correspond to fluorescence signal colors. (lower) Chimera fluorescence in the presence of various mono- and divalent cations (all chloride salts) normalized to KCl signal. Each cation present at 50 mM in buffer (Methods). (lower; inset) Chimera fluorescence in the presence of 50 mM $K^+$ is unchanged with or without the presence of competing 50 mM $Na^+$. (D) Fluorescence-based KCl titrations of $2'OH_{py/pu}$, $2'F_{py}$, 1 : 1 mixture, and annealed chimera. Both the 1 : 1 mixture and the annealed chimera split the difference between two monomeric chemistries with respect to apparent $K_D$ and broader detection range. Difference between 1 : 1 mixture and chimera is negligible, confirming 3′extension and annealing does not alter $K^+$ association and tripartite complex formation for individual chimeric subunits. (inset) Linear regression shows the chimera to have a linear range of detection from ~150 μM to ~100 mM. Inset axes identical to outset.

fusion product to a single nucleotide chemistry. This work represents, to our knowledge, the first example of a mixed chemistry chimera, generated *in vitro* via multiple synthesis reactions, expanding the aptamer toolkit.

Functionally, the RNA chimera is a valuable addition to the available methods for fluorescent $K^+$ sensing, which contains several small molecule chelators, a single protein-based construct employing FRET or a split-eGFP, and three DNA constructs, which are the most chemically similar (Table S1, ESI†).[27–37] The DNA sensors PSO-1,2 and TBA also employ G-quadruplex folds for $K^+$ docking, though, unlike the chimera, the fluorescence of these DNA sensors is based on termini-linked FRET pair interaction upon $K^+$-induced G-quadruplex folding.[34–37] While this detection scheme has a lower background signal than the chimera, the linear range of function $K^+$

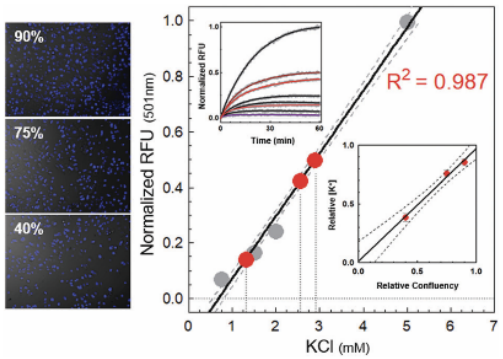



Fig. 4 *In vitro* detection of $K^+$ flux by cellular hypotonic lysis. HEK293T cells were cultured to ~40, 75, 90% confluence and media replaced with KCl free lysis buffer (see Methods). (upper inset) Raw fluorescence signal gain was collected (red fits) next to KCl positive standards (grey fits) and negative controls (purple fit). Unknown $K^+$ concentrations (red dots) were interpolated from a standard curve (grey dots) generated from nonlinear regression of positive controls. (lower inset) Approximate cell confluence and interpolated $K^+$ concentrations were directly proportional over the range tested.

detection of these tools is within a single order of magnitude, ~1–10 and ~100–200 mM.[34,37] The complementary molecular geometry of G-quadruplexes and desolvated potassium is sufficiently favorable and selective to suggest any nucleic acid sensor engineered for detecting $K^+$ in the millimolar range will likely employ the motif. Tuning such sensors will be heavily based on modulating the number of G-quadruplex planes and optimizing proximal residues, changing the number of bound potassium and their binding affinity. In the case of this chimera, additional tuning can be achieved by optimizing construct preference for alternate DFHBI derivatives with differential affinity, excitation/emission, brightness, and photo-bleaching.[20,38–40]

Nucleic acid aptamers have ligand affinity with selectivity that rival antibodies. Using aptamers to detect or "sense" their molecular targets in ways similar to antibodies has been commonplace for some time. However nucleic acid sensors for the accurate quantification of their ligands are less developed, for which fluorescent aptamers are ideal. We have found the unaltered sequence of Broccoli aptamer allows for quantification of $K^+$ when annealed to a modified $2'F_{py}$ construct of the same sequence. A chimera of the two annealed RNAs produces a stoichiometric unit that can be maintained in non-equilibrium conditions. While the properties of the monomeric subunits are unaltered, their individual signals coalesce into a single readout spanning a linear range of $K^+$ detection of nearly three orders of magnitude. This platform provides a promising foundation for future ion-selective optimization, which could yield a flight of RNA sensors for the real-time fluorescent quantification of cellular cation flux.

The authors thank Drs Dave Farrens, Matt Thayer, and Larry David for their discussion and use of instrumentation.

## Conflicts of interest

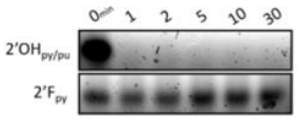
The authors of this manuscript have no conflicts to declare.

## Notes and references

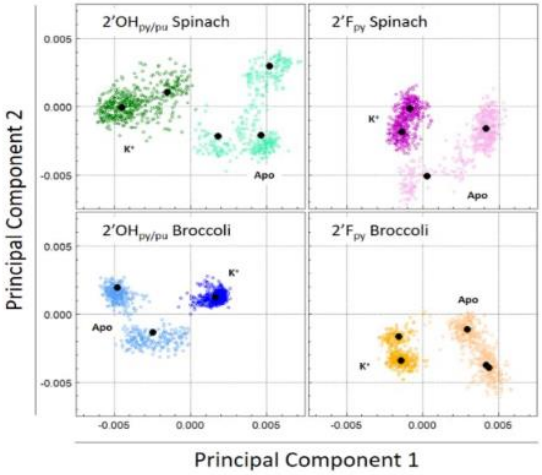
1 G. S. Filonov, J. D. Moon, N. Svensen and S. R. Jaffrey, *J. Am. Chem. Soc.*, 2014, **136**, 16299–16308.
2 J. S. Paige, K. Y. Wu and S. R. Jaffrey, *Science*, 2011, **333**, 642–646.
3 K. D. Warner, M. C. Chen, W. Song, R. L. Strack, A. Thorn, S. R. Jaffrey and A. R. Ferré-D'Amaré, *Nat. Struct. Mol. Biol.*, 2014, **21**, 658–663.
4 H. Huang, N. B. Suslov, N.-S. Li, S. A. Shelke, M. E. Evans, Y. Koldobskaya, P. A. Rice and J. A. Piccirilli, *Nat. Chem. Biol.*, 2014, **10**, 686–691.
5 J. C. Savage, M. A. Davare and U. Shinde, *Chem. Commun.*, 2020, **56**, 2634–2637.
6 G. S. Filonov, W. Song and S. R. Jaffrey, *Biochemistry*, 2019, **58**, 1560–1564.
7 H. Huang, N. B. Suslov, N.-S. Li, S. A. Shelke, M. E. Evans, Y. Koldobskaya, P. A. Rice and J. A. Piccirilli, *Nat. Chem. Biol.*, 2014, **10**, 686–691.
8 P. Fernandez-Millan, A. Autour, E. Ennifar, E. Westhof and M. Ryckelynck, *RNA*, 2017, **23**, 1788–1795.
9 G. F. Joyce, *Gene*, 1989, **82**, 83–87.
10 A. D. Ellington and J. W. Szostak, *Nature*, 1990, **346**, 818–822.
11 C. Tuerk and L. Gold, *Science*, 1990, **249**, 505–510.
12 D. L. Robertson and G. F. Joyce, *Nature*, 1990, **344**, 467–468.
13 R. R. Breaker, in *The RNA World*, ed. R. F. Gesteland, T. R. Cech and J. F. Atkins, Cold Spring Harbor Laboratory Press, 3rd edn, 2006, pp. 89–108.
14 R. R. Breaker, *Cold Spring Harbor Perspect. Biol.*, 2012, **4**, a003566.
15 J. C. Savage, P. Shinde, H. P. Bächinger, M. A. Davare and U. Shinde, *Chem. Commun.*, 2019, **55**, 5882–5885.
16 R. Sousa and R. Padilla, *EMBO J.*, 1995, **14**, 4609–4621.
17 R. Padilla and R. Sousa, *Nucleic Acids Res.*, 2002, **30**, e138.
18 A. Patra, M. Paolillo, K. Charisse, M. Manoharan, E. Rozners and M. Egli, *Angew. Chem., Int. Ed.*, 2012, **51**, 11863–11866.
19 P. S. Pallan, E. M. Greene, P. A. Jicman, R. K. Pandey, M. Manoharan, E. Rozners and M. Egli, *Nucleic Acids Res.*, 2011, **39**, 3482–3495.
20 K. Y. Han, B. J. Leslie, J. Fei, J. Zhang and T. Ha, *J. Am. Chem. Soc.*, 2013, **135**, 19033–19038.
21 M. L. Gunz, L. Rabinowitz and C. S. Wingo, *N. Engl. J. Med.*, 2015, **373**, 60–72.
22 M. A. Dellafiore, J. M. Montserrat and A. M. Iribarren, *Front. Chem.*, 2016, **4**, 18.
23 S. Ni, H. Yao, L. Wang, J. Lu, F. Jiang, A. Lu and G. Zhang, *Int. J. Mol. Sci.*, 2017, **18**, 1683–1704.
24 C. A. Kellenberger, S. C. Wilson, J. Sales-Lee and M. C. Hammond, *J. Am. Chem. Soc.*, 2013, **135**, 58.
25 J. S. Paige, T. Nguyen-Duc, W. Song and S. R. Jaffrey, *Science*, 2012, **335**, 1194.
26 C. A. Kellenberger, S. C. Wilson, J. Sales-Lee and M. C. Hammond, *J. Am. Chem. Soc.*, 2013, **135**, 4906–4909.
27 P. Sekar, D. Y. Huang, S. F. Chang and W. W. Lin, *OncoTargets Ther.*, 2018, **9**, 12718–12731.
28 S. E. Kasner and M. B. Ganz, *Am. J. Physiol.: Renal, Fluid Electrolyte Physiol.*, 1992, **262**, 462–467.
29 X. Kong, F. Su, L. Zhang, J. Yaron, F. Lee, Z. Shi, Y. Tian and D. R. Meldrum, *Angew. Chem., Int. Ed.*, 2015, **54**, 12053–12057.
30 X. Zhou, F. Su, Y. Tian, C. Youngbull, R. H. Johnson and D. R. Meldrum, *J. Am. Chem. Soc.*, 2011, **133**, 18530–18533.
31 T. S. Rimmele and J.-Y. Chatton, *PLoS One*, 2014, **9**, e109243.
32 H. Bischof, M. Rehberg, S. Stryeck, K. Artinger, E. Eroglu, M. Waldeck-Weiermair, B. Gottschalk, R. Rost, A. T. Deak, T. Niedrist, N. Vujic, H. Lindermuth, R. Prassl, B. Pelzmann, K. Groschner, D. Kratky, K. Eller, A. R. Rosenkranz, T. Madl, N. Plesnila, W. F. Graier and R. Malli, *Nat. Commun.*, 2017, **8**, 1–12.
33 Y. Shen, S. Y. Wu, V. Rancic, A. Aggarwal, Y. Qian, S. I. Miyashita, K. Ballanyi, R. E. Campbell and M. Dong, *Commun. Biol.*, 2019, **2**, 18.
34 H. Ueyama, M. Takagi and S. Takenaka, *J. Am. Chem. Soc.*, 2002, **124**, 14286–14287.
35 S. Nagatoishi, T. Nojima, B. Juskowiak and S. Takenaka, *Angew. Chem.*, 2005, **117**, 5195–5198.
36 T. Nojima, H. Ueyama, M. Takagi and S. Takenaka, *Nucleic Acids Res. Suppl.*, 2002, 125–126.
37 K. Ohtsuka, S. Sato, Y. Sato, K. Sota, S. Ohzawa, T. Matsuda, K. Takemoto, N. Takamune, B. Juskowiak, T. Nagai and S. Takenaka, *Chem. Commun.*, 2012, **48**, 4740–4742.
38 X. Li, H. Kim, J. L. Litke, J. Wu and S. R. Jaffrey, *Angew. Chem., Int. Ed.*, 2020, **59**, 4511–4518.
39 W. Song, R. L. Strack, N. Svensen and S. R. Jaffrey, *J. Am. Chem. Soc.*, 2014, **136**, 1198–1201.
40 P. Swetha, Z. Fan, F. Wang and J. H. Jiang, *J. Mater. Chem. B*, 2020, **8**, 3382–3392.

# Supplemental Material

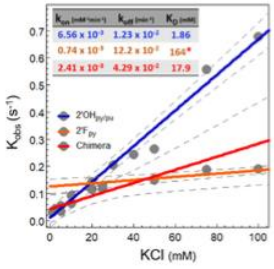
## Figures



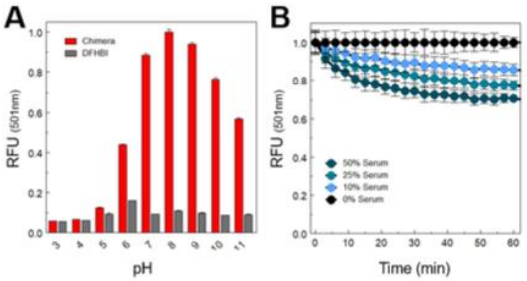

Supplemental 1. RNase stability of 2'hydroxylated and 2'fluorinated pyrimidine Broccoli RNAs. Each construct was incubated with purified RNase A for given time before being quenched with EDTA and loading dye. Broccoli transcribed with 100% 2'OH rNTPs ($2'OH_{py/pu}$) was completely degraded within 1min of RNase A exposure, whereas the incorporation of 2'F pyrimidine NTPs resulted in no visible degradation for more than 30minutes.




Supplemental 2. 250ns Molecular dynamic simulations were performed to observe structural differences between 2'chemistries in Spinach and Broccoli over time. Principal component analysis of G-quadruplex motifs under 3 variable conditions are presented (RNA identity, 2'chemistry, and $K^+$ coordination). Each colored dot corresponds to one G-quadruplex structure, sampled every 100ps; Black dots correspond to the structural average of each unique population.




Supplemental 3. Chimera $K^+$ binding kinetics. Following the identical protocol described for $2'OH_{py/pu}$ and $2'F_{py}$, 0.25μM Chimera was incubated with various KCl concentrations, and tripartite complex formation was observed by fluorescence. Fluorescent traces of individual KCl concentrations were fit to acquire $K_{obs}$ values. These $K_{obs}$ are plotted above against their respective KCl concentrations, overlaid with identical data for $2'OH_{py/pu}$ and $2'F_{py}$ monomers adapted from figure 2. The Chimera behaves as a mixture of the two component monomers, as expected.




Supplemental 4. Assessment of Chimera stability. (A) pH-dependent fluorescence of Chimera shows construct displays maximal complex fluorescence at pH 8. (B) Chimera fluorescence was observed over time in the presences of increasing cell culture serum concentrations, 10, 25, and 50% by volume. Buffer and KCl concentrations held constant for each serum concentration. Serum was a 1:1 mixture of Bovine Growth Serum and Fetal Calf Serum.

Table 1. Comparative traits of published fluorescent $K^+$ sensors. Values as reported in stated supplemental references.

| Name | Chemistry | $K_D$ (mM) | Functional Range (mM) | λ Fluorescent Emission (nm) | Reference |
|---|---|---|---|---|---|
| NK1 | Small Molecule | 200 | 10-50 | 572 | 1 |
| KS6 | Small Molecule | — | 30-500 | 540 | 2 |
| TAC | Small Molecule | — | — | 540 | 3 |
| TAC-Red | Small Molecule | — | 0-50 | 574 | 4 |
| TAC-Chrimson | Small Molecule | — | 20-60 | 600 | 5 |
| PBFI | Small Molecule | 8 | — | 500 | 6 |
| CDC222 | Small Molecule | 15-34 | 2.5-6.5 | 445 | 7 |
| GEPII (family) | Protein | 0.1-20 | 0.01-1000 | 550 | 8 |
| KIRIN1 | Protein | 1.67 | 0.1-10 | 530 | 9 |
| GINKO | Protein | 0.4 | 3-100 | 514 | 9 |
| PSO-1 | DNA | — | — | 581 | 10 |
| PSO-2 | DNA | 7.3 | 0.5-10 | 480 | 11 |
| TBA | DNA | — | 100-200 | 583 | 11 |
| Broccoli Chimera | RNA (Modified) | 8-18 | 0.15-100 | 501 | — |

## Methods

**In vitro Transcription and purification:** RNA transcripts were prepared using linear dsDNA PCR templates. Sequences of the forward DNA strand are as follows (T7 promoter is underlined, construct in bold font): Spinach 5'—GCGCGCGAATTCTAATACGACTCACTATA**GGAGGACGCGACCGAAATGGTGAAGGACGGGTCCAGTGCGAAACACGCACTGTTGAGTAG**AGTGTGAGCTCCGTAACTGGTCGCGTC—3'; and Broccoli 5'—GCGCGCGAATTCTAATACGACTCACTATA**GGGAGACGGTCGGGTCCAGGCACACAAAAATGTTGCCTGTTGAGTAGAGTGTGGGCTCC**—3'. Prior to transcription, PCR templates were gel purified, phenol:chloroform cleaned, and desalted, similar to RNA clean

up (described below). Transcription reactions were assembled at room temperature in the following order, to final concentrations of: water, 1X T7 RNAP transcription buffer (NEB), 24mM $MgCl_2$, 4mM each NTP (2'OH purines from NEB, 2'F pyrimidines from Trilink Biotechnologies), 15-25ng/µl dsDNA, 5mM DTT, 1U/µl murine RNase inhibitor (NEB), 2.5mU/µl Yeast Inorganic Pyrophosphatase (NEB), 5U/µl WT T7 RNA polymerase (NEB) or 1.25U/µl mutant T7 R&D polymerase (Lucigen, Inc). RNA was transcribed at 42°C for 2-5 hours. Transcriptions were quenched with DNase I treatment, and observed on a 4% agarose sodium borate gel, stained with Sybr Gold nucleic acid dye. Target transcript bands were purified by gel excision and electro-eluted into 3.5kDa MWCO dialysis tubing. RNA was concentrated with a 3.5KDa MWCO concentrating spin column (Millipore) or Speed Vac, and cleaned with acidified phenol:chloroform:isoamyl alcohol (Ambion). RNA was then desalted into 5mM Tris (pH 8.0) using a 7kDa MWCO desalting spin column (Zeba, Thermo Scientific), and quantified by UV-absorbance at 260nm and fluorescence (Qubit HS RNA kit; Invitrogen), and stored at -80°C. Prior to experimental use of frozen stocks, RNA was folded following a basic protocol: dilution into reaction buffer, then successive thermocycler incubations at 95°C (20sec), 50°C (10sec), 37°C (10sec), addition of DFHBI, 4°C (≥30min).

**Fluorescence:** Single sample fluorescent measurements were performed on a Photon Technologies International QM-1 steady state fluorescent spectrophotometer. All fluorescent measurements were performed with the following solution with various cation additions: 0.5µM RNA, 20mM Tris (pH 8.0), 5mM $MgCl_2$, 20µM DFHBI. Samples were read at 20°C, held constant by a circulating water bath, and excited at 468nm in a 2mm quartz fluorescence cuvette. KCl titrations were performed by folding RNA in the presence of various KCl concentrations in individual reaction tubes. Samples were then incubated on ice for 2hr to ensure equilibrium was reached.

*RNaseA-dependent stability:* 0.5µM RNA was incubated with 1.25mU/µl purified RNase A (Qiagen). At stated time points, 5µl of reaction was removed, quenched with 10mM EDTA in 2X gel loading dye, and frozen at -20°C. Samples were analyzed via 4% agarose gel stained with Sybr Gold nucleic acid dye.

*pH-dependent stability:* 0.5µM Chimera was assayed for fluorescent yield at 9 pH points in a universal buffer composed of 20mM acetate, 20mM HEPES, 20mM Borate, 100mM KCl, 5mM $MgCl_2$, 20mM DFHBI. Samples were incubated at stated pH for 30minutes, then analyzed simultaneously in a SpectroMax M2$^e$ Microplate reader (Molecular Devices), Ex/Em at 468/501nm, respectively.

*Serum-dependent Stability:* 0.5µM Chimera was assayed for fluorescence over time in increasing concentrations of cell culture serum, 0, 10, 25, and 50%. Serum was composed of a 1:1 mixture of Bovine Growth Serum and Fatal Calf Serum (Atlanta Biologicals). Fluorescence was measured in a SpectroMax M2$^e$ Microplate reader (Molecular Devices). Samples were Top-read at 25°C, every 60 seconds for 60min with Ex/Em at 468/501nm. Data plotted in GraphPad Prism 5.

**Tripartite complex association kinetics:** Kinetic fluorescence was performed by rapid, in-cuvette mixing of folded Apo-state 0.25µM RNA in buffer with a 10% volume of 10X KCl in water. Emission was read at 501nm every 29.75 seconds with a 0.25 second shutter exposure. Association curves of individual KCl concentrations were fit in GraphPad Prism 5 using the following built in equation for specific binding:

$$Y = [(B_{max})(X)]/(K_D+X) \quad (1)$$

This fit assumes any non-specific binding does not contribute to fluorescence. Individual kinetic fits provided K observed ($K_{obs}$) rate values. To determine the individual kinetic parameters of the $K_{obs}$ (binding and unbinding, or $K_{on}$ and $K_{off}$, respectively), $K_{obs}$ values for each KCl concentration were plotted against their respective KCl concentration. Linear regression of the resulting plot provides $K_{on}$ and $K_{off}$, via the slope and y-intercept, respectively.

**Circular Dichroism:** Measurements were collected with an Aviv 215 CD Spectrophotometer at 20°C. RNA samples were assayed in 10mM Cacodylate (pH 7.4), 100mM KCl, 5mM $MgCl_2$. RNA was folded as described. 135ng/µl RNA samples were scanned in 1mm quartz CD cuvette from 320 to 190nm by a 0.2nm step, with a 3 second integrated read per step. Data presented as average of three individual scans. Control scans of buffer were subtracted from spectra. Values plotted in GraphPad Prism 5.

**Construct Design:** Complementary 3' extensions were engineered into Broccoli construct by PCR. The two sequences added to the 3' terminus of dsDNA T7 Broccoli constructs: (1) 5'-**AGACAGACAGT**-3' and (2) 5'-**ACTGTCTGTCT**-3'.

**Cell Culture:** HEK293T cells were cultured in DMEM with 10% fetal bovine serum, penicillin and streptomycin, and supplemented 5mM glutamine, at 37°C with 5% CO2. Cells were maintained between 40-80% confluent between passages.

**Hypotonic Lysis:** HEK293T cells were cultured to various degrees of confluence in 96-well, clear-bottom, black-walled, tissue culture plates. Cells were quickly and gently washed once with 1X PBS then once with hypotonic buffer (20mM Tris pH 8.0, 5mM $MgCl_2$, 0.01% Triton X-100), before 200ul of hypotonic lysis solution containing 3µM DFHBI, 0.5µM Broccoli chimera RNA, and 80mU/µl murine RNase Inhibitor was added to each well. Plate immediately placed in SpectroMax M2$^e$ Microplate reader (Molecular Devices). Samples were bottom-read at 25°C, every 60 seconds for 60min with Ex/Em at 468/501nm. Positive control wells of varying KCl concentrations with no cells were simultaneously read, as were negative control wells containing either no cells, no RNA, or neither of both. Positive control samples used fit to a simple nonlinear regression model in GraphPad Prism 5:

$$Y = Y_{intercept} + X*Slope \quad (2)$$

Unknown samples values were interpolated from the resulting plot in GraphPad Prism 5.

**Molecular Dynamic Simulations:** Broccoli structural model adapted from Savage et al, 2020.[12] 2'fluorinated Broccoli structural model was created in the YASARA software package as described.[12,13] Simulations were performed for 50ns at 298K with an improved Berendsen thermostat for temperature control and a "densostat" for pressure control, with an integration time step of 2.5fs. Ionic concentration was 0.9% w/v NaCl concentration; electrostatics were evaluated using Particle Mesh Ewald with an 8Å cut off for long range interactions.[14] A modified AMBER14 force field for RNA was employed.[15]

**Model Analysis:** *G-Quadruplex bonding analysis:* The distances between hydrogen bond donors and acceptors for each canonical hydrogen bond within the two-plane G-quadruplex motif was analyzed for each snapshot and plotted as a time trace. Distances within 2Å were considered bound via hydrogen

bonding, and interaction distances outside this range were considered unbound.

*Principle Component Analysis:* The core residues of the G-quadruplex and DHFBI binding pocket were extracted for every simulation snapshot and aligned to the starting broccoli model, to remove any translations and rotations that occurred over the course of the simulation. The MD trajectories for the aptamer with and without $K^+$ were combined to form a single trajectory upon which a principal component analysis was conducted using ProDy.[16] Cluster analysis was conducted on the projections of all snapshots onto the first three principal components using Wolfram Mathematica FindCluster. The structures corresponding to cluster averages were then back-calculated using the coordinates of individual elements within the cluster in order to find potential intermediate states.

## References:


1 J. Ning and Y. Tian, *Sensors Actuators, B Chem.*, 2020, **307**, 127659.
2 X. Kong, F. Su, L. Zhang, J. Yaron, F. Lee, Z. Shi, Y. Tian and D. R. Meldrum, *Angew. Chemie - Int. Ed.*, 2015, **54**, 12053–12057.
3 H. He, M. A. Mortellaro, M. J. P. Leiner, R. J. Fraatz and J. K. Tusa, *J. Am. Chem. Soc.*, 2003, **125**, 1468–1469.
4 P. Padmawar, X. Yao, O. Bloch, G. T. Manley and A. S. Verkman, *Nat. Methods*, 2005, **2**, 825–827.
5 M. Magzoub, P. Padmawar, J. A. Dix and A. S. Verkman, *J. Phys. Chem. B*, 2006, **110**, 21216–21221.
6 A. Minta, R. Y. Tsien, *J Biol Chem.*, 1989, **264**, 19449–19457.
7 H. Szmacinski and J. R. Lakowicz, *Sensors Actuators, B Chem.*, 1999, **60**, 8–18.
8 H. Bischof, M. Rehberg, S. Stryeck, K. Artinger, E. Eroglu, M. Waldeck-Weiermair, B. Gottschalk, R. Rost, A. T. Deak, T. Niedrist, N. Vujic, H. Lindermuth, R. Prassl, B. Pelzmann, K. Groschner, D. Kratky, K. Eller, A. R. Rosenkranz, T. Madl, N. Plesnila, W. F. Graier and R. Malli, *Nat. Commun.*, 2017, **8**, 1–12.
9 Y. Shen, S. Y. Wu, V. Rancic, A. Aggarwal, Y. Qian, S. I. Miyashita, K. Ballanyi, R. E. Campbell and M. Dong, *Commun. Biol.*, 2019, **4**, 18.
10 H. Ueyama, M. Takagi and S. Takenaka, *J. Am. Chem. Soc.*, 2002, **124**, 14286–14287.
11 S. Nagatoishi, T. Nojima, B. Juskowiak and S. Takenaka, *Angew. Chemie*, 2005, **117**, 5195–5198.
12 J. C. Savage, M. A. Davare and U. Shinde, *Chem. Commun.*, 2020, **56**, 2634–2637.
13 J. C. Savage, P. Shinde, H. P. Bächinger, M. A. Davare and U. Shinde, *Chem. Commun.*, 2019, **55**, 5882–5885.
14 T. Darden, L. Perera, L. Li and P. Lee, *Structure*, 1999, **7**, R55–R60.
15 D. Tan, S. Piana, R. M. Dirks and D. E. Shaw, *Proc. Natl. Acad. Sci.*, 2018, **115**, E1346–E1355.
16 A. Bakan, L. M. Meireles and I. Bahar, *Bioinformatics*, 2011, **27**, 1575–1577.

# 8

# References


(1) Gilbert, W. Origins of Life: The RNA World. *Nature* **1986**, *319*, 618.

(2) White, H. B. Coenzymes as Fossils of an Earlier Metabolic State. *J. Mol. Evol.* **1976**, *7* (2), 101–104. https://doi.org/10.1007/BF01732468.

(3) Joyce, G. F. Amplification, Mutation and Selection of Catalytic RNA. *Gene* **1989**, *82* (1), 83–87. https://doi.org/10.1016/0378-1119(89)90033-4.

(4) Robertson, D. L.; Joyce, G. F. Selection in Vitro of an RNA Enzyme That Specifically Cleaves Single-Stranded DNA. *Nature* **1990**, *344* (6265), 467–468. https://doi.org/10.1038/344467a0.

(5) Ellington, A. D.; Szostak, J. W. In Vitro Selection of RNA Molecules That Bind Specific Ligands. *Nature* **1990**, *346* (6287), 818–822. https://doi.org/10.1038/346818a0.

(6) Tuerk, C.; Gold, L. Systematic Evolution of Ligands by Exponential Enrichment: RNA Ligands to Bacteriophage T4 DNA Polymerase. *Science* **1990**, *249* (4968), 505–510.

(7) Schmidt, M. Xenobiology: A New Form of Life as the Ultimate Biosafety Tool. *BioEssays*. **2010**, pp 322–331. https://doi.org/10.1002/bies.200900147.

(8) Pinheiro, V. B.; Taylor, A. I.; Cozens, C.; Abramov, M.; Renders, M.; Zhang, S.; Chaput, J. C.; Wengel, J.; Peak-Chew, S. Y.; McLaughlin, S. H.; Herdewijn, P.; Holliger, P. Synthetic Genetic Polymers Capable of Heredity and Evolution. *Science (80-. ).* **2012**, *336* (6079), 341–344. https://doi.org/10.1126/science.1217622.

(9) Smith, G. P. Filamentous Fusion Phage: Novel Expression Vectors That Display Cloned Antigens on the Virion Surface. *Science (80-. ).* **1985**, *228* (4705), 1315–1317. https://doi.org/10.1126/science.4001944.

(10) Smith, G. P.; Petrenko, V. A. Phage Display. *Chem. Rev.* **1997**, *97* (2), 391–410. https://doi.org/10.1021/cr960065d.

(11) Lorsch, J. R.; Szostak, J. W. Chance and Necessity in the Selection of Nucleic Acid Catalysts. *Accounts of chemical research*. Acc Chem Res **1996**, pp 103–110. https://doi.org/10.1021/ar9501378.

(12) Chen, L.; Rashid, F.; Shah, A.; Awan, H. M.; Wu, M.; Liu, A.; Wang, J.; Zhu, T.; Luo, Z.; Shan, G. The Isolation of an RNA Aptamer Targeting to P53 Protein with Single Amino Acid Mutation. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (32), 10002–10007. https://doi.org/10.1073/pnas.1502159112.

(13) Gelinas, A. D.; Davies, D. R.; Janjic, N. Embracing Proteins: Structural Themes in Aptamer-Protein Complexes. *Current Opinion in Structural Biology*. Elsevier Ltd February 1, **2016**, pp 122–132.

(14) Joyce, G. F. Directed Evolution of Nucleic Acid Enzymes. *Annual Review of Biochemistry*. Annu Rev Biochem **2004**, pp 791–836. https://doi.org/10.1146/annurev.biochem.73.011303.073717.

(15) Wilson, D. S.; Szostak, J. W. In Vitro Selection of Functional Nucleic Acids. *Annu. Rev. Biochem.* **1999**, *68* (1), 611–647. https://doi.org/10.1146/annurev.biochem.68.1.611.

(16) Keefe, A. D.; Pai, S.; Ellington, A. Aptamers as Therapeutics. *Nat. Rev. Drug Discov.* **2010**, *9* (7), 537–550. https://doi.org/10.1038/nrd3141.

(17) Cho, E. J.; Lee, J. W.; Ellington, A. D. Applications of Aptamers as Sensors. *Annu. Rev. Anal. Chem.* **2009**, *2*, 241–264. https://doi.org/10.1146/annurev.anchem.1.031207.112851.

(18) Darmostuk, M.; Rimpelova, S.; Gbelcova, H.; Ruml, T. Current Approaches in SELEX: An Update to Aptamer Selection Technology. *Biotechnol. Adv.* **2015**, *33* (6), 1141–1161. https://doi.org/10.1016/j.biotechadv.2015.02.008.

(19) Nimjee, S. M.; White, R. R.; Becker, R. C.; Sullenger, B. A. Aptamers as Therapeutics. *Annual Review of Pharmacology and Toxicology*. Annual Reviews Inc. January 6, **2017**, pp 61–79. https://doi.org/10.1146/annurev-pharmtox-010716-104558.

(20) Ferrada, E.; Wagner, A. A Comparison of Genotype-Phenotype Maps for RNA and Proteins. *Biophys. J.* **2012**, *102* (8), 1916–1925. https://doi.org/10.1016/j.bpj.2012.01.047.

(21) Fontana, W.; Konings, D. A. M.; Stadler, P. F.; Schuster, P. Statistics of RNA Secondary Structures. *Biopolymers* **1993**, *33* (9), 1389–1404. https://doi.org/10.1002/bip.360330909.

(22) Breaker, R. R. In Vitro Selection of Catalytic Polynucleotides. *Chem. Rev.* **1997**, *97* (2), 371–390. https://doi.org/10.1021/cr960008k.

(23) Prudent, J. R.; Uno, T.; Schultz, P. G. Expanding the Scope of RNA Catalysis.

*Science (80-. ).* **1994**, *264* (5167), 1924–1927. https://doi.org/10.1126/science.8009223.

(24) Breaker, R. R.; Joyce, G. F. A DNA Enzyme That Cleaves RNA. *Chem. Biol.* **1994**, *1* (4), 223–229. https://doi.org/10.1016/1074-5521(94)90014-0.

(25) Illangasekare, M.; Sanchez, G.; Nickles, T.; Yarus, M. Aminoacyl-RNA Synthesis Catalyzed by an RNA. *Science (80-. ).* **1995**, *267* (5198), 643–647. https://doi.org/10.1126/science.7530860.

(26) Saito, H.; Kourouklis, D.; Suga, H. An in Vitro Evolved Precursor TRNA with Aminoacylation Activity. *EMBO J.* **2001**, *20* (7), 1797–1806. https://doi.org/10.1093/emboj/20.7.1797.

(27) Dai, X.; De Mesmaeker, A.; Joyce, G. F. Cleavage of an Amide Bond by a Ribozyme. *Science (80-. ).* **1995**, *267* (5195), 237–240. https://doi.org/10.1126/science.7809628.

(28) Lohse, P. A.; Szostak, J. W. Ribozyme-Catalysed Amino-Acid Transfer Reactions. *Nature* **1996**, *381* (6581), 442–444. https://doi.org/10.1038/381442a0.

(29) Conn, M. M.; Prudent, J. R.; Schultz, P. G. Porphyrin Metalation Catalyzed by a Small RNA Molcule. *J. Am. Chem. Soc.* **1996**, *118* (29), 7012–7013. https://doi.org/10.1021/ja961249c.

(30) Zhang, B.; Cech, T. R. Peptide Bond Formation by in Vitro Selected Ribozymes. *Nature* **1997**, *390* (6655), 96–100. https://doi.org/10.1038/36375.

(31) Seelig, B.; Jäschke, A. A Small Catalytic RNA Motif with Diels-Alderase Activity. *Chem. Biol.* **1999**, *6* (3), 167–176. https://doi.org/10.1016/S1074-5521(99)89008-5.

(32) Fusz, S.; Eisenführ, A.; Srivatsan, S. G.; Heckel, A.; Famulok, M. A Ribozyme for the Aldol Reaction. *Chem. Biol.* **2005**, *12* (8), 941–950. https://doi.org/10.1016/j.chembiol.2005.06.008.

(33) Chumachenko, N. V.; Novikov, Y.; Yarus, M. Rapid and Simple Ribozymic Aminoacylation Using Three Conserved Nucleotides. *J. Am. Chem. Soc.* **2009**, *131* (14), 5257–5263. https://doi.org/10.1021/ja809419f.

(34) Turk, R. M.; Chumachenko, N. V.; Yarus, M. Multiple Translational Products from a Five-Nucleotide Ribozyme. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (10), 4585–4589. https://doi.org/10.1073/pnas.0912895107.

(35) Wolk, S. K.; Mayfield, W. S.; Gelinas, A. D.; Astling, D.; Guillot, J.; Brody, E. N.; Janjic, N.; Gold, L. Modified Nucleotides May Have Enhanced Early RNA Catalysis. *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences April 14, **2020**, pp 8236–8242. https://doi.org/10.1073/pnas.1809041117.

(36) Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R. Self-Splicing RNA: Autoexcision and Autocyclization of the Ribosomal RNA Intervening Sequence of Tetrahymena. *Cell* **1982**, *31* (1), 147–157. https://doi.org/10.1016/0092-8674(82)90414-7.

(37) Ward, W. L.; Plakos, K.; Derose, V. J. Nucleic Acid Catalysis: Metals, Nucleobases, and Other Cofactors. *Chemical Reviews*. American Chemical Society April 23, **2014**, pp 4318–4342. https://doi.org/10.1021/cr400476k.

(38) Micura, R.; Höbartner, C. Fundamental Studies of Functional Nucleic Acids: Aptamers, Riboswitches, Ribozymes and DNAzymes. *Chem. Soc. Rev.* **2020**, *49* (20), 7331. https://doi.org/10.1039/d0cs00617c.

(39) Jenison, R. D.; Gill, S. C.; Pardi, A.; Polisky, B. High-Resolution Molecular Discrimination by RNA. *Science (80-. ).* **1994**, *263* (5152), 1425–1429. https://doi.org/10.1126/science.7510417.

(40) Gold, L.; Ayers, D.; Bertino, J.; Bock, C.; Bock, A.; Brody, E. N.; Carter, J.; Dalby, A. B.; Eaton, B. E.; Fitzwater, T.; Flather, D.; Forbes, A.; Foreman, T.; Fowler, C.; Gawande, B.; Goss, M.; Gunn, M.; Gupta, S.; Halladay, D.; Heil, J.; Heilig, J.; Hicke, B.; Husar, G.; Janjic, N.; Jarvis, T.; Jennings, S.; Katilius, E.; Keeney, T. R.; Kim, N.; Koch, T. H.; Kraemer, S.; Kroiss, L.; Le, N.; Levine, D.; Lindsey, W.; Lollo, B.; Mayfield, W.; Mehan, M.; Mehler, R.; Nelson, S. K.; Nelson, M.; Nieuwlandt, D.; Nikrad, M.; Ochsner, U.; Ostroff, R. M.; Otis, M.; Parker, T.; Pietrasiewicz, S.; Resnicow, D. I.; Rohloff, J.; Sanders, G.; Sattin, S.; Schneider, D.; Singer, B.; Stanton, M.; Sterkel, A.; Stewart, A.; Stratford, S.; Vaught, J. D.; Vrkljan, M.; Walker, J. J.; Watrobka, M.; Waugh, S.; Weiss, A.; Wilcox, S. K.; Wolfson, A.; Wolk, S. K.; Zhang, C.; Zichi, D. Aptamer-Based Multiplexed Proteomic Technology for Biomarker Discovery. *PLoS One* **2010**, *5* (12). https://doi.org/10.1371/journal.pone.0015004.

(41) Nissen, S. B.; Magidson, T.; Gross, K.; Bergstrom, C. T. Publication Bias and the Canonization of False Facts. *Elife* **2016**, *5*. https://doi.org/10.7554/eLife.21451.

(42) Lee, J. F.; Hesselberth, J. R.; Meyers, L. A.; Ellington, A. D. Aptamer Database. *Nucleic Acids Res.* **2004**, *32* (DATABASE ISS.), D95–D100. https://doi.org/10.1093/nar/gkh094.

(43) Debiais, M.; Lelievre, A.; Smietana, M.; Müller, S. Splitting Aptamers and Nucleic Acid Enzymes for the Development of Advanced Biosensors. *Nucleic Acids Research*. Oxford University Press April 17, **2020**, pp 3400–3422. https://doi.org/10.1093/nar/gkaa132.

(44) Schoukroun-Barnes, L. R.; Macazo, F. C.; Gutierrez, B.; Lottermoser, J.; Liu, J.; White, R. J. Reagentless, Structure-Switching, Electrochemical Aptamer-Based Sensors. *Annual Review of Analytical Chemistry*. Annual Reviews Inc. June 12, **2016**, pp 163–181. https://doi.org/10.1146/annurev-anchem-071015-041446.

(45) Jhaveri, S.; Rajendran, M.; Ellington, A. D. In Vitro Selection of Signaling Aptamers. *Nat. Biotechnol.* **2000**, *18* (12), 1293–1297. https://doi.org/10.1038/82414.

(46) Torimura, M.; Kurata, S.; Yamada, K.; Yokomaku, T.; Kamagata, Y.; Kanagawa, T.; Kurane, R. Fluorescence-Quenching Phenomenon by Photoinduced Electron Transfer between a Fluorescent Dye and a Nucleotide Base. *Anal. Sci.* **2001**, *17* (1), 155–160. https://doi.org/10.2116/analsci.17.155.

(47) Merino, E. J.; Weeks, K. M. Facile Conversion of Aptamers into Sensors Using a 2'-Ribose-Linked Fluorophore. *J. Am. Chem. Soc.* **2005**, *127* (37), 12766–12767. https://doi.org/10.1021/ja053189t.

(48) Katilius, E.; Katiliene, Z.; Woodbury, N. W. Signaling Aptamers Created Using Fluorescent Nucleotide Analogues. *Anal. Chem.* **2006**, *78* (18), 6484–6489. https://doi.org/10.1021/ac060859k.

(49) Jhaveri, S. D.; Kirby, R.; Conrad, R.; Maglott, E. J.; Bowser, M.; Kennedy, R. T.; Glick, G.; Ellington, A. D. Designed Signaling Aptamers That Transduce Molecular Recognition to Changes in Fluorescence Intensity. *J. Am. Chem. Soc.* **2000**, *122* (11), 2469–2473. https://doi.org/10.1021/ja992393b.

(50) Rajendran, M.; Ellington, A. D. In Vitro Selection of Molecular Beacons. *Nucleic Acids Res.* **2003**, *31* (19), 5700–5713. https://doi.org/10.1093/nar/gkg764.

(51) Yamana, K.; Iwai, T.; Ohtani, Y.; Sato, S.; Nakamura, M.; Nakano, H. Bis-Pyrene-Labeled Oligonucleotides: Sequence Specificity of Excimer and Monomer Fluorescence Changes upon Hybridization with DNA. *Bioconjug. Chem.* **2002**, *13* (6), 1266–1273. https://doi.org/10.1021/bc025530u.

(52) Stojanovic, M. N.; de Prada, P.; Landry, D. W. Aptamer-Based Folding Fluorescent Sensor for Cocaine. *J. Am. Chem. Soc.* **2001**, *123* (21), 4928–4931. https://doi.org/10.1021/ja0038171.

(53) Li, J. J.; Fang, X.; Tan, W. Molecular Aptamer Beacons for Real-Time Protein Recognition. *Biochem. Biophys. Res. Commun.* **2002**, *292* (1), 31–40. https://doi.org/10.1006/bbrc.2002.6581.

(54) Hamaguchi, N.; Ellington, A.; Stanton, M. Aptamer Beacons for the Direct Detection of Proteins. *Anal. Biochem.* **2001**, *294* (2), 126–131. https://doi.org/10.1006/abio.2001.5169.

(55) Fang, X.; Sen, A.; Vicens, M.; Tan, W. Synthetic DNA Aptamers to Detect Protein Molecular Variants in a High-Throughput Fluorescence Quenching Assay. *ChemBioChem* **2003**, *4* (9), 829–834. https://doi.org/10.1002/cbic.200300615.

(56) Nutiu, R.; Li, Y. Structure-Switching Signaling Aptamers. *J. Am. Chem. Soc.* **2003**, *125* (16), 4771–4778. https://doi.org/10.1021/ja028962o.

(57) Rajendran, M.; Ellington, A. D. Selection of Fluorescent Aptamer Beacons That

Light up in the Presence of Zinc. *Anal. Bioanal. Chem.* **2008**, *390* (4), 1067–1075. https://doi.org/10.1007/s00216-007-1735-8.

(58) Nutiu, R.; Li, Y. In Vitro Selection of Structure-Switching Signaling Aptamers. *Angew. Chemie Int. Ed.* **2005**, *44* (7), 1061–1065. https://doi.org/10.1002/anie.200461848.

(59) Kent, A. D.; Spiropulos, N. G.; Heemstra, J. M. General Approach for Engineering Small-Molecule-Binding DNA Split Aptamers. *Anal. Chem.* **2013**, *85* (20), 9916–9923. https://doi.org/10.1021/ac402500n.

(60) Chen, A.; Yan, M.; Yang, S. Split Aptamers and Their Applications in Sandwich Aptasensors. *TrAC - Trends in Analytical Chemistry*. Elsevier B.V. June 1, **2016**, pp 581–593. https://doi.org/10.1016/j.trac.2016.04.006.

(61) Levy, M.; Cater, S. F.; Ellington, A. D. Quantum-Dot Aptamer Beacons for the Detection of Proteins. *ChemBioChem* **2005**, *6* (12), 2163–2166. https://doi.org/10.1002/cbic.200500218.

(62) Banerjee, A.; Grazon, C.; Nadal, B.; Pons, T.; Krishnan, Y.; Dubertret, B. Fast, Efficient, and Stable Conjugation of Multiple DNA Strands on Colloidal Quantum Dots. *Bioconjug. Chem.* **2015**, *26* (8), 1582–1589. https://doi.org/10.1021/acs.bioconjchem.5b00221.

(63) Chu, T. C.; Shieh, F.; Lavery, L. A.; Levy, M.; Richards-Kortum, R.; Korgel, B. A.; Ellington, A. D. Labeling Tumor Cells with Fluorescent Nanocrystal-Aptamer Bioconjugates. *Biosens. Bioelectron.* **2006**, *21*, 1859–1866. https://doi.org/10.1016/j.bios.2005.12.015.

(64) Ikanovic, M.; Rudzinski, W. E.; Bruno, J. G.; Allman, A.; Carrillo, M. P.; Dwarakanath, S.; Bhahdigadi, S.; Rao, P.; Kiel, J. L.; Andrews, C. J. Fluorescence Assay Based on Aptamer-Quantum Dot Binding to Bacillus Thuringiensis Spores. *J. Fluoresc.* **2007**, *17* (2), 193–199. https://doi.org/10.1007/s10895-007-0158-4.

(65) Choi, J. H.; Chen, K. H.; Strano, M. S. Aptamer-Capped Nanocrystal Quantum Dots: A New Method for Label-Free Protein Detection. *J. Am. Chem. Soc.* **2006**, *128* (49), 15584–15585. https://doi.org/10.1021/ja066506k.

(66) Liu, J.; Lu, Y. Fast Colorimetric Sensing of Adenosine and Cocaine Based on a General Sensor Design Involving Aptamers and Nanoparticles. *Angew. Chemie Int. Ed.* **2006**, *45* (1), 90–94. https://doi.org/10.1002/anie.200502589.

(67) Zhao, W.; Chiuman, W.; Brook, M. A.; Li, Y. Simple and Rapid Colorimetric Biosensors Based on DNA Aptamer and Noncrosslinking Gold Nanoparticle Aggregation. *ChemBioChem* **2007**, *8* (7), 727–731. https://doi.org/10.1002/cbic.200700014.

(68) Wang, W.; Chen, C.; Qian, M.; Zhao, X. S. Aptamer Biosensor for Protein Detection Using Gold Nanoparticles. *Anal. Biochem.* **2008**, *373* (2), 213–219.

https://doi.org/10.1016/j.ab.2007.11.013.

(69) Sano, T.; Smith, C. L.; Cantor, C. R. Immuno-PCR: Very Sensitive Antigen Detection by Means of Specific Antibody-DNA Conjugates. *Science (80-. ).* **1992**, *258* (5079), 120–122. https://doi.org/10.1126/science.1439758.

(70) Fischer, N. O.; Tarasow, T. M.; Tok, J. B. H. Protein Detection via Direct Enzymatic Amplification of Short DNA Aptamers. *Anal. Biochem.* **2008**, *373* (1), 121–128. https://doi.org/10.1016/j.ab.2007.09.035.

(71) Niemeyer, C. M.; Adler, M.; Wacker, R. Detecting Antigens by Quantitative Immuno-PCR. *Nat. Protoc.* **2007**, *2* (8), 1918–1930. https://doi.org/10.1038/nprot.2007.267.

(72) Yoshida, Y.; Horii, K.; Sakai, N.; Masuda, H.; Furuichi, M.; Waga, I. Antibody-Specific Aptamer-Based PCR Analysis for Sensitive Protein Detection. *Anal. Bioanal. Chem.* **2009**, *395* (4), 1089–1096. https://doi.org/10.1007/s00216-009-3041-0.

(73) Gullberg, M.; Fredriksson, S.; Taussig, M.; Jarvius, J.; Gustafsdottir, S.; Landegren, U. A Sense of Closeness: Protein Detection by Proximity Ligation. *Current Opinion in Biotechnology*. Elsevier Ltd **2003**, pp 82–86. https://doi.org/10.1016/S0958-1669(02)00011-3.

(74) Gustafsdottir, S. M.; Schallmeiner, E.; Fredriksson, S.; Gullberg, M.; Söderberg, O.; Jarvius, M.; Jarvius, J.; Howell, M.; Landegren, U. Proximity Ligation Assays for Sensitive and Specific Protein Analyses. In *Analytical Biochemistry*; Academic Press Inc., **2005**; Vol. 345, pp 2–9. https://doi.org/10.1016/j.ab.2005.01.018.

(75) Landegren, U.; Schallmeiner, E.; Nilsson, M.; Fredriksson, S.; Banér, J.; Gullberg, M.; Jarvius, J.; Gustafsdottir, S.; Dahl, F.; Söderberg, O.; Ericsson, O.; Stenberg, J. Molecular Tools for a Molecular Medicine: Analyzing Genes, Transcripts and Proteins Using Padlock and Proximity Probes. In *Journal of Molecular Recognition*; J Mol Recognit, **2004**; Vol. 17, pp 194–197. https://doi.org/10.1002/jmr.664.

(76) Söderberg, O.; Leuchowius, K. J.; Gullberg, M.; Jarvius, M.; Weibrecht, I.; Larsson, L. G.; Landegren, U. Characterizing Proteins and Their Interactions in Cells and Tissues Using the in Situ Proximity Ligation Assay. *Methods* **2008**, *45* (3), 227–232. https://doi.org/10.1016/j.ymeth.2008.06.014.

(77) Gustafsdottir, S. M.; Nordengrahn, A.; Fredriksson, S.; Wallgren, P.; Rivera, E.; Schallmeiner, E.; Merza, M.; Landegren, U. Detection of Individual Microbial Pathogens by Proximity Ligation. *Clin. Chem.* **2006**, *52* (6), 1152–1160. https://doi.org/10.1373/clinchem.2005.065847.

(78) Pai, S. S.; Ellington, A. D. Using RNA Aptamers and the Proximity Ligation Assay for the Detection of Cell Surface Antigens. *Methods Mol. Biol.* **2009**, *504*, 385–398. https://doi.org/10.1007/978-1-60327-569-9_21.

(79) Fredriksson, S.; Gullberg, M.; Jarvius, J.; Olsson, C.; Pietras, K.; Gústafsdóttir, S. M.; Östman, A.; Landegren, U. Protein Detection Using Proximity-Dependent DNA Ligation Assays. *Nat. Biotechnol.* **2002**, *20* (5), 473–477. https://doi.org/10.1038/nbt0502-473.

(80) Xu, D.; Xu, D.; Yu, X.; Liu, Z.; He, W.; Ma, Z. Label-Free Electrochemical Detection for Aptamer-Based Array Electrodes. *Anal. Chem.* **2005**, *77* (16), 5107–5113. https://doi.org/10.1021/ac050192m.

(81) Ikebukuro, K.; Kiyohara, C.; Sode, K. Novel Electrochemical Sensor System for Protein Using the Aptamers in Sandwich Manner. In *Biosensors and Bioelectronics*; Elsevier Ltd, **2005**; Vol. 20, pp 2168–2172. https://doi.org/10.1016/j.bios.2004.09.002.

(82) Numnuam, A.; Chumbimuni-Torres, K. Y.; Xiang, Y.; Bash, R.; Thavarungkul, P.; Kanatharana, P.; Pretsch, E.; Wang, J.; Bakker, E. Aptamer-Based Potentiometric Measurements of Proteins Using Ion-Selective Microelectrodes. *Anal. Chem.* **2008**, *80* (3), 707–712. https://doi.org/10.1021/ac701910r.

(83) Feng, K.; Sun, C.; Kang, Y.; Chen, J.; Jiang, J. H.; Shen, G. L.; Yu, R. Q. Label-Free Electrochemical Detection of Nanomolar Adenosine Based on Target-Induced Aptamer Displacement. *Electrochem. commun.* **2008**, *10* (4), 531–535. https://doi.org/10.1016/j.elecom.2008.01.024.

(84) Wang, X.; Zhou, J.; Yun, W.; Xiao, S.; Chang, Z.; He, P.; Fang, Y. Detection of Thrombin Using Electrogenerated Chemiluminescence Based on Ru(Bpy)32+-Doped Silica Nanoparticle Aptasensor via Target Protein-Induced Strand Displacement. *Anal. Chim. Acta* **2007**, *598* (2), 242–248. https://doi.org/10.1016/j.aca.2007.07.050.

(85) Lai, R. Y.; Plaxco, K. W.; Heeger, A. J. Aptamer-Based Electrochemical Detection of Picomolar Platelet-Derived Growth Factor Directly in Blood Serum. *Anal. Chem.* **2007**, *79* (1), 229–233. https://doi.org/10.1021/ac061592s.

(86) Swensen, J. S.; Xiao, Y.; Ferguson, B. S.; Lubin, A. A.; Lai, R. Y.; Heeger, A. J.; Plaxco, K. W.; Soh, H. T. Continuous, Real-Time Monitoring of Cocaine in Undiluted Blood Serum via a Microfluidic, Electrochemical Aptamer-Based Sensor. *J. Am. Chem. Soc.* **2009**, *131* (12), 4262–4266. https://doi.org/10.1021/ja806531z.

(87) Lubin, A. A.; Lai, R. Y.; Baker, B. R.; Heeger, A. J.; Plaxco, K. W. Sequence-Specific, Electronic Detection of Oligonucleotides in Blood, Soil, and Foodstuffs with the Reagentless, Reusable E-DNA Sensor. *Anal. Chem.* **2006**, *78* (16), 5671–5677. https://doi.org/10.1021/ac0601819.

(88) Ferguson, B. S.; Hoggarth, D. A.; Maliniak, D.; Ploense, K.; White, R. J.; Woodward, N.; Hsieh, K.; Bonham, A. J.; Eisenstein, M.; Kippin, T. E.; Plaxco, K. W.; Soh, H. T. Real-Time, Aptamer-Based Tracking of Circulating Therapeutic Agents in Living Animals. *Sci. Transl. Med.* **2013**, *5* (213), 213ra165.

https://doi.org/10.1126/scitranslmed.3007095.

(89) Li, Y.; Breaker, R. R. Kinetics of RNA Degradation by Specific Base Catalysis of Transesterification Involving the 2γ-Hydroxyl Group. *J. Am. Chem. Soc.* **1999**, *121* (23), 5364–5372. https://doi.org/10.1021/ja990592p.

(90) Velikyan, I.; Acharya, S.; Trifonova, A.; Földesi, A.; Chattopadhyaya, J. The PKa's of 2'-Hydroxyl Group in Nucleosides and Nucleotides. *Journal of the American Chemical Society*. American Chemical Society **2001**, pp 2893–2894. https://doi.org/10.1021/ja0036312.

(91) Guschlbauer, W.; Jankowski, K. Nucleoside Conformation Is Determined by the Electronegativity of the Sugar Substituent. *Nucleic Acids Res.* **1980**, *8* (6), 1421–1433. https://doi.org/10.1093/nar/8.6.1421.

(92) Gevertz, J.; Gan, H. H.; Schlick, T. In Vitro RNA Random Pools Are Not Structurally Diverse: A Computational Analysis. *RNA* **2005**, *11* (6), 853–863. https://doi.org/10.1261/rna.7271405.

(93) Takahashi, M.; Wu, X.; Ho, M.; Chomchan, P.; Rossi, J. J.; Burnett, J. C.; Zhou, J. High Throughput Sequencing Analysis of RNA Libraries Reveals the Influences of Initial Library and PCR Methods on SELEX Efficiency. *Sci. Rep.* **2016**, *6* (1), 1–14. https://doi.org/10.1038/srep33697.

(94) Thiel, W. H.; Bair, T.; Wyatt Thiel, K.; Dassie, J. P.; Rockey, W. M.; Howell, C. A.; Liu, X. Y.; Dupuy, A. J.; Huang, L.; Owczarzy, R.; Behlke, M. A.; McNamara, J. O.; Giangrande, P. H. Nucleotide Bias Observed with a Short SELEX RNA Aptamer Library. *Nucleic Acid Ther.* **2011**, *21* (4), 253–263. https://doi.org/10.1089/nat.2011.0288.

(95) Stevens, A. J.; Stuffrein-Roberts, S.; Cree, S. L.; Gibb, A.; Miller, A. L.; Doudney, K.; Aitchison, A.; Eccles, M. R.; Joyce, P. R.; Filichev, V. V.; Kennedy, M. A. G-Quadruplex Structures and CpG Methylation Cause Drop-Out of the Maternal Allele in Polymerase Chain Reaction Amplification of the Imprinted MEST Gene Promoter. *PLoS One* **2014**, *9* (12), e113955. https://doi.org/10.1371/journal.pone.0113955.

(96) Latham, J. A.; Johnson, R.; Toole, J. J. The Application of a Modified Nucleotide in Aptamer Selection: Novel Thrombin Aptamers Containing -(1 -Pentynyl)-2'-Deoxyuridine. *Nucleic Acids Res.* **1994**, *22* (14), 2817–2822. https://doi.org/10.1093/nar/22.14.2817.

(97) Sousa, R.; Padilla, R. A Mutant T7 RNA Polymerase as a DNA Polymerase. *EMBO J.* **1995**, *14* (18), 4609–4621.

(98) Padilla, R.; Sousa, R. A Y639F/H784A T7 RNA Polymerase Double Mutant Displays Superior Properties for Synthesizing RNAs with Non-Canonical NTPs. *Nucleic Acids Res.* **2002**, *30* (24), e138. https://doi.org/10.1093/NAR/GNF138.

(99) Meyer, A. J.; Garry, D. J.; Hall, B.; Byrom, M. M.; McDonald, H. G.; Yang, X.; Yin, Y. W.; Ellington, A. D. Transcription Yield of Fully 2'-Modified RNA Can Be Increased by the Addition of Thermostabilizing Mutations to T7 RNA Polymerase Mutants. *Nucleic Acids Res.* **2015**, *43* (15), 7480–7488. https://doi.org/10.1093/nar/gkv734.

(100) Eaton, B. E.; Gold, L.; Hicke, B. J.; Janjić, N.; Jucker, F. M.; Sebesta, D. P.; Tarasow, T. M.; Willis, M. C.; Zichi, D. A. Post-SELEX Combinatorial Optimization of Aptamers. In *Bioorganic and Medicinal Chemistry*; Pergamon, **1997**; Vol. 5, pp 1087–1096. https://doi.org/10.1016/S0968-0896(97)00044-8.

(101) Verma, S.; Eckstein, F. *MODIFIED OLIGONUCLEOTIDES: Synthesis and Strategy for Users*; **1998**; Vol. 67.

(102) Keefe, A. D.; Cload, S. T. SELEX with Modified Nucleotides. *Current Opinion in Chemical Biology*. Curr Opin Chem Biol August **2008**, pp 448–456. https://doi.org/10.1016/j.cbpa.2008.06.028.

(103) Dellafiore, M. A.; Montserrat, J. M.; Iribarren, A. M. Modified Nucleoside Triphosphates for In-Vitro Selection Techniques. *Frontiers in Chemistry*. Frontiers Media S. A May 4, **2016**, p 18. https://doi.org/10.3389/fchem.2016.00018.

(104) Odeh, F.; Nsairat, H.; Alshaer, W.; Ismail, M. A.; Esawi, E.; Qaqish, B.; Bawab, A. Al; Ismail, S. I. Aptamers Chemistry: Chemical Modifications and Conjugation Strategies. *Molecules*. MDPI AG **2020**. https://doi.org/10.3390/molecules25010003.

(105) Beaucage, S. L.; Caruthers, M. H. Deoxynucleoside Phosphoramidites-A New Class of Key Intermediates for Deoxypolynucleotide Synthesis. *Tetrahedron Lett.* **1981**, *22* (20), 1859–1862. https://doi.org/10.1016/S0040-4039(01)90461-7.

(106) Cadwell, R. C.; Joyce, G. F. Randomization of Genes by PCR Mutagenesis. *Genome Res.* **1992**, *2* (1), 28–33. https://doi.org/10.1101/gr.2.1.28.

(107) Singer, B. S.; Shtatland, T.; Brown, D.; Gold, L. Libraries for Genomic SELEX. *Nucleic Acids Res.* **1997**, *25* (4), 781–786. https://doi.org/10.1093/nar/25.4.781.

(108) Lorenz, C.; von Pelchrzim, F.; Schroeder, R. Genomic Systematic Evolution of Ligands by Exponential Enrichment (Genomic SELEX) for the Identification of Protein-Binding RNAs Independent of Their Expression Levels. *Nat. Protoc.* **2006**, *1* (5), 2204–2212. https://doi.org/10.1038/nprot.2006.372.

(109) Dunitz, J. D.; Joyce, G. F. Leslie Eleazer Orgel. 12 January 1927 — 27 October 2007. *Biogr. Mem. Fellows R. Soc.* **2013**, *59*, 277–289. https://doi.org/10.1098/rsbm.2013.0002.

(110) Ollis, D. L.; Brick, P.; Hamlin, R.; Xuong, N. G.; Steitz, T. A. Structure of Large Fragment of Escherichia Coli DNA Polymerase I Complexed with DTMP. *Nature* **1985**, *313* (6005), 762–766. https://doi.org/10.1038/313762a0.

(111) Beard, W. A.; Wilson, S. H. Structural Insights into the Origins of DNA Polymerase Fidelity. *Structure*. Cell Press May 1, **2003**, pp 489–496. https://doi.org/10.1016/S0969-2126(03)00051-0.

(112) Joyce, C. M. Choosing the Right Sugar: How Polymerases Select a Nucleotide Substrate. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94* (5), 1619–1622. https://doi.org/10.1073/pnas.94.5.1619.

(113) Pagratis, N. C.; Bell, C.; Chang, Y. F.; Jennings, S.; Fitzwater, T.; Jellinek, D.; Dang, C. Potent 2′-Amino-, and 2′-Fluoro-2-Deoxyribonucleotide Rna Inhibitors of Keratinocyte Growth Factor. *Nat. Biotechnol.* **1997**, *15* (1), 68–73. https://doi.org/10.1038/nbt0197-68.

(114) Lin, Y.; Qiu, Q.; Gill, S. C.; Jayasena, S. D. Modified RNA Sequence Pools for in Vitro Selection. *Nucleic Acids Res.* **1994**, *22* (24), 5229–5234. https://doi.org/10.1093/nar/22.24.5229.

(115) Codington, J. F.; Doerr, I. L.; Fox, J. J. Nucleosides. XVIII. Synthesis of 2′-Fluorothymidine, 2′-Fluorodeoxyuridine, and Other 2′-Halogeno-2′-Deoxy Nucleosides. *J. Org. Chem.* **1964**, *29* (3), 558–564. https://doi.org/10.1021/jo01026a009.

(116) Janik, B.; Kotick, M. P.; Kreiser, T. H.; Reverman, L. F.; Sommer, R. G.; Wilson, D. P. Synthesis and Properties of Poly 2′-Fluoro-2′-Deoxyuridylic Acid. *Biochem. Biophys. Res. Commun.* **1972**, *46* (3), 1153–1160. https://doi.org/10.1016/S0006-291X(72)80095-0.

(117) Pieken, W. A.; Olsen, D. B.; Benseler, F.; Aurup, H.; Eckstein, F. Kinetic Characterization of Ribonuclease-Resistant 2′-Modified Hammerhead Ribozymes. *Science (80-. ).* **1991**, *253* (5017), 314–317. https://doi.org/10.1126/science.1857967.

(118) Monia, B. P.; Lesnik, E. A.; Gonzalez, C.; Lima, W. F.; McGee, D.; Guinosso, C. J.; Kawasaki, A. M.; Cook, P. D.; Freier, S. M. Evaluation of 2′-Modified Oligonucleotides Containing 2′-Deoxy Gaps as Antisense Inhibitors of Gene Expression. *J. Biol. Chem.* **1993**, *268* (19), 14514–14522. https://doi.org/10.1016/s0021-9258(19)85268-7.

(119) Patra, A.; Paolillo, M.; Charisse, K.; Manoharan, M.; Rozners, E.; Egli, M. 2′-Fluoro RNA Shows Increased Watson-Crick H-Bonding Strength and Stacking Relative to RNA: Evidence from NMR and Thermodynamic Data. *Angew. Chemie - Int. Ed.* **2012**, *51* (47), 11863–11866. https://doi.org/10.1002/anie.201204946.

(120) Pallan, P. S.; Greene, E. M.; Jicman, P. A.; Pandey, R. K.; Manoharan, M.; Rozners, E.; Egli, M. Unexpected Origins of the Enhanced Pairing Affinity of 2′-Fluoro-Modified RNA. *Nucleic Acids Res.* **2011**, *39* (8), 3482–3495. https://doi.org/10.1093/nar/gkq1270.

(121) Manoharan, M.; Akinc, A.; Pandey, R. K.; Qin, J.; Hadwiger, P.; John, M.; Mills, K.; Charisse, K.; Maier, M. A.; Nechev, L.; Greene, E. M.; Pallan, P. S.; Rozners, E.; Rajeev, K. G.; Egli, M. Unique Gene-Silencing and Structural Properties of 2'-Fluoro-Modified SiRNAs. *Angew. Chemie - Int. Ed.* **2011**, *50* (10), 2284–2288. https://doi.org/10.1002/anie.201006519.

(122) Verheyden, J. P. H.; Wagner, D.; Mofpatt, J. G. Synthesis of Some Pyrimidine 2'-Amino-2'-Deoxynucleosides. *J. Org. Chem.* **1971**, *36* (2), 250–254. https://doi.org/10.1021/jo00801a002.

(123) Aurup, H.; Tuschl, T.; Benseler, F.; Ludwig, J.; Eckstein, F. Oligonucleotide Duplexes Containing 2'-Amino-2'-Deoxycytidines: Thermal Stability and Chemical Reactivity. *Nucleic Acids Res.* **1994**, *22* (1), 20–24. https://doi.org/10.1093/nar/22.1.20.

(124) Padilla, R.; Sousa, R. Efficient Synthesis of Nucleic Acids Heavily Modified with Non-Canonical Ribose 2'-Groups Using a Mutant T7 RNA Polymerase (RNAP). *Nucleic Acids Res.* **1999**, *27* (6), 1561–1563. https://doi.org/10.1093/nar/27.6.1561.

(125) Green, L. S.; Jellinek, D.; Bell, C.; Beebe, L. A.; Feistner, B. D.; Gill, S. C.; Jucker, F. M.; Janjić, N. Nuclease-Resistant Nucleic Acid Ligands to Vascular Permeability Factor/Vascular Endothelial Growth Factor. *Chem. Biol.* **1995**, *2* (10), 683–695. https://doi.org/10.1016/1074-5521(95)90032-2.

(126) Suzuki, T. The Expanding World of TRNA Modifications and Their Disease Relevance. *Nature Reviews Molecular Cell Biology*. Nature Research March 3, **2021**, pp 1–18. https://doi.org/10.1038/s41580-021-00342-0.

(127) Kratschmer, C.; Levy, M. Effect of Chemical Modifications on Aptamer Stability in Serum. *Nucleic Acid Ther.* **2017**, *27* (6), 335–344. https://doi.org/10.1089/nat.2017.0680.

(128) Ruckman, J.; Green, L. S.; Beeson, J.; Waugh, S.; Gillette, W. L.; Henninger, D. D.; Claesson-Welsh, L.; Janjic, N. 2'-Fluoropyrimidine RNA-Based Aptamers to the 165-Amino Acid Form of Vascular Endothelial Growth Factor (VEGF165): INHIBITION OF RECEPTOR BINDING AND VEGF-INDUCED VASCULAR PERMEABILITY THROUGH INTERACTIONS REQUIRING THE EXON 7-ENCODED DOMAIN. *J. Biol. Chem.* **1998**, *273* (32), 20556–20567. https://doi.org/10.1074/jbc.273.32.20556.

(129) Ng, E. W. M.; Shima, D. T.; Calias, P.; Cunningham, E. T.; Guyer, D. R.; Adamis, A. P. Pegaptanib, a Targeted Anti-VEGF Aptamer for Ocular Vascular Disease. *Nature Reviews Drug Discovery*. Nat Rev Drug Discov February **2006**, pp 123–132. https://doi.org/10.1038/nrd1955.

(130) Ammar, M. J.; Hsu, J.; Chiang, A.; Ho, A. C.; Regillo, C. D. Age-Related Macular Degeneration Therapy: A Review. *Current Opinion in Ophthalmology*. Lippincott

Williams and Wilkins May 1, **2020**, pp 215–221. https://doi.org/10.1097/ICU.0000000000000657.

(131) Healy, J. M.; Lewis, S. D.; Kurz, M.; Boomer, R. M.; Thompson, K. M.; Wilson, C.; McCauley, T. G. Pharmacokinetics and Biodistribution of Novel Aptamer Compositions. *Pharm. Res.* **2004**, *21* (12), 2234–2246.

(132) Drolet, D. W.; Nelson, J.; Tucker, C. E.; Zack, P. M.; Nixon, K.; Bolin, R.; Judkins, M. B.; Farmer, J. A.; Wolf, J. L.; Gill, S. C.; Bendele, R. A. Pharmacokinetics and Safety of an Anti-Vascular Endothelial Growth Factor Aptamer (NX1838) Following Injection into the Vitreous Humor of Rhesus Monkeys. *Pharm. Res.* **2000**, *17* (12), 1503–1510.

(133) Kraemer, S.; Vaught, J. D.; Bock, C.; Gold, L.; Katilius, E.; Keeney, T. R.; Kim, N.; Saccomano, N. A.; Wilcox, S. K.; Zichi, D.; Sanders, G. M. From SOMAmer-Based Biomarker Discovery to Diagnostic and Clinical Applications: A SOMAmer-Based, Streamlined Multiplex Proteomic Assay. *PLoS One* **2011**, *6* (10), e26332. https://doi.org/10.1371/journal.pone.0026332.

(134) Davies, D. R.; Gelinas, A. D.; Zhang, C.; Rohloff, J. C.; Carter, J. D.; O'Connell, D.; Waugh, S. M.; Wolk, S. K.; Mayfield, W. S.; Burgin, A. B.; Edwards, T. E.; Stewart, L. J.; Gold, L.; Janjic, N.; Jarvis, T. C. Unique Motifs and Hydrophobic Interactions Shape the Binding of Modified DNA Ligands to Protein Targets. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (49), 19971–19976. https://doi.org/10.1073/pnas.1213933109.

(135) Gelinas, A. D.; Davies, D. R.; Edwards, T. E.; Rohloff, J. C.; Carter, J. D.; Zhang, C.; Gupta, S.; Ishikawa, Y.; Hirota, M.; Nakaishi, Y.; Jarvis, T. C.; Janjic, N. Crystal Structure of Interleukin-6 in Complex with a Modified Nucleic Acid Ligand. *J. Biol. Chem.* **2014**, *289* (12), 8720–8734. https://doi.org/10.1074/jbc.M113.532697.

(136) Jarvis, T. C.; Davies, D. R.; Hisaminato, A.; Resnicow, D. I.; Gupta, S.; Waugh, S. M.; Nagabukuro, A.; Wadatsu, T.; Hishigaki, H.; Gawande, B.; Zhang, C.; Wolk, S. K.; Mayfield, W. S.; Nakaishi, Y.; Burgin, A. B.; Stewart, L. J.; Edwards, T. E.; Gelinas, A. D.; Schneider, D. J.; Janjic, N. Non-Helical DNA Triplex Forms a Unique Aptamer Scaffold for High Affinity Recognition of Nerve Growth Factor. *Structure* **2015**, *23* (7), 1293–1304. https://doi.org/10.1016/j.str.2015.03.027.

(137) Rohloff, J. C.; Gelinas, A. D.; Jarvis, T. C.; Ochsner, U. A.; Schneider, D. J.; Gold, L.; Janjic, N. Nucleic Acid Ligands with Protein-like Side Chains: Modified Aptamers and Their Use as Diagnostic and Therapeutic Agents. *Molecular Therapy - Nucleic Acids*. Nature Publishing Group January 1, **2014**, p e201. https://doi.org/10.1038/mtna.2014.49.

(138) Mian, I. S.; Bradwell, A. R.; Olson, A. J. Structure, Function and Properties of Antibody Binding Sites. *J. Mol. Biol.* **1991**, *217* (1), 133–151. https://doi.org/10.1016/0022-2836(91)90617-F.

(139) Ramaraj, T.; Angel, T.; Dratz, E. A.; Jesaitis, A. J.; Mumey, B. Antigen-Antibody Interface Properties: Composition, Residue Interactions, and Features of 53 Non-Redundant Structures. *Biochim. Biophys. Acta - Proteins Proteomics* **2012**, *1824* (3), 520–532. https://doi.org/10.1016/j.bbapap.2011.12.007.

(140) Tonegawa, S.; Steinberg, C.; Dube, S.; Bernardini, A. Evidence for Somatic Generation of Antibody Diversity. *Proc. Natl. Acad. Sci. U. S. A.* **1974**, *71* (10), 4027–4031. https://doi.org/10.1073/pnas.71.10.4027.

(141) Hozumi, N.; Tonegawa, S. Evidence for Somatic Rearrangement of Immunoglobulin Genes Coding for Variable and Constant Regions. *Proc. Natl. Acad. Sci. U. S. A.* **1976**, *73* (10), 3628–3632. https://doi.org/10.1073/pnas.73.10.3628.

(142) Jung, D.; Alt, F. W. Unraveling V(D)J Recombination: Insights into Gene Regulation. *Cell*. Elsevier B.V. January 23, **2004**, pp 299–311. https://doi.org/10.1016/S0092-8674(04)00039-X.

(143) Wolk, S. K.; Shoemaker, R. K.; Mayfield, W. S.; Mestdagh, A. L.; Janjic, N. Influence of 5-N-Carboxamide Modifications on the Thermodynamic Stability of Oligonucleotides. *Nucleic Acids Res.* **2015**, *43* (19), 9107–9122. https://doi.org/10.1093/nar/gkv981.

(144) Editorial: The 18th Annual Nucleic Acids Research Web Server Issue 2020. *Nucleic acids research*. NLM (Medline) July 2, **2020**, pp W1–W4. https://doi.org/10.1093/nar/gkaa528.

(145) Davenport, D.; Nicol, J. A. C. Luminescence in Hydromedusae. *Proc. R. Soc. London. Ser. B - Biol. Sci.* **1955**, *144* (916), 399–411. https://doi.org/10.1098/rspb.1955.0066.

(146) Shimomura, O.; Johnson, F. H.; Saiga, Y. Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan,Aequorea. *J. Cell. Comp. Physiol.* **1962**, *59* (3), 223–239. https://doi.org/10.1002/jcp.1030590302.

(147) Prasher, D. C.; Eckenrode, V. K.; Ward, W. W.; Prendergast, F. G.; Cormier, M. J. Primary Structure of the Aequorea Victoria Green-Fluorescent Protein. *Gene* **1992**, *111* (2), 229–233. https://doi.org/10.1016/0378-1119(92)90691-H.

(148) Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W. W.; Prasher, D. C. Green Fluorescent Protein as a Marker for Gene Expression. *Science (80-. ).* **1994**, *263* (5148), 802–805. https://doi.org/10.1126/science.8303295.

(149) Matz, M. V.; Fradkov, A. F.; Labas, Y. A.; Savitsky, A. P.; Zaraisky, A. G.; Markelov, M. L.; Lukyanov, S. A. Fluorescent Proteins from Nonbioluminescent Anthozoa Species. *Nat. Biotechnol.* **1999**, *17* (10), 969–973. https://doi.org/10.1038/13657.

(150) Tsien, R. Y. The Green Fluorescent Protein. *Annual Review of Biochemistry*.

Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA November 28, **1998**, pp 509–544. https://doi.org/10.1146/annurev.biochem.67.1.509.

(151) Shimomura, O. Structure of the Chromophore of Aequorea Green Fluorescent Protein. *FEBS Lett.* **1979**, *104* (2), 220–222. https://doi.org/10.1016/0014-5793(79)80818-2.

(152) Ormö, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. J. Crystal Structure of the Aequorea Victoria Green Fluorescent Protein. *Science (80-. ).* **1996**, *273* (5280), 1392–1395. https://doi.org/10.1126/science.273.5280.1392.

(153) Meech, S. R. Excited State Reactions in Fluorescent Proteins. *Chem. Soc. Rev.* **2009**, *38* (10), 2922–2934. https://doi.org/10.1039/b820168b.

(154) Bertrand, E.; Chartrand, P.; Schaefer, M.; Shenoy, S. M.; Singer, R. H.; Long, R. M. Localization of ASH1 MRNA Particles in Living Yeast. *Mol. Cell* **1998**, *2* (4), 437–445. https://doi.org/10.1016/S1097-2765(00)80143-4.

(155) Holeman, L. A.; Robinson, S. L.; Szostak, J. W.; Wilson, C. Isolation and Characterization of Fluorophore-Binding RNA Aptamers. *Fold. Des.* **1998**, *3* (6), 423–431. https://doi.org/10.1016/S1359-0278(98)00059-5.

(156) Grate, D.; Wilson, C. Laser-Mediated, Site-Specific Inactivation of RNA Transcripts. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (11), 6131–6136. https://doi.org/10.1073/pnas.96.11.6131.

(157) Babendure, J. R.; Adams, S. R.; Tsien, R. Y. Aptamers Switch on Fluorescence of Triphenylmethane Dyes. *J. Am. Chem. Soc.* **2003**, *125* (48), 14716–14717. https://doi.org/10.1021/ja037994o.

(158) 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), 1561–1564. https://doi.org/10.1021/ol702920e.

(159) Eydeler, K.; Magbanua, E.; Werner, A.; Ziegelmüller, P.; Hahn, U. Fluorophore Binding Aptamers as a Tool for RNA Visualization. *Biophys. J.* **2009**, *96* (9), 3703–3707. https://doi.org/10.1016/j.bpj.2009.01.041.

(160) Paige, J. S.; Wu, K. Y.; Jaffrey, S. R. RNA Mimics of Green Fluorescent Protein. *Science (80-. ).* **2011**, *333* (6042), 642–646. https://doi.org/10.1126/science.1207339.

(161) Han, K. Y.; Leslie, B. J.; Fei, J.; Zhang, J.; Ha, T. Understanding the Photophysics of the Spinach–DFHBI RNA Aptamer–Fluorogen Complex To Improve Live-Cell RNA Imaging. *J. Am. Chem. Soc.* **2013**, *135* (50), 19033–19038. https://doi.org/10.1021/ja411060p.

(162) 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), 1198–1201. https://doi.org/10.1021/ja410819x.

(163) 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), 1219–1224. https://doi.org/10.1038/nmeth.2701.

(164) Autour, A.; Westhof, E.; Ryckelynck, M. ISpinach: A Fluorogenic RNA Aptamer Optimized for in Vitro Applications. *Nucleic Acids Res.* **2016**, *44* (6), 2491–2500. https://doi.org/10.1093/nar/gkw083.

(165) 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), 16299–16308. https://doi.org/10.1021/ja508478x.

(166) Dolgosheina, E. V.; Jeng, S. C. Y.; Panchapakesan, S. S. S.; Cojocaru, R.; Chen, P. S. K.; 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), 2412–2420. https://doi.org/10.1021/cb500499x.

(167) Warner, K. D.; Sjekloća, L.; Song, W.; Filonov, G. S.; Jaffrey, S. R.; Ferré-D'Amaré, A. R. A Homodimer Interface without Base Pairs in an RNA Mimic of Red Fluorescent Protein. *Nat. Chem. Biol.* **2017**, *13* (11), 1195–1201. https://doi.org/10.1038/nchembio.2475.

(168) 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. https://doi.org/10.1021/jacs.7b04211.

(169) Steinmetzger, C.; Bessi, I.; Lenz, A. K.; Höbartner, C. Structure-Fluorescence Activation Relationships of a Large Stokes Shift Fluorogenic RNA Aptamer. *Nucleic Acids Res.* **2019**, *47* (22), 11538–11550. https://doi.org/10.1093/nar/gkz1084.

(170) Okuda, M.; Fourmy, D.; Yoshizawa, S. Use of Baby Spinach and Broccoli for Imaging of Structured Cellular RNAs. *Nucleic Acids Res.* **2017**, *45* (3), 1404–1415. https://doi.org/10.1093/nar/gkw794.

(171) Panchapakesan, S. S. S.; Ferguson, M. L.; Hayden, E. J.; Chen, X.; Hoskins, A. A.; Unrau, P. J. Ribonucleoprotein Purification and Characterization Using RNA Mango. *RNA* **2017**, *23* (10), 1592–1599. https://doi.org/10.1261/rna.062166.117.

(172) Nilaratanakul, V.; Hauer, D. A.; Griffin, D. E. Development and Characterization of Sindbis Virus with Encoded Fluorescent RNA Aptamer Spinach2 for Imaging of Replication and Immune-Mediated Changes in Intracellular Viral RNA. *J. Gen. Virol.* **2017**, *98* (5), 992–1003. https://doi.org/10.1099/jgv.0.000755.

(173) Autour, A.; Jeng, S. C. Y.; Cawte, A. D.; Abdolahzadeh, A.; Galli, A.; Panchapakesan, S. S. S.; Rueda, D.; Ryckelynck, M.; Unrau, P. J. Fluorogenic RNA Mango Aptamers for Imaging Small Non-Coding RNAs in Mammalian Cells. *Nat. Commun.* **2018**, *9* (1). https://doi.org/10.1038/s41467-018-02993-8.

(174) Porciani, D.; Cardwell, L. N.; Tawiah, K. D.; Alam, K. K.; Lange, M. J.; Daniels, M. A.; Burke, D. H. Modular Cell-Internalizing Aptamer Nanostructure Enables Targeted Delivery of Large Functional RNAs in Cancer Cell Lines. *Nat. Commun.* **2018**, *9* (1), 1–13. https://doi.org/10.1038/s41467-018-04691-x.

(175) Ouellet, J. RNA Fluorescence with Light-Up Aptamers. *Frontiers in Chemistry*. Frontiers Media S. A **2016**. https://doi.org/10.3389/fchem.2016.00029.

(176) Warner, K. D.; Chen, M. C.; Song, W.; Strack, R. L.; Thorn, A.; Jaffrey, S. R.; Ferré-D'Amaré, A. R. Structural Basis for Activity of Highly Efficient RNA Mimics of Green Fluorescent Protein. *Nat. Struct. Mol. Biol.* **2014**, *21* (8), 658–663. https://doi.org/10.1038/nsmb.2865.

(177) Huang, H.; Suslov, N. B.; Li, N.-S.; Shelke, S. A.; Evans, M. E.; Koldobskaya, Y.; Rice, P. A.; Piccirilli, J. A. A G-Quadruplex–Containing RNA Activates Fluorescence in a GFP-like Fluorophore. *Nat. Chem. Biol.* **2014**, *10* (8), 686–691. https://doi.org/10.1038/nchembio.1561.

(178) You, M.; Jaffrey, S. R. Structure and Mechanism of RNA Mimics of Green Fluorescent Protein. *Annu. Rev. Biophys.* **2015**, *44*, 187–206. https://doi.org/10.1146/annurev-biophys-060414-033954.

(179) Cuesta, J.; Read, M.; Neidle, S. The Design of G-Quadruplex Ligands as Telomerase Inhibitors. *Mini-Reviews Med. Chem.* **2012**, *3* (1), 11–21. https://doi.org/10.2174/1389557033405502.

(180) Paige, J. S.; Nguyen-Duc, T.; Song, W.; Jaffrey, S. R. Fluorescence Imaging of Cellular Metabolites with RNA. *Science*. American Association for the Advancement of Science March 9, **2012**, p 1194. https://doi.org/10.1126/science.1218298.

(181) Mccown, P. J.; Corbino, K. A.; Stav, S.; Sherlock, M. E.; Breaker, R. R. Riboswitch Diversity and Distribution. *RNA* **2017**, *23* (7), 995–1011. https://doi.org/10.1261/rna.061234.117.

(182) Kellenberger, C. A.; Hammond, M. C. In Vitro Analysis of Riboswitch-Spinach Aptamer Fusions as Metabolite-Sensing Fluorescent Biosensors. In *Methods in Enzymology*; Academic Press Inc., **2015**; Vol. 550, pp 147–172. https://doi.org/10.1016/bs.mie.2014.10.045.

(183) Kellenberger, C. A.; Wilson, S. C.; Sales-Lee, J.; Hammond, M. C. RNA-Based Fluorescent Biosensors for Live Cell Imaging of Second Messengers Cyclic Di-GMP and Cyclic AMP-GMP. *J. Am. Chem. Soc.* **2013**, *135* (13), 4906–4909.

https://doi.org/10.1021/ja311960g.

(184) You, M.; Litke, J. L.; Jaffrey, S. R. Imaging Metabolite Dynamics in Living Cells Using a Spinach-Based Riboswitch. *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences May 26, **2015**, pp E2756–E2765. https://doi.org/10.1073/pnas.1504354112.

(185) Li, X.; Mo, L.; Litke, J. L.; Dey, S. K.; Suter, S. R.; Jaffrey, S. R. Imaging Intracellular S-Adenosyl Methionine Dynamics in Live Mammalian Cells with a Genetically Encoded Red Fluorescent RNA-Based Sensor. *J. Am. Chem. Soc.* **2020**, *142* (33), 14117–14124. https://doi.org/10.1021/jacs.0c02931.

(186) Ghosh, I.; Hamilton, A. D.; Regan, L. Antiparallel Leucine Zipper-Directed Protein Reassembly: Application to the Green Fluorescent Protein [12]. *Journal of the American Chemical Society*. Varshavsky June 14, **2000**, pp 5658–5659. https://doi.org/10.1021/ja994421w.

(187) Rogers, T. A.; Andrews, G. E.; Jaeger, L.; Grabow, W. W. Fluorescent Monitoring of RNA Assembly and Processing Using the Split-Spinach Aptamer. *ACS Synth. Biol.* **2015**, *4* (2), 162–166. https://doi.org/10.1021/sb5000725.

(188) Kikuchi, N.; Kolpashchikov, D. M. Split Spinach Aptamer for Highly Selective Recognition of DNA and RNA at Ambient Temperatures. *ChemBioChem* **2016**, *17* (17), 1589–1592. https://doi.org/10.1002/cbic.201600323.

(189) Alam, K. K.; Tawiah, K. D.; Lichte, M. F.; Porciani, D.; Burke, D. H. A Fluorescent Split Aptamer for Visualizing RNA-RNA Assembly in Vivo. *ACS Synth. Biol.* **2017**, *6* (9), 1710–1721. https://doi.org/10.1021/acssynbio.7b00059.

(190) Chandler, M.; Lyalina, T.; Halman, J.; Rackley, L.; Lee, L.; Dang, D.; Ke, W.; Sajja, S.; Woods, S.; Acharya, S.; Baumgarten, E.; Christopher, J.; Elshalia, E.; Hrebien, G.; Kublank, K.; Saleh, S.; Stallings, B.; Tafere, M.; Striplin, C.; Afonin, K. A. Broccoli Fluorets: Split Aptamers as a User-Friendly Fluorescent Toolkit for Dynamic RNA Nanotechnology. *Molecules* **2018**, *23* (12). https://doi.org/10.3390/molecules23123178.

(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. Chemie - Int. Ed.* **2018**, *57* (4), 972–976. https://doi.org/10.1002/anie.201707795.

(192) Ausländer, S.; Fuchs, D.; Hürlemann, S.; Ausländer, D.; Fussenegger, M. Engineering a Ribozyme Cleavage-Induced Split Fluorescent Aptamer Complementation Assay. *Nucleic Acids Res.* **2016**, *44* (10). https://doi.org/10.1093/nar/gkw117.

(193) Halman, J. R.; Satterwhite, E.; Roark, B.; Chandler, M.; Viard, M.; Ivanina, A.;

Bindewald, E.; Kasprzak, W. K.; Panigaj, M.; Bui, M. N.; Lu, J. S.; Miller, J.; Khisamutdinov, E. F.; Shapiro, B. A.; Dobrovolskaia, M. A.; Afonin, K. A. Functionally-Interdependent Shape-Switching Nanoparticles with Controllable Properties. *Nucleic Acids Res.* **2017**, *45* (4), 2210–2220. https://doi.org/10.1093/nar/gkx008.

(194) O'Hara, J. M.; Marashi, D.; Morton, S.; Jaeger, L.; Grabow, W. W. Optimization of the Split-Spinach Aptamer for Monitoring Nanoparticle Assembly Involving Multiple Contiguous RNAs. *Nanomaterials* **2019**, *9* (3). https://doi.org/10.3390/nano9030378.

(195) Kikuchi, N.; Kolpashchikov, D. M. A Universal Split Spinach Aptamer (USSA) for Nucleic Acid Analysis and DNA Computation. *Chem. Commun.* **2017**, *53* (36), 4977–4980. https://doi.org/10.1039/c7cc01540b.

(196) Karunanayake Mudiyanselage, A. P. K. K.; Yu, Q.; Leon-Duque, M. A.; Zhao, B.; Wu, R.; You, M. Genetically Encoded Catalytic Hairpin Assembly for Sensitive RNA Imaging in Live Cells. *J. Am. Chem. Soc.* **2018**, *140* (28), 8739–8745. https://doi.org/10.1021/jacs.8b03956.

(197) Song, W.; Filonov, G. S.; Kim, H.; Hirsch, M.; Li, X.; Moon, J. D.; Jaffrey, S. R. Imaging RNA Polymerase III Transcription Using a Photostable RNA-Fluorophore Complex. *Nat. Chem. Biol.* **2017**, *13* (11), 1187–1194. https://doi.org/10.1038/nchembio.2477.

(198) Kim, H.; Jaffrey, S. R. A Fluorogenic RNA-Based Sensor Activated by Metabolite-Induced RNA Dimerization. *Cell Chem. Biol.* **2019**, *26* (12), 1725-1731.e6. https://doi.org/10.1016/j.chembiol.2019.09.013.

(199) Bhattacharyya, D.; Arachchilage, G. M.; Basu, S. Metal Cations in G-Quadruplex Folding and Stability. *Frontiers in Chemistry*. Frontiers Media S. A **2016**, p 38. https://doi.org/10.3389/fchem.2016.00038.

(200) Largy, E.; Mergny, J. L.; Gabelica, V. Role of Alkali Metal Ions in G-Quadruplex Nucleic Acid Structure and Stability. In *Metal Ions in Life Sciences*; Walter de Gruyter GmbH, **2016**; Vol. 16, pp 203–258. https://doi.org/10.1007/978-3-319-21756-7_7.

(201) Zhou, W.; Saran, R.; Liu, J. Metal Sensing by DNA. *Chem. Rev.* **2017**, *117* (12), 8272–8325. https://doi.org/10.1021/acs.chemrev.7b00063.

(202) DasGupta, S.; Shelke, S. A.; Li, N.; Piccirilli, J. A. Spinach RNA Aptamer Detects Lead( Ii ) with High Selectivity. *Chem. Commun.* **2015**, *51* (43), 9034–9037. https://doi.org/10.1039/C5CC01526J.

(203) Watson, J. D.; Crick, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171* (4356), 737–738. https://doi.org/10.1038/171737a0.

(204) Hoogsteen, K. The Crystal and Molecular Structure of a Hydrogen-Bonded Complex between 1-Methylthymine and 9-Methyladenine. *Acta Crystallogr.* **1963**, *16* (9), 907–916. https://doi.org/10.1107/s0365110x63002437.

(205) Miao, Z.; Westhof, E. RNA Structure: Advances and Assessment of 3D Structure Prediction. *Annual Review of Biophysics*. Annual Reviews Inc. May 22, **2017**, pp 483–503. https://doi.org/10.1146/annurev-biophys-070816-034125.

(206) Sen, D.; Gilbert, W. Formation of Parallel Four-Stranded Complexes by Guanine-Rich Motifs in DNA and Its Implications for Meiosis. *Nature* **1988**, *334* (6180), 364–366. https://doi.org/10.1038/334364a0.

(207) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. Quadruplex DNA: Sequence, Topology and Structure. *Nucleic Acids Res.* **2006**, *34* (19), 5402–5415. https://doi.org/10.1093/nar/gkl655.

(208) Henderson, E.; Hardin, C. C.; Walk, S. K.; Tinoco, I.; Blackburn, E. H. Telomeric DNA Oligonucleotides Form Novel Intramolecular Structures Containing Guanine·guanine Base Pairs. *Cell* **1987**, *51* (6), 899–908. https://doi.org/10.1016/0092-8674(87)90577-0.

(209) Wang, Y.; Patel, D. J. Solution Structure of a Parallel-Stranded G-Quadruplex DNA. *J. Mol. Biol.* **1993**, *234* (4), 1171–1183. https://doi.org/10.1006/jmbi.1993.1668.

(210) Simonsson, T.; Pecinka, P.; Kubista, M. DNA Tetraplex Formation in the Control Region of C-Myc. *Nucleic Acids Res.* **1998**, *26* (5), 1167–1172. https://doi.org/10.1093/nar/26.5.1167.

(211) Cogoi, S.; Xodo, L. E. G-Quadruplex Formation within the Promoter of the KRAS Proto-Oncogene and Its Effect on Transcription. *Nucleic Acids Res.* **2006**, *34* (9), 2536–2549. https://doi.org/10.1093/nar/gkl286.

(212) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. Direct Evidence for a G-Quadruplex in a Promoter Region and Its Targeting with a Small Molecule to Repress c-MYC Transcription. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (18), 11593–11598. https://doi.org/10.1073/pnas.182256799.

(213) Dai, J.; Chen, D.; Jones, R. A.; Hurley, L. H.; Yang, D. NMR Solution Structure of the Major G-Quadruplex Structure Formed in the Human BCL2 Promoter Region. *Nucleic Acids Res.* **2006**, *34* (18), 5133–5144. https://doi.org/10.1093/nar/gkl610.

(214) Balasubramanian, S.; Hurley, L. H.; Neidle, S. Targeting G-Quadruplexes in Gene Promoters: A Novel Anticancer Strategy? *Nat. Rev. Drug Discov.* **2011**, *10* (4), 261–275. https://doi.org/10.1038/nrd3428.

(215) Guo, J. U.; Bartel, D. P. RNA G-Quadruplexes Are Globally Unfolded in Eukaryotic Cells and Depleted in Bacteria. *Science (80-. ).* **2016**, *353* (6306). https://doi.org/10.1126/science.aaf5371.

(216) Yang, S. Y.; Lejault, P.; Chevrier, S.; Boidot, R.; Robertson, A. G.; Wong, J. M. Y.; Monchaud, D. Transcriptome-Wide Identification of Transient RNA G-Quadruplexes in Human Cells. *Nat. Commun.* **2018**, *9* (1), 4730. https://doi.org/10.1038/s41467-018-07224-8.

(217) Reshetnikov, R. V; Kopylov, A. M.; Golovin, A. V. Classification of G-Quadruplex DNA on the Basis of the Quadruplex Twist Angle and Planarity of G-Quartets. *Acta Naturae* **2010**, *2* (4), 72–81. https://doi.org/10.32607/20758251-2010-2-4-72-81.

(218) Gray, R. D.; Chaires, J. B. Linkage of Cation Binding and Folding in Human Telomeric Quadruplex DNA. *Biophys. Chem.* **2011**, *159* (1), 205–209. https://doi.org/10.1016/j.bpc.2011.06.012.

(219) Gu, J.; Leszczynski, J.; Bansal, M. A New Insight into the Structure and Stability of Hoogsteen Hydrogen-Bonded G-Tetrad: An Ab Initio SCF Study. *Chem. Phys. Lett.* **1999**, *311* (3–4), 209–214. https://doi.org/10.1016/S0009-2614(99)00821-0.

(220) Laughlan, G.; Murchie, A. I. H.; Norman, D. G.; Moore, M. H.; Moody, P. C. E.; Lilley, D. M. J.; Luisi, B. The High-Resolution Crystal Structure of a Parallel-Stranded Guanine Tetraplex. *Science (80-. ).* **1994**, *265* (5171), 520–524. https://doi.org/10.1126/science.8036494.

(221) Hud, N. V; Smith, F. W.; Anet, F. A.; Feigon, J. The Selectivity for K+ versus Na+ in DNA Quadruplexes Is Dominated by Relative Free Energies of Hydration: A Thermodynamic Analysis by 1H NMR. *Biochemistry* **1996**, *35* (48), 15383–15390. https://doi.org/10.1021/bi9620565.

(222) Doyle, D. A.; Cabral, J. M.; Pfuetzner, R. A.; Kuo, A.; Gulbis, J. M.; Cohen, S. L.; Chait, B. T.; MacKinnon, R. The Structure of the Potassium Channel: Molecular Basis of K+ Conduction and Selectivity. *Science (80-. ).* **1998**, *280* (5360), 69–77. https://doi.org/10.1126/science.280.5360.69.

(223) Zhou, M.; Morais-Cabral, J. H.; Mann, S.; MacKinnon, R. Potassium Channel Receptor Site for the Inactivation Gate and Quaternary Amine Inhibitors. *Nature* **2001**, *411* (6838), 657–661. https://doi.org/10.1038/35079500.

(224) Thomas, M.; Jayatilaka, D.; Corry, B. The Predominant Role of Coordination Number in Potassium Channel Selectivity. *Biophys. J.* **2007**, *93* (8), 2635–2643. https://doi.org/10.1529/biophysj.107.108167.

(225) Chaudhari, M. I.; Rempe, S. B. Strontium and Barium in Aqueous Solution and a Potassium Channel Binding Site. *J. Chem. Phys.* **2018**, *148* (22), 222831. https://doi.org/10.1063/1.5023130.

(226) Kankia, B. I.; Marky, L. A. Folding of the Thrombin Aptamer into a G-Quadruplex with Sr2+: Stability, Heat, and Hydration. *J. Am. Chem. Soc.* **2001**, *123* (44), 10799–10804. https://doi.org/10.1021/ja010008o.

(227) Liu, W.; Zhu, H.; Zheng, B.; Cheng, S.; Fu, Y.; Li, W.; Lau, T. C.; Liang, H. Kinetics

and Mechanism of G-Quadruplex Formation and Conformational Switch in a G-Quadruplex of PS2.M Induced by Pb2+. *Nucleic Acids Res.* **2012**, *40* (9), 4229–4236. https://doi.org/10.1093/nar/gkr1310.

(228) Liu, W.; Fu, Y.; Zheng, B.; Cheng, S.; Li, W.; Lau, T. C.; Liang, H. Kinetics and Mechanism of Conformational Changes in a G-Quadruplex of Thrombin-Binding Aptamer Induced by Pb2+. *J. Phys. Chem. B* **2011**, *115* (44), 13051–13056. https://doi.org/10.1021/jp2074489.

(229) Venczel, E. A.; Sen, D. Parallel and Antiparallel G-DNA Structures from a Complex Telomeric Sequence. *Biochemistry* **1993**, *32* (24), 6220–6228. https://doi.org/10.1021/bi00075a015.

(230) Deng, J.; Xiong, Y.; Sundaralingam, M. X-Ray Analysis of an RNA Tetraplex (UGGGGU)4 with Divalent Sr2+ Ions at Subatomic Resolution (0.61 Å). *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (24), 13665–13670. https://doi.org/10.1073/pnas.241374798.

(231) Mergny, J. L.; Phan, A. T.; Lacroix, L. Following G-Quartet Formation by UV-Spectroscopy. *FEBS Lett.* **1998**, *435* (1), 74–78. https://doi.org/10.1016/S0014-5793(98)01043-6.

(232) Vorlíčková, M.; Kejnovská, I.; Sagi, J.; Renčiuk, D.; Bednářová, K.; Motlová, J.; Kypr, J. Circular Dichroism and Guanine Quadruplexes. *Methods*. May **2012**, pp 64–75. https://doi.org/10.1016/j.ymeth.2012.03.011.

(233) Del Villar-Guerra, R.; Trent, J. O.; Chaires, J. B. G-Quadruplex Secondary Structure Obtained from Circular Dichroism Spectroscopy. *Angew. Chem. Int. Ed. Engl.* **2018**, *57* (24), 7171–7175. https://doi.org/10.1002/anie.201709184.

(234) Bock, L. C.; Griffin, L. C.; Latham, J. A.; Vermaas, E. H.; Toole, J. J. Selection of Single-Stranded DNA Molecules That Bind and Inhibit Human Thrombin. *Nature* **1992**, *355* (6360), 564–566. https://doi.org/10.1038/355564a0.

(235) Griffin, L.; Tidmarsh, G.; Bock, L.; Toole, J.; Leung, L. In Vivo Anticoagulant Properties of a Novel Nucleotide-Based Thrombin Inhibitor and Demonstration of Regional Anticoagulation in Extracorporeal Circuits. *Blood* **1993**, *81* (12), 3271–3276. https://doi.org/10.1182/blood.v81.12.3271.3271.

(236) Padmanabhan, K.; Padmanabhan, K. P.; Ferrara, J. D.; Sadler, J. E.; Tulinsky, A. The Structure of α-Thrombin Inhibited by a 15-Mer Single-Stranded DNA Aptamer. *J. Biol. Chem.* **1993**, *268* (24), 17651–17654. https://doi.org/10.1016/s0021-9258(17)46749-4.

(237) Ueyama, H.; Takagi, M.; Takenaka, S. A Novel Potassium Sensing in Aqueous Media with a Synthetic Oligonucleotide Derivative. Fluorescence Resonance Energy Transfer Associated with Guanine Quartet-Potassium Ion Complex Formation. *J. Am. Chem. Soc.* **2002**, *124* (48), 14286–14287.

https://doi.org/10.1021/ja026892f.

(238) Nojima, T.; Ueyama, H.; Takagi, M.; Takenaka, S. Patassium Sensing Oligonucleotide, PSO, Based on DNA Tetraplex Formation. *Nucleic Acids Res. Suppl.* **2002**, No. 2, 125–126. https://doi.org/10.1093/nass/2.1.125.

(239) Takenaka, S.; Ueyama, H.; Nojima, T.; Takagi, M. Comparison of Potassium Ion Preference of Potassium-Sensing Oligonucleotides, PSO-1 and PSO-2, Carrying the Human and Oxytricha Telomeric Sequence, Respectively. In *Analytical and Bioanalytical Chemistry*; Anal Bioanal Chem, **2003**; Vol. 375, pp 1006–1010. https://doi.org/10.1007/s00216-003-1799-z.

(240) Nagatoishi, S.; Nojima, T.; Galezowska, E.; Juskowiak, B.; Takenaka, S. G Quadruplex-Based FRET Probes with the Thrombin-Binding Aptamer (TBA) Sequence Designed for the Efficient Fluorometric Detection of the Potassium Ion. *ChemBioChem* **2006**, *7* (11), 1730–1737. https://doi.org/10.1002/cbic.200600179.

(241) Nagatoishi, S.; Nojima, T.; Galezowska, E.; Gluszynska, A.; Juskowiak, B.; Takenaka, S. Fluorescence Energy Transfer Probes Based on the Guanine Quadruplex Formation for the Fluorometric Detection of Potassium Ion. *Anal. Chim. Acta* **2007**, *581* (1), 125–131. https://doi.org/10.1016/j.aca.2006.08.010.

(242) Hayashida, H.; Paczesny, J.; Juskowiak, B.; Takenaka, S. Interactions of Sodium and Potassium Ions with Oligonucleotides Carrying Human Telomeric Sequence and Pyrene Moieties at Both Termini. *Bioorganic Med. Chem.* **2008**, *16* (22), 9871–9881. https://doi.org/10.1016/j.bmc.2008.08.035.

(243) He, F.; Tang, Y.; Wang, S.; Li, Y.; Zhu, D. Fluorescent Amplifying Recognition for DNA G-Quadruplex Folding with a Cationic Conjugated Polymer: A Platform for Homogeneous Potassium Detection. *J. Am. Chem. Soc.* **2005**, *127* (35), 12343–12346. https://doi.org/10.1021/ja051507i.

(244) Świtalska, A.; Dembska, A.; Fedoruk-Wyszomirska, A.; Juskowiak, B. Cholesterol-Bearing Fluorescent g-Quadruplex Potassium Probes for Anchoring at the Langmuir Monolayer and Cell Membrane. *Sensors (Switzerland)* **2018**, *18* (7). https://doi.org/10.3390/s18072201.

(245) Radi, A. E.; O'Sullivan, C. K. Aptamer Conformational Switch as Sensitive Electrochemical Biosensor for Potassium Ion Recognition. *Chem. Commun.* **2006**, No. 32, 3432–3434. https://doi.org/10.1039/b606804a.

(246) Zavyalova, E.; Tagiltsev, G.; Reshetnikov, R.; Arutyunyan, A.; Kopylov, A. Cation Coordination Alters the Conformation of a Thrombin-Binding G-Quadruplex DNA Aptamer That Affects Inhibition of Thrombin. *Nucleic Acid Ther.* **2016**, *26* (5), 299–308. https://doi.org/10.1089/nat.2016.0606.

(247) Wagh, A. A.; Fernandes, M. 2'-5'-Isomerically Linked Thrombin-Binding Aptamer (IsoTBA) Forms a Stable Unimolecular Parallel G-Quadruplex in the Presence of Sr

$^{2+}$ Ions. *ChemistrySelect* **2019**, *4* (36), 10668–10673. https://doi.org/10.1002/slct.201902005.

(248) Brenneman, K. L.; Poduri, S.; Stroscio, M. A.; Dutta, M. Optical Detection of Lead(II) Ions Using DNA-Based Nanosensor. *IEEE Sens. J.* **2013**, *13* (5), 1783–1786. https://doi.org/10.1109/JSEN.2013.2241757.

(249) Wu, Z. S.; Chen, C. R.; Shen, G. L.; Yu, R. Q. Reversible Electronic Nanoswitch Based on DNA G-Quadruplex Conformation: A Platform for Single-Step, Reagentless Potassium Detection. *Biomaterials* **2008**, *29* (17), 2689–2696. https://doi.org/10.1016/j.biomaterials.2008.02.024.

(250) Wang, L.; Liu, X.; Hu, X.; Song, S.; Fan, C. Unmodified Gold Nanoparticles as a Colorimetric Probe for Potassium DNA Aptamers. *Chem. Commun.* **2006**, No. 36, 3780–3782. https://doi.org/10.1039/b607448k.

(251) Jacobi, Z. E.; Li, L.; Liu, J. Visual Detection of Lead(II) Using a Label-Free DNA-Based Sensor and Its Immobilization within a Monolithic Hydrogel. *Analyst* **2012**, *137* (3), 704–709. https://doi.org/10.1039/c2an15754c.

(252) Li, T.; Dong, S.; Wang, E. A Lead(II)-Driven DNA Molecular Device for Turn-on Fluorescence Detection of Lead(II) Ion with High Selectivity and Sensitivity. *J. Am. Chem. Soc.* **2010**, *132* (38), 13156–13157. https://doi.org/10.1021/ja105849m.

(253) Liu, L.; Shao, Y.; Peng, J.; Huang, C.; Liu, H.; Zhang, L. Molecular Rotor-Based Fluorescent Probe for Selective Recognition of Hybrid G-Quadruplex and as a K+ Sensor. *Anal. Chem.* **2014**, *86* (3), 1622–1631. https://doi.org/10.1021/ac403326m.

(254) Qin, H.; Ren, J.; Wang, J.; Luedtke, N. W.; Wang, E. G-Quadruplex-Modulated Fluorescence Detection of Potassium in the Presence of a 3500-Fold Excess of Sodium Ions. *Anal. Chem.* **2010**, *82* (19), 8356–8360. https://doi.org/10.1021/ac101894b.

(255) Yang, L.; Qing, Z.; Liu, C.; Tang, Q.; Li, J.; Yang, S.; Zheng, J.; Yang, R.; Tan, W. Direct Fluorescent Detection of Blood Potassium by Ion-Selective Formation of Intermolecular G-Quadruplex and Ligand Binding. *Anal. Chem.* **2016**, *88* (18), 9285–9292. https://doi.org/10.1021/acs.analchem.6b02667.

(256) Dimroth, K.; Jaenicke, L.; Heinzel, D. Die Spaltung Der Pentose-Nucleinsäure Der Hefe Mit Bleihydroxyd. *Justus Liebigs Ann. Chem.* **1950**, *566* (2), 206–210. https://doi.org/10.1002/jlac.19505660209.

(257) Werner, C.; Krebs, B.; Keith, G.; Dirheimer, G. Specific Cleavages of Pure TRNAs by Plumbous Ions. *BBA Sect. Nucleic Acids Protein Synth.* **1976**, *432* (2), 161–175. https://doi.org/10.1016/0005-2787(76)90158-1.

(258) Brown, R. S.; Dewan, J. C.; Klug, A. Crystallographic and Biochemical Investigation of the Lead(II)-Catalyzed Hydrolysis of Yeast Phenylalanine TRNA. *Biochemistry*

**1985**, *24* (18), 4785–4801. https://doi.org/10.1021/bi00339a012.

(259) Ciesiołka, J. Metal Ion - Induced Cleavages in Probing of RNA Structure. In *RNA Biochemistry and Biotechnology*; Springer Netherlands, **1999**; pp 111–121. https://doi.org/10.1007/978-94-011-4485-8_7.

(260) Gomicki, P.; Baudin, F.; Romby, P.; Wiewiorowski, M.; Kryzosiak, W.; Ebel, J. P.; Ehresmann, C.; Ehresmann, B. Use of Lead(H) to Probe the Structure of Large Rna's. Conformation of the 3' Terminal Domain of e.Coli 16s Rrna and Its Involvement in Building the Trna Binding Sites. *J. Biomol. Struct. Dyn.* **1989**, *6* (5), 971–984. https://doi.org/10.1080/07391102.1989.10506525.

(261) Brunel, C.; Romby, P.; Westhof, E.; Ehresmann, C.; Ehresmann, B. Three-Dimensional Model of Escherichia Coli Ribosomal 5 S RNA as Deduced from Structure Probing in Solution and Computer Modeling. *J. Mol. Biol.* **1991**, *221* (1), 293–308. https://doi.org/10.1016/0022-2836(91)80220-O.

(262) CIESIOLKA, J.; LORENZ, S.; ERDMANN, V. A. Different Conformational Forms of Escherichia Coli and Rat Liver 5S RRNA Revealed by Pb(II)-induced Hydrolysis. *Eur. J. Biochem.* **1992**, *204* (2), 583–589. https://doi.org/10.1111/j.1432-1033.1992.tb16671.x.

(263) Lindell, M.; Brännvall, M.; Wagner, E. G. H.; Kirsebom, L. A. Lead(II) Cleavage Analysis of RNase P RNA in Vivo. *RNA* **2005**, *11* (9), 1348–1354. https://doi.org/10.1261/rna.2590605.

(264) Brown, R. S.; Hingerty, B. E.; Dewan, J. C.; Klug, A. Pb(II)-Catalysed Cleavage of the Sugar-Phosphate Backbone of Yeast TRNAPhe-Implications for Lead Toxicity and Self-Splicing RNA. *Nature* **1983**, *303* (5917), 543–546. https://doi.org/10.1038/303543a0.

(265) Winter, D.; Polacek, N.; Halama, I.; Streicher, B.; Barta, A. Lead-Catalysed Specific Cleavage of Ribosomal RNAs. *Nucleic Acids Res.* **1997**, *25* (9), 1817–1824. https://doi.org/10.1093/nar/25.9.1817.

(266) Abeydeera, N. D.; Egli, M.; Cox, N.; Mercier, K.; Conde, J. N.; Pallan, P. S.; Mizurini, D. M.; Sierant, M.; Hibti, F. E.; Hassell, T.; Wang, T.; Liu, F. W.; Liu, H. M.; Martinez, C.; Sood, A. K.; Lybrand, T. P.; Frydman, C.; Monteiro, R. Q.; Gomer, R. H.; Nawrot, B.; Yang, X. Evoking Picomolar Binding in RNA by a Single Phosphorodithioate Linkage. *Nucleic Acids Res.* **2016**, *44* (17), 8052–8064. https://doi.org/10.1093/nar/gkw725.

(267) Santulli-Maraotto, S.; Nair, S. K.; Rusconi, C.; Sullenger, B.; Gilboa, E. Multivalent RNA Aptamers That Inhibit CTLA-4 and Enhance Tumor Immunity. *Cancer Res.* **2003**, *63* (21), 7483–7489.

(268) Mallikaratchy, P. R.; Ruggiero, A.; Gardner, J. R.; Kuryavyi, V.; Maguire, W. F.; Heaney, M. L.; McDevitt, M. R.; Patel, D. J.; Scheinberg, D. A. A Multivalent DNA

Aptamer Specific for the B-Cell Receptor on Human Lymphoma and Leukemia. *Nucleic Acids Res.* **2011**, *39* (6), 2458–2469. https://doi.org/10.1093/nar/gkq996.

(269) Nonaka, Y.; Yoshida, W.; Abe, K.; Ferri, S.; Schulze, H.; Bachmann, T. T.; Ikebukuro, K. Affinity Improvement of a VEGF Aptamer by in Silico Maturation for a Sensitive VEGF-Detection System. *Anal. Chem.* **2013**, *85* (2), 1132–1137. https://doi.org/10.1021/ac303023d.

(270) Müller, J.; Wulffen, B.; Pötzsch, B.; Mayer, G. Multidomain Targeting Generates a High-Affinity Thrombin-Inhibiting Bivalent Aptamer. *ChemBioChem* **2007**, *8* (18), 2223–2226. https://doi.org/10.1002/cbic.200700535.

(271) Kim, Y.; Cao, Z.; Tan, W. Molecular Assembly for High-Performance Bivalent Nucleic Acid Inhibitor. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (15), 5664–5669. https://doi.org/10.1073/pnas.0711803105.

(272) Liu, Y.; Li, J.; Chen, Z.; Huang, W.; Cai, Z. Synthesizing Artificial Devices That Redirect Cellular Information at Will. *Elife* **2018**, *7*. https://doi.org/10.7554/eLife.31936.

(273) Miyakawa, S.; Oguro, A.; Ohtsu, T.; Imataka, H.; Sonenberg, N.; Nakamura, Y. RNA Aptamers to Mammalian Initiation Factor 4G Inhibit Cap-Dependent Translation by Blocking the Formation of Initiation Factor Complexes. *RNA* **2006**, *12* (10), 1825–1834. https://doi.org/10.1261/rna.2169406.

(274) Berens, C.; Thain, A.; Schroeder, R. A Tetracycline-Binding RNA Aptamer. *Bioorganic Med. Chem.* **2001**, *9* (10), 2549–2556. https://doi.org/10.1016/S0968-0896(01)00063-3.

(275) Müller, M.; Weigand, J. E.; Weichenrieder, O.; Suess, B. Thermodynamic Characterization of an Engineered Tetracycline-Binding Riboswitch. *Nucleic Acids Res.* **2006**, *34* (9), 2607–2617. https://doi.org/10.1093/nar/gkl347.

(276) Li, T.; Wang, E.; Dong, S. Parallel G-Quadruplex-Specific Fluorescent Probe for Monitoring DNA Structural Changes and Label-Free Detection of Potassium Ion. *Anal. Chem.* **2010**, *82* (18), 7576–7580. https://doi.org/10.1021/ac1019446.

(277) Li, J.; Lu, Y. A Highly Sensitive and Selective Catalytic DNA Biosensor for Lead Ions [9]. *Journal of the American Chemical Society*. American Chemical Society October 25, **2000**, pp 10466–10467. https://doi.org/10.1021/ja0021316.

(278) Breaker, R. R.; Joyce, G. F. The Expanding View of RNA and DNA Function. *Chemistry and Biology*. Elsevier Ltd September 18, **2014**, pp 1059–1065. https://doi.org/10.1016/j.chembiol.2014.07.008.

(279) Joyce, G. F. Evolution in an RNA World. In *Cold Spring Harbor Symposia on Quantitative Biology*; Cold Spring Harb Symp Quant Biol, **2009**; Vol. 74, pp 17–23. https://doi.org/10.1101/sqb.2009.74.004.

(280) Farkas, W. R. Depolymerization of Ribonucleic Acid by Plumbous Ion. *Biochim. Biophys. Acta* **1968**, *155* (2), 401–409.

(281) Ni, S.; Yao, H.; Wang, L.; Lu, J.; Jiang, F.; Lu, A.; Zhang, G. Chemical Modifications of Nucleic Acid Aptamers for Therapeutic Purposes. *International Journal of Molecular Sciences*. MDPI AG August 2, 2017. https://doi.org/10.3390/ijms18081683.

(282) Lipi, F.; Chen, S.; Chakravarthy, M.; Rakesh, S.; Veedu, R. N. In Vitro Evolution of Chemically-Modified Nucleic Acid Aptamers: Pros and Cons, and Comprehensive Selection Strategies. *RNA Biology*. Taylor and Francis Inc. December 1, **2016**, pp 1232–1245. https://doi.org/10.1080/15476286.2016.1236173.

(283) Tan, D.; Piana, S.; Dirks, R. M.; Shaw, D. E. RNA Force Field with Accuracy Comparable to State-of-the-Art Protein Force Fields. *Proc. Natl. Acad. Sci.* **2018**, *115* (7), E1346–E1355. https://doi.org/10.1073/pnas.1713027115.

(284) Du, Y.; Dong, S. Nucleic Acid Biosensors: Recent Advances and Perspectives. *Analytical Chemistry*. American Chemical Society January 3, **2017**, pp 189–215. https://doi.org/10.1021/acs.analchem.6b04190.

(285) Šponer, J.; Krepl, M.; Banáš, P.; Kührová, P.; Zgarbová, M.; Jurečka, P.; Havrila, M.; Otyepka, M. How to Understand Atomistic Molecular Dynamics Simulations of RNA and Protein-RNA Complexes? *Wiley Interdiscip. Rev. RNA* **2017**, *8* (3), e1405. https://doi.org/10.1002/wrna.1405.

(286) Dalvit, C.; Invernizzi, C.; Vulpetti, A. Fluorine as a Hydrogen-Bond Acceptor: Experimental Evidence and Computational Calculations. *Chem. - A Eur. J.* **2014**, *20* (35), 11058–11068. https://doi.org/10.1002/chem.201402858.

(287) 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. https://doi.org/10.1261/rna.063008.117.

(288) Tchounwou, P. B.; Yedjou, C. G.; Patlolla, A. K.; Sutton, D. J. Heavy Metal Toxicity and the Environment. *EXS* **2012**, *101*, 133–164. https://doi.org/10.1007/978-3-7643-8340-4_6.

(289) Wani, A. L.; Ara, A.; Usmani, J. A. Lead Toxicity: A Review. *Interdiscip. Toxicol.* **2015**, *8* (2), 55–64. https://doi.org/10.1515/intox-2015-0009.

(290) Needleman, H. Lead Poisoning. *Annu. Rev. Med.* **2004**, *55* (1), 209–222. https://doi.org/10.1146/annurev.med.55.091902.103653.

(291) Bellinger, D.; Leviton, A.; Waternaux, C.; Needleman, H.; Rabinowitz, M. Longitudinal Analyses of Prenatal and Postnatal Lead Exposure and Early Cognitive Development. *N. Engl. J. Med.* **1987**, *316* (17), 1037–1043. https://doi.org/10.1056/NEJM198704233161701.

(292) Needleman, H. L.; Schell, A.; Bellinger, D.; Leviton, A.; Allred, E. N. The Long-Term Effects of Exposure to Low Doses of Lead in Childhood. *N. Engl. J. Med.* **1990**, *322* (2), 83–88. https://doi.org/10.1056/NEJM199001113220203.

(293) Selander, S.; Cramér, K. Determination of Lead in Blood by Atomic Absorption Spectrophotometry. *Br. J. Ind. Med.* **1968**, *25* (3), 209–213.

(294) Hour, R. S. *Elemental and Isotopic Analysis by Inductively Coupled Plasma Mass Spectrometry*; **1994**; Vol. 27.

(295) Chen, C.-T.; Huang, W.-P. A Highly Selective Fluorescent Chemosensor for Lead Ions. *J. AM. CHEM. SOC* **2002**, *124*, 6246–6247. https://doi.org/10.1021/ja025710e.

(296) Qiwen He; Evan W. Miller; Audrey P. Wong, and; Chang*, C. J. A Selective Fluorescent Sensor for Detecting Lead in Living Cells. **2006**. https://doi.org/10.1021/JA063029X.

(297) Huang, H.; Suslov, N. B.; Li, N.-S.; Shelke, S. A.; Evans, M. E.; Koldobskaya, Y.; Rice, P. A.; Piccirilli, J. A. A G-Quadruplex–Containing RNA Activates Fluorescence in a GFP-like Fluorophore. *Nat. Chem. Biol.* **2014**, *10* (8), 686–691. https://doi.org/10.1038/nchembio.1561.

(298) Filonov, G. S.; Song, W.; Jaffrey, S. R. Spectral Tuning by a Single Nucleotide Controls the Fluorescence Properties of a Fluorogenic Aptamer. *Biochemistry* **2019**, *58* (12), 1560–1564. https://doi.org/10.1021/acs.biochem.9b00048.

(299) Vendeix, F. A. P.; Munoz, A. M.; Agris, P. F. Free Energy Calculation of Modified Base-Pair Formation in Explicit Solvent: A Predictive Model. *RNA* **2009**, *15* (12), 2278–2287. https://doi.org/10.1261/rna.1734309.

(300) Savage, J. C.; Shinde, P.; Bächinger, H. P.; Davare, M. A.; Shinde, U. A Ribose Modification of Spinach Aptamer Accelerates Lead(Ii) Cation Association in Vitro. *Chem. Commun. (Camb).* **2019**, *55* (42), 5882–5885. https://doi.org/10.1039/c9cc01697j.

(301) Fay, M. M.; Lyons, S. M.; Ivanov, P. RNA G-Quadruplexes in Biology: Principles and Molecular Mechanisms. *J. Mol. Biol.* **2017**, *429* (14), 2127–2147. https://doi.org/10.1016/j.jmb.2017.05.017.

(302) Simone, R.; Fratta, P.; Neidle, S.; Parkinson, G. N.; Isaacs, A. M. G-Quadruplexes: Emerging Roles in Neurodegenerative Diseases and the Non-Coding Transcriptome. *FEBS Letters*. Elsevier June 3, **2015**, pp 1653–1668. https://doi.org/10.1016/j.febslet.2015.05.003.

(303) Kharel, P.; Balaratnam, S.; Beals, N.; Basu, S. The Role of RNA G-quadruplexes in Human Diseases and Therapeutic Strategies. *Wiley Interdiscip. Rev. RNA* **2019**. https://doi.org/10.1002/wrna.1568.

(304) Bugaut, A.; Balasubramanian, S. 5'-UTR RNA G-Quadruplexes: Translation Regulation and Targeting. *Nucleic Acids Research*. June **2012**, pp 4727–4741. https://doi.org/10.1093/nar/gks068.

(305) Morris, M. J.; Negishi, Y.; Pazsint, C.; Schonhoft, J. D.; Basu, S. An RNA G-Quadruplex Is Essential for Cap-Independent Translation Initiation in Human VEGF IRES. *J. Am. Chem. Soc.* **2010**, *132* (50), 17831–17839. https://doi.org/10.1021/ja106287x.

(306) Crenshaw, E.; Leung, B. P.; Kwok, C. K.; Sharoni, M.; Olson, K.; Sebastian, N. P.; Ansaloni, S.; Schweitzer-Stenner, R.; Akins, M. R.; Bevilacqua, P. C.; Saunders, A. J. Amyloid Precursor Protein Translation Is Regulated by a 3'UTR Guanine Quadruplex. *PLoS One* **2015**, *10* (11), e0143160. https://doi.org/10.1371/journal.pone.0143160.

(307) Arora, A.; Suess, B. An RNA G-Quadruplex in the 3' UTR of the Proto-Oncogene PIM1 Represses Translation. *RNA Biol.* **2011**, *8* (5). https://doi.org/10.4161/rna.8.5.16038.

(308) Majerová, T.; Streckerová, T.; Bednárová, L.; Curtis, E. A. Sequence Requirements of Intrinsically Fluorescent G-Quadruplexes. *Biochemistry* **2018**, *57* (28), 4052–4062. https://doi.org/10.1021/acs.biochem.8b00252.

(309) Dao, N. T.; Haselsberger, R.; Michel-Beyerle, M. E.; Phan, A. T. Following G-Quadruplex Formation by Its Intrinsic Fluorescence. *FEBS Lett.* **2011**, *585* (24), 3969–3977. https://doi.org/10.1016/j.febslet.2011.11.004.

(310) Arora, A.; Maiti, S. Differential Biophysical Behavior of Human Telomeric RNA and DNA Quadruplex. *J. Phys. Chem. B* **2009**, *113* (30), 10515–10520. https://doi.org/10.1021/jp810638n.

(311) Guet, D.; Burns, L. T.; Maji, S.; Boulanger, J.; Hersen, P.; Wente, S. R.; Salamero, J.; Dargemont, C. Combining Spinach-Tagged RNA and Gene Localization to Image Gene Expression in Live Yeast. *Nat. Commun.* **2015**, *6*. https://doi.org/10.1038/ncomms9882.

(312) Popenda, M.; Szachniuk, M.; Antczak, M.; Purzycka, K. J.; Lukasiak, P.; Bartol, N.; Blazewicz, J.; Adamiak, R. W. Automated 3D Structure Composition for Large RNAs. *Nucleic Acids Res.* **2012**, *40* (14). https://doi.org/10.1093/nar/gks339.

(313) Rother, M.; Rother, K.; Puton, T.; Bujnicki, J. M. ModeRNA: A Tool for Comparative Modeling of RNA 3D Structure. *Nucleic Acids Res.* **2011**, *39* (10), 4007–4022. https://doi.org/10.1093/nar/gkq1320.

(314) Boniecki, M. J.; Lach, G.; Dawson, W. K.; Tomala, K.; Lukasz, P.; Soltysinski, T.; Rother, K. M.; Bujnicki, J. M. SimRNA: A Coarse-Grained Method for RNA Folding Simulations and 3D Structure Prediction. *Nucleic Acids Res.* **2015**, *44* (7). https://doi.org/10.1093/nar/gkv1479.

(315) Lu, D.; Searles, M. A.; Klug, A. Crystal Structure of a Zinc-Finger-RNA Complex Reveals Two Modes of Molecular Recognition. *Nature* **2003**, *426* (6962), 96–100. https://doi.org/10.1038/nature02088.

(316) Lukasiak, P.; Antczak, M.; Ratajczak, T.; Szachniuk, M.; Popenda, M.; Adamiak, R. W.; Blazewicz, J. RNAssess - A Web Server for Quality Assessment of RNA 3D Structures. *Nucleic Acids Res.* **2015**, *43* (W1), W502–W506. https://doi.org/10.1093/nar/gkv557.

(317) Savage, J. C.; Davare, M. A.; Shinde, U. Subtle Sequence Variations Alter Tripartite Complex Kinetics and G-Quadruplex Dynamics in RNA Aptamer Broccoli. *Chem. Commun.* **2020**, *56* (17), 2634–2637. https://doi.org/10.1039/c9cc09375c.

(318) Ellington, A. D.; Szostak, J. W. In Vitro Selection of RNA Molecules That Bind Specific Ligands. *Nature* **1990**, *346* (6287), 818–822. https://doi.org/10.1038/346818a0.

(319) Breaker, R. R. Riboswitches and the RNA World. In *The RNA World, 3rd Edition*; Gesteland, Raymond F., Cech, Thomas R., Atkins, J. F., Ed.; Cold Spring Harbor Laboratory Press, **2006**; pp 89–108.

(320) Breaker, R. R. Riboswitches and the RNA World. *Cold Spring Harb. Perspect. Biol.* **2012**, *4* (2). https://doi.org/10.1101/cshperspect.a003566.

(321) Gumz, M. L.; Rabinowitz, L.; Wingo, C. S. An Integrated View of Potassium Homeostasis. *N. Engl. J. Med.* **2015**, *373* (1), 60–72. https://doi.org/10.1056/nejmra1313341.

(322) Ni, S.; Yao, H.; Wang, L.; Lu, J.; Jiang, F.; Lu, A.; Zhang, G. Chemical Modifications of Nucleic Acid Aptamers for Therapeutic Purposes. *Int. J. Mol. Sci.* **2017**, *18* (8). https://doi.org/10.3390/ijms18081683.

(323) Kellenberger, C. A.; Wilson, S. C.; Sales-Lee, J.; Hammond, M. C. RNA-Based Fluorescent Biosensors for Live Cell Imaging of Second Messengers Cyclic Di-GMP and Cyclic AMP-GMP. *J. Am. Chem. Soc* **2013**, *135*, 58. https://doi.org/10.1021/ja311960g.

(324) Sekar, P.; Huang, D. Y.; Chang, S. F.; Lin, W. W. Coordinate Effects of P2X7 and Extracellular Acidification in Microglial Cells. *Oncotarget* **2018**, *9* (16), 12718–12731. https://doi.org/10.18632/oncotarget.24331.

(325) Kasner, S. E.; Ganz, M. B. Regulation of Intracellular Potassium in Mesangial Cells: A Fluorescence Analysis Using the Dye, PBFI. *Am. J. Physiol. - Ren. Fluid Electrolyte Physiol.* **1992**, *262* (3 31-3). https://doi.org/10.1152/ajprenal.1992.262.3.f462.

(326) Kong, X.; Su, F.; Zhang, L.; Yaron, J.; Lee, F.; Shi, Z.; Tian, Y.; Meldrum, D. R. A Highly Selective Mitochondria-Targeting Fluorescent K+ Sensor. *Angew. Chemie - Int. Ed.* **2015**, *54* (41), 12053–12057. https://doi.org/10.1002/anie.201506038.

(327) Zhou, X.; Su, F.; Tian, Y.; Youngbull, C.; Johnson, R. H.; Meldrum, D. R. A New Highly Selective Fluorescent K+ Sensor. *J. Am. Chem. Soc.* **2011**, *133* (46), 18530–18533. https://doi.org/10.1021/ja207345s.

(328) Rimmele, T. S.; Chatton, J.-Y. A Novel Optical Intracellular Imaging Approach for Potassium Dynamics in Astrocytes. *PLoS One* **2014**, *9* (10), e109243. https://doi.org/10.1371/journal.pone.0109243.

(329) Bischof, H.; Rehberg, M.; Stryeck, S.; Artinger, K.; Eroglu, E.; Waldeck-Weiermair, M.; Gottschalk, B.; Rost, R.; Deak, A. T.; Niedrist, T.; Vujic, N.; Lindermuth, H.; Prassl, R.; Pelzmann, B.; Groschner, K.; Kratky, D.; Eller, K.; Rosenkranz, A. R.; Madl, T.; Plesnila, N.; Graier, W. F.; Malli, R. Novel Genetically Encoded Fluorescent Probes Enable Real-Time Detection of Potassium in Vitro and in Vivo. *Nat. Commun.* **2017**, *8* (1), 1–12. https://doi.org/10.1038/s41467-017-01615-z.

(330) Shen, Y.; Wu, S. Y.; Rancic, V.; Aggarwal, A.; Qian, Y.; Miyashita, S. I.; Ballanyi, K.; Campbell, R. E.; Dong, M. Genetically Encoded Fluorescent Indicators for Imaging Intracellular Potassium Ion Concentration. *Commun. Biol.* **2019**, *2* (1). https://doi.org/10.1038/s42003-018-0269-2.

(331) Nagatoishi, S.; Nojima, T.; Juskowiak, B.; Takenaka, S. A Pyrene-Labeled G-Quadruplex Oligonucleotide as a Fluorescent Probe for Potassium Ion Detection in Biological Applications. *Angew. Chemie* **2005**, *117* (32), 5195–5198. https://doi.org/10.1002/ange.200501506.

(332) Ohtsuka, K.; Sato, S.; Sato, Y.; Sota, K.; Ohzawa, S.; Matsuda, T.; Takemoto, K.; Takamune, N.; Juskowiak, B.; Nagai, T.; Takenaka, S. Fluorescence Imaging of Potassium Ions in Living Cells Using a Fluorescent Probe Based on a Thrombin Binding Aptamer–Peptide Conjugate. *Chem. Commun.* **2012**, *48* (39), 4740–4742. https://doi.org/10.1039/c2cc30536d.

(333) Li, X.; Kim, H.; Litke, J. L.; Wu, J.; Jaffrey, S. R. Fluorophore-Promoted RNA Folding and Photostability Enables Imaging of Single Broccoli-Tagged MRNAs in Live Mammalian Cells. *Angew. Chemie Int. Ed.* **2020**, *59* (11), 4511–4518. https://doi.org/10.1002/anie.201914576.

(334) Swetha, P.; Fan, Z.; Wang, F.; Jiang, J. H. Genetically Encoded Light-up RNA Aptamers and Their Applications for Imaging and Biosensing. *Journal of Materials Chemistry B*. Royal Society of Chemistry April 28, **2020**, pp 3382–3392. https://doi.org/10.1039/c9tb02668a.

(335) Savage, J. C.; Davare, M. A.; Shinde, U. Subtle Sequence Variations Alter Tripartite Complex Kinetics and G-Quadruplex Dynamics in RNA Aptamer Broccoli. *Chem. Commun.* **2020**, *56* (17), 2634–2637. https://doi.org/10.1039/c9cc09375c.

(336) Darden, T.; Perera, L.; Li, L.; Lee, P. New Tricks for Modelers from the Crystallography Toolkit: The Particle Mesh Ewald Algorithm and Its Use in Nucleic Acid Simulations. *Structure* **1999**, *7* (3), R55–R60. https://doi.org/10.1016/S0969-

2126(99)80033-1.

(337) Bakan, A.; Meireles, L. M.; Bahar, I. ProDy: Protein Dynamics Inferred from Theory and Experiments. *Bioinformatics* **2011**, *27* (11), 1575–1577. https://doi.org/10.1093/bioinformatics/btr168.

(338) Savage, J. C.; Shinde, P.; Yao, Y.; Davare, M. A.; Shinde, U. A Broccoli Aptamer Chimera Yields a Fluorescent K+sensor Spanning Physiological Concentrations. *Chem. Commun.* **2021**, *57* (11), 1344–1347. https://doi.org/10.1039/d0cc07042d.

(339) Gholamalipour, Y.; Karunanayake Mudiyanselage, A.; Martin, C. T. 3' End Additions by T7 RNA Polymerase Are RNA Self-Templated, Distributive and Diverse in Character—RNA-Seq Analyses. *Nucleic Acids Res.* **2018**, *46* (18), 9253–9263. https://doi.org/10.1093/nar/gky796.

(340) Milligan, J. F.; Groebe, D. R.; Witherell, G. W.; Uhlenbeck, O. C. Oligoribonucleotide Synthesis Using T7 RNA Polymerase and Synthetic DNA Templates. *Nucleic Acids Res.* **1987**, *15* (21), 8783–8798. https://doi.org/10.1093/nar/15.21.8783.

(341) Lima, W. F.; Crooke, S. T. Binding Affinity and Specificity of Escherichia Coli RNase H1: Impact on the Kinetics of Catalysis of Antisense Oligonucleotide-RNA Hybrids. *Biochemistry* **1997**, *36* (2), 390–398. https://doi.org/10.1021/bi962230p.

(342) Schürer, H.; Lang, K.; Schuster, J.; Mörl, M. A Universal Method to Produce in Vitro Transcripts with Homogeneous 3' Ends. *Nucleic Acids Res.* **2002**, *30* (12), e56. https://doi.org/10.1093/nar/gnf055.

(343) Wichlacz, A.; Legiewicz, M.; Ciesiołka, J. Generating in Vitro Transcripts with Homogenous 3' Ends Using Trans-Acting Antigenomic Delta Ribozyme. *Nucleic Acids Res.* **2004**, *32* (3), e39–e39. https://doi.org/10.1093/nar/gnh037.

(344) Walker, S. C.; Avis, J. M.; Conn, G. L. General Plasmids for Producing RNA in Vitro Transcripts with Homogeneous Ends. *Nucleic Acids Res.* **2003**, *31* (15). https://doi.org/10.1093/nar/gng082.

(345) Perrotta, A. T.; Been, M. D. Cleavage of Oligoribonucleotides by a Ribozyme Derived from the Hepatitis Δ Virus RNA Sequence. *Biochemistry* **1992**, *31* (1), 16–21. https://doi.org/10.1021/bi00116a004.

(346) Suh, Y. ah; Kumar, P. K. R.; Taira, K.; Nishikawa, S. Self-Cleavage Activity of the Genomic HDV Ribozyme in the Presence of Various Divalent Metal Ions. *Nucleic Acids Res.* **1993**, *21* (14), 3277–3280. https://doi.org/10.1093/nar/21.14.3277.

(347) Wrzesinski, J.; Legiewicz, M.; Smólska, B.; Ciesiolka, J. *Catalytic Cleavage of Cis- and Trans-Acting Antigenomic Delta Ribozymes in the Presence of Various Divalent Metal Ions*; **2001**; Vol. 29.

(348) Burgin, A. B.; Gonzalez, C.; Matulic-Adamic, J.; Karpeisky, A. M.; Usman, N.;

McSwiggen, J. A.; Beigelman, L. Chemically Modified Hammerhead Ribozymes with Improved Catalytic Rates. *Biochemistry* **1996**, *35* (45), 14090–14097. https://doi.org/10.1021/bi961264u.

(349) Ferré-D'Amaré, A. R.; Zhou, K.; Doudna, J. A. Crystal Structure of a Hepatitis Delta Virus Ribozyme. *Nature* **1998**, *395* (6702), 567–574. https://doi.org/10.1038/26912.

(350) Perrotta, A. T.; Shih, I. H.; Been, M. D. Imidazole Rescue of a Cytosine Mutation in a Self-Cleaving Ribozyme. *Science (80-. ).* **1999**, *286* (5437), 123–126. https://doi.org/10.1126/science.286.5437.123.

(351) Chen, J. H.; Yajima, R.; Chadalavada, D. M.; Chase, E.; Bevilacqua, P. C.; Golden, B. L. A 1.9 Å Crystal Structure of the HDV Ribozyme Precleavage Suggests Both Lewis Acid and General Acid Mechanisms Contribute to Phosphodiester Cleavage. *Biochemistry* **2010**, *49* (31), 6508–6518. https://doi.org/10.1021/bi100670p.

(352) Thill, G.; Vasseur, M.; Tanner, N. K. Structural and Sequence Elements Required for the Self-Cleaving Activity of the Hepatitis Delta Virus Ribozyme. *Biochemistry* **1993**, *32* (16), 4254–4262. https://doi.org/10.1021/bi00067a013.

(353) Puttaraju, M.; Perrotta, A. T.; Been, M. D. A Circular Trans-Acting Hepatitis Delta Virus Ribozyme. *Nucleic Acids Res.* **1993**, *21* (18), 4253–4258. https://doi.org/10.1093/nar/21.18.4253.

(354) Ning, X.; Guo, J.; Wolfert, M. A.; Boons, G. J. Visualizing Metabolically Labeled Glycoconjugates of Living Cells by Copper-Free and Fast Huisgen Cycloadditions. *Angew. Chemie - Int. Ed.* **2008**, *47* (12), 2253–2255. https://doi.org/10.1002/anie.200705456.

(355) Chen, X.; Zhang, X.; Wang, H. Y.; Chen, Z.; Wu, F. G. Subcellular Fate of a Fluorescent Cholesterol-Poly(Ethylene Glycol) Conjugate: An Excellent Plasma Membrane Imaging Reagent. *Langmuir* **2016**, *32* (39), 10126–10135. https://doi.org/10.1021/acs.langmuir.6b02288.

(356) Vabbilisetty, P.; Boron, M.; Nie, H.; Ozhegov, E.; Sun, X. L. Chemical Reactive Anchoring Lipids with Different Performance for Cell Surface Re-Engineering Application. *ACS Omega* **2018**, *3* (2), 1589–1599. https://doi.org/10.1021/acsomega.7b01886.

(357) Jia, H. R.; Zhu, Y. X.; Xu, K. F.; Pan, G. Y.; Liu, X.; Qiao, Y.; Wu, F. G. Efficient Cell Surface Labelling of Live Zebrafish Embryos: Wash-Free Fluorescence Imaging for Cellular Dynamics Tracking and Nanotoxicity Evaluation. *Chem. Sci.* **2019**, *10* (14), 4062–4068. https://doi.org/10.1039/C8SC04884C.

(358) Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. A Strain-Promoted [3 + 2] Azide-Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. *J. Am. Chem. Soc.* **2004**, *126* (46), 15046–15047. https://doi.org/10.1021/ja044996f.

(359) Hawkins, M. E. Fluorescent Pteridine Nucleoside Analogs: A Window on DNA Interactions. *Cell Biochem. Biophys.* **2001**, *34* (2), 257–281. https://doi.org/10.1385/CBB:34:2:257.

(360) Myers, J. C.; Moore, S. A.; Shamoo, Y. Structure-Based Incorporation of 6-Methyl-8-(2-Deoxy-Î²-Ribofuranosyl)Isoxanthopteridine into the Human Telomeric Repeat DNA as a Probe for UP1 Binding and Destabilization of G-Tetrad Structures*. *J. Biol. Chem.* **2003**, *278*, 42300–42306. https://doi.org/10.1074/jbc.M306147200.

(361) Vrma, S.; Vaish, N. k.; Eckstein, F. Applications of Ribonucleotide Analogues in RNA Biochemistry. In *RNA*; Elsevier, **2001**; pp 259–275. https://doi.org/10.1016/b978-008043408-7/50036-8.

(362) Hawkins, M. E. Fluorescent Nucleoside Analoguesas DNA Probes. In *151Topics in Fluorescence Spectroscopy, Volume 7: DNA Technology*; Lakowicz, J. R., Ed.; Klumer Academic/Plenum Publishers: New York, **2003**; pp 151–175.

(363) Gros, J.; Rosu, F.; Amrane, S.; De Cian, A.; Gabelica, V.; Lacroix, L.; Mergny, J. L. Guanines Are a Quartet's Best Friend: Impact of Base Substitutions on the Kinetics and Stability of Tetramolecular Quadruplexes. *Nucleic Acids Res.* **2007**, *35* (9), 3064–3075. https://doi.org/10.1093/nar/gkm111.

(364) Han, J. H.; Chitrapriya, N.; Lee, H. S.; Lee, Y.-A.; Kim, S. K.; Jung, M.-J. Behavior of the Guanine Base in G-Quadruplexes Probed by the Fluorescent Guanine Analog, 6-Methyl Isoxanthopterin. *Bull. Korean Chem. Soc.* **2017**, *38* (2), 183–190. https://doi.org/10.1002/bkcs.11060.

(365) Olsen, C. M.; Gmeiner, W. H.; Marky, L. A. Unfolding of G-Quadruplexes: Energetic, and Ion and Water Contributions of G-Quartet Stacking. *J. Phys. Chem. B* **2006**, *110* (13), 6962–6969. https://doi.org/10.1021/jp0574697.

(366) Cang, X.; Šponer, J.; Cheatham, T. E. Explaining the Varied Glycosidic Conformational, G-Tract Length and Sequence Preferences for Anti-Parallel G-Quadruplexes. *Nucleic Acids Res.* **2011**, *39* (10), 4499–4512. https://doi.org/10.1093/nar/gkr031.

(367) Cheng, M.; Cheng, Y.; Hao, J.; Jia, G.; Zhou, J.; Mergny, J. L.; Li, C. Loop Permutation Affects the Topology and Stability of G-Quadruplexes. *Nucleic Acids Res.* **2018**, *46* (18), 9264–9275. https://doi.org/10.1093/nar/gky757.

(368) Bergues-Pupo, A. E.; Arias-Gonzalez, J. R.; Morón, M. C.; Fiasconaro, A.; Falo, F. Role of the Central Cations in the Mechanical Unfolding of DNA and RNA G-Quadruplexes. *Nucleic Acids Res.* **2015**, *43* (15), 7638–7647. https://doi.org/10.1093/nar/gkv690.

(369) Ren, J.; Qu, X.; Trent, J. O.; Chaires, J. B. Tiny Telomere DNA. *Nucleic Acids Res.* **2002**, *30* (11), 2307–2315. https://doi.org/10.1093/nar/30.11.2307.

(370) König, S. L. B.; Huppert, J. L.; Sigel, R. K. O.; Evans, A. C. Distance-Dependent

Duplex DNA Destabilization Proximal to G-Quadruplex/i-Motif Sequences. *Nucleic Acids Res.* **2013**, *41* (15), 7453–7461. https://doi.org/10.1093/nar/gkt476.

(371) Zhao, Y.; Kan, Z. Y.; Zeng, Z. X.; Hao, Y. H.; Chen, H.; Tan, Z. Determining the Folding and Unfolding Rate Constants of Nucleic Acids by Biosensor. Application Telomere G-Quadruplex. *J. Am. Chem. Soc.* **2004**, *126* (41), 13255–13264. https://doi.org/10.1021/ja048398c.

(372) Zhang, A. Y. Q.; Balasubramanian, S. The Kinetics and Folding Pathways of Intramolecular G-Quadruplex Nucleic Acids. *J. Am. Chem. Soc.* **2012**, *134* (46), 19297–19308. https://doi.org/10.1021/ja309851t.

(373) Ponzo, I.; Möller, F. M.; Daub, H.; Matscheko, N. A DNA-Based Biosensor Assay for the Kinetic Characterization of Ion-Dependent Aptamer Folding and Protein Binding. *Molecules* **2019**, *24* (16). https://doi.org/10.3390/molecules24162877.

(374) Gray, R. D.; Chaires, J. B. Kinetics and Mechanism of K+ and Na +-Induced Folding of Models of Human Telomeric DNA into G-Quadruplex Structures. *Nucleic Acids Res.* **2008**, *36* (12), 4191–4203. https://doi.org/10.1093/nar/gkn379.

(375) Phan, A. T.; Patel, D. J. Two-Repeat Human Telomeric d(TAGGGTTAGGGT) Sequence Forms Interconverting Parallel and Antiparallel G-Quadruplexes in Solution: Distinct Topologies, Thermodynamic Properties, and Folding/Unfolding Kinetics. *J. Am. Chem. Soc.* **2003**, *125* (49), 15021–15027. https://doi.org/10.1021/ja037616j.

(376) König, S. L. B.; Evans, A. C.; Huppert, J. L. Seven Essential Questions on G-Quadruplexes. *Biomolecular Concepts*. De Gruyter Mouton August 1, **2010**, pp 197–213. https://doi.org/10.1515/bmc.2010.011.

(377) Mayer, G.; Ahmed, M. S. L.; Dolf, A.; Endl, E.; Knolle, P. A.; Famulok, M. Fluorescence-Activated Cell Sorting for Aptamer SELEX with Cell Mixtures. *Nat. Protoc.* **2010**, *5* (12), 1993–2004. https://doi.org/10.1038/nprot.2010.163.

(378) Schultes, E.; Hraber, P. T.; Labean, T. H. Global Similarities in Nucleotide Base Composition among Disparate Functional Classes of Single-Stranded RNA Imply Adaptive Evolutionary Convergence. *RNA* **1997**, *3* (7), 792–806.

(379) Davis, J. H.; Szostak, J. W. Isolation of High-Affinity GTP Aptamers from Partially Structured RNA Libraries. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (18), 11616–11621. https://doi.org/10.1073/pnas.182095699.

(380) Carothers, J. M.; Davis, J. H.; Chou, J. J.; Szostak, J. W. Solution Structure of an Informationally Complex High-Affinity RNA Aptamer to GTP. *RNA* **2006**, *12* (4), 567–579. https://doi.org/10.1261/rna.2251306.

(381) Ruff, K. M.; Snyder, T. M.; Liu, D. R. Enhanced Functional Potential of Nucleic Acid Aptamer Libraries Patterned to Increase Secondary Structure. *J. Am. Chem. Soc.* **2010**, *132* (27), 9453–9464. https://doi.org/10.1021/ja103023m.

(382) Zhu, L.; Li, C.; Zhu, Z.; Liu, D.; Zou, Y.; Wang, C.; Fu, H.; Yang, C. J. In Vitro Selection of Highly Efficient G-Quadruplex-Based DNAzymes. *Anal. Chem.* **2012**, *84* (19), 8383–8390. https://doi.org/10.1021/ac301899h.

(383) McManus, S. A.; Li, Y. Assessing the Amount of Quadruplex Structures Present within G2-Tract Synthetic Random-Sequence DNA Libraries. *PLoS One* **2013**, *8* (5), e64131. https://doi.org/10.1371/journal.pone.0064131.

(384) Ferré-D'Amaré, A. R.; Scott, W. G. Small Self-Cleaving Ribozymes. *Cold Spring Harbor perspectives in biology*. Cold Spring Harbor Laboratory Press **2010**. https://doi.org/10.1101/cshperspect.a003574.

(385) Sabeti, P. C.; Unrau, P. J.; Bartel, D. P. Accessing Rare Activities from Random RNA Sequences: The Importance of the Length of Molecules in the Starting Pool. *Chem. Biol.* **1997**, *4* (10), 767–774. https://doi.org/10.1016/S1074-5521(97)90315-X.

(386) Higgs, P. G. RNA Secondary Structure: A Comparison of Real and Random Sequences. *J. Phys. I* **1993**, *3* (1), 43–59. https://doi.org/10.1051/jp1:1993116.

(387) Seffens, W.; Digby, D. MRNAs Have Greater Negative Folding Free Energies than Shuffled or Codon Choice Randomized Sequences. *Nucleic Acids Res.* **1999**, *27* (7), 1578–1584. https://doi.org/10.1093/nar/27.7.1578.

(388) Green, R.; Szostak, J. W. Selection of a Ribozyme That Functions as a Superior Template in a Self-Copying Reaction. *Science (80-. ).* **1992**, *258* (5090), 1910–1915. https://doi.org/10.1126/science.1470913.

(389) Connell, G. J.; Yarus, M. RNAs with Dual Specificity and Dual RNAs with Similar Specificity. *Science (80-. ).* **1994**, *264* (5162), 1137–1141. https://doi.org/10.1126/science.7513905.

(390) Lorsch, J. R.; Szostak, J. W. In Vitro Evolution of New Ribozymes with Polynucleotide Kinase Activity. *Nature* **1994**, *371* (6492), 31–36. https://doi.org/10.1038/371031a0.

(391) Jaeger, L.; Wright, M. C.; Joyce, G. F. A Complex Ligase Ribozyme Evolved in Vitro from a Group I Ribozyme Domain. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (26), 14712–14717. https://doi.org/10.1073/pnas.96.26.14712.

(392) Ohuchi, S. J.; Ikawa, Y.; Shiraishi, H.; Inoue, T. Modular Engineering of a Group I Intron Ribozyme. *Nucleic Acids Res.* **2002**, *30* (15), 3473–3480. https://doi.org/10.1093/nar/gkf453.

(393) Yoshioka, W.; Ikawa, Y.; Jaeger, L.; Shiraishi, H.; Inoue, T. Generation of a Catalytic Module on a Self-Folding RNA. *RNA* **2004**, *10* (12), 1900–1906. https://doi.org/10.1261/rna.7170304.

(394) Horning, D. P.; Joyce, G. F. Amplification of RNA by an RNA Polymerase Ribozyme.

*Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (35), 9786–9791. https://doi.org/10.1073/pnas.1610103113.

(395) Porter, E. B.; Polaski, J. T.; Morck, M. M.; Batey, R. T. Recurrent RNA Motifs as Scaffolds for Genetically Encodable Small-Molecule Biosensors. *Nat. Chem. Biol.* **2017**, *13* (3), 295–301. https://doi.org/10.1038/nchembio.2278.

(396) Robertson, M. P.; Joyce, G. F. The Origins of the RNA World. *Cold Spring Harb. Perspect. Biol.* **2012**, *4* (5), 1. https://doi.org/10.1101/cshperspect.a003608.

(397) Ye, J. D.; Tereshko, V.; Frederiksen, J. K.; Koide, A.; Fellouse, F. A.; Sidhu, S. S.; Koide, S.; Kossiakoff, A. A.; Piccirilli, J. A. Synthetic Antibodies for Specific Recognition and Crystallization of Structured RNA. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (1), 82–87. https://doi.org/10.1073/pnas.0709082105.

(398) White, R.; Rusconi, C.; Scardino, E.; Wolberg, A.; Lawson, J.; Hoffman, M.; Sullenger, B. Generation of Species Cross-Reactive Aptamers Using "Toggle" SELEX. *Mol. Ther.* **2001**, *4* (6), 567–573. https://doi.org/10.1006/mthe.2001.0495.

(399) Long, S. B.; Long, M. B.; White, R. R.; Sullenger, B. A. Crystal Structure of an RNA Aptamer Bound to Thrombin. *RNA* **2008**, *14* (12), 2504–2512. https://doi.org/10.1261/rna.1239308.

(400) Nazarenko, I. A.; Uhlenbeck, O. C. Defining a Smaller RNA Substrate for Elongation Factor Tu. *Biochemistry* **1995**, *34* (8), 2545–2552. https://doi.org/10.1021/bi00008a019.

(401) Bullock, T. L.; Sherlin, L. D.; Perona, J. J. *Tertiary Core Rearrangements in a Tight Binding Transfer RNA Aptamer*; Nature Publishing Group, **2000**; Vol. 7. https://doi.org/10.1038/nsb0600_497.

(402) Hwang, B.; Lee, S. W. Improvement of RNA Aptamer Activity against Myasthenic Autoantibodies by Extended Sequence Selection. *Biochem. Biophys. Res. Commun.* **2002**, *290* (2), 656–662. https://doi.org/10.1006/bbrc.2001.6252.

(403) Rowsell, S.; Stonehouse, N. J.; Convery, M. A.; Adams, C. J.; Ellington, A. D.; Hirao, I.; Peabody, D. S.; Stockley, P. G.; Phillips, S. E. V. Crystal Structures of a Series of RNA Aptamers Complexed to the Same Protein Target. *Nat. Struct. Biol.* **1998**, *5* (11), 970–975. https://doi.org/10.1038/2946.

(404) Carothers, J. M.; Oestreich, S. C.; Szostak, J. W. Aptamers Selected for Higher-Affinity Binding Are Not More Specific for the Target Ligand. *J. Am. Chem. Soc.* **2006**, *128* (24), 7929–7937. https://doi.org/10.1021/ja060952q.

(405) Barrick, J. E.; Corbino, K. A.; Winkler, W. C.; Nahvi, A.; Mandal, M.; Collins, J.; Lee, M.; Roth, A.; Sudarsan, N.; Jona, I.; Wickiser, J. K.; Breaker, R. R. New RNA Motifs Suggest an Expanded Scope for Riboswitches in Bacterial Genetic Control. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (17), 6421–6426. https://doi.org/10.1073/pnas.0308014101.

(406) Peselis, A.; Serganov, A. YkkC Riboswitches Employ an Add-on Helix to Adjust Specificity for Polyanionic Ligands. *Nat. Chem. Biol.* **2018**, *14* (9), 887–894. https://doi.org/10.1038/s41589-018-0114-4.

(407) Knappenberger, A. J.; Reiss, C. W.; Strobel, S. A. Structures of Two Aptamers with Differing Ligand Specificity Reveal Ruggedness in the Functional Landscape of RNA. *Elife* **2018**, *7*. https://doi.org/10.7554/eLife.36381.

(408) Serganov, A.; Yuan, Y. R.; Pikovskaya, O.; Polonskaia, A.; Malinina, L.; Phan, A. T.; Hobartner, C.; Micura, R.; Breaker, R. R.; Patel, D. J. Structural Basis for Discriminative Regulation of Gene Expression by Adenine- and Guanine-Sensing MRNAs. *Chem. Biol.* **2004**, *11* (12), 1729–1741. https://doi.org/10.1016/j.chembiol.2004.11.018.

(409) Fellouse, F. A.; Li, B.; Compaan, D. M.; Peden, A. A.; Hymowitz, S. G.; Sidhu, S. S. Molecular Recognition by a Binary Code. *J. Mol. Biol.* **2005**, *348* (5), 1153–1162. https://doi.org/10.1016/j.jmb.2005.03.041.

(410) Hoshika, S.; Leal, N. A.; Kim, M. J.; Kim, M. S.; Karalkar, N. B.; Kim, H. J.; Bates, A. M.; Watkins, N. E.; SantaLucia, H. A.; Meyer, A. J.; DasGupta, S.; Piccirilli, J. A.; Ellington, A. D.; SantaLucia, J.; Georgiadis, M. M.; Benner, S. A. Hachimoji DNA and RNA: A Genetic System with Eight Building Blocks. *Science (80-. ).* **2019**, *363* (6429), 884–887. https://doi.org/10.1126/science.aat0971.

(411) Sefah, K.; Yang, Z.; Bradley, K. M.; Hoshika, S.; Jiménez, E.; Zhang, L.; Zhu, G.; Shanker, S.; Yu, F.; Turek, D.; Tan, W.; Benner, S. A. In Vitro Selection with Artificial Expanded Genetic Information Systems. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (4), 1449–1454. https://doi.org/10.1073/pnas.1311778111.

(412) Biondi, E.; Lane, J. D.; Das, D.; Dasgupta, S.; Piccirilli, J. A.; Hoshika, S.; Bradley, K. M.; Krantz, B. A.; Benner, S. A. Laboratory Evolution of Artificially Expanded DNA Gives Redesignable Aptamers That Target the Toxic Form of Anthrax Protective Antigen. *Nucleic Acids Res.* **2016**, *44* (20), 9565–9577. https://doi.org/10.1093/nar/gkw890.

(413) Matsunaga, K. I.; Kimoto, M.; Hirao, I. High-Affinity DNA Aptamer Generation Targeting von Willebrand Factor A1-Domain by Genetic Alphabet Expansion for Systematic Evolution of Ligands by Exponential Enrichment Using Two Types of Libraries Composed of Five Different Bases. *J. Am. Chem. Soc.* **2017**, *139* (1), 324–334. https://doi.org/10.1021/jacs.6b10767.

(414) Baker, J. L.; Sudarsan, N.; Weinberg, Z.; Roth, A.; Stockbridge, R. B.; Breaker, R. R. Widespread Genetic Switches and Toxicity Resistance Proteins for Fluoride. *Science (80-. ).* **2012**, *335* (6065), 233–235. https://doi.org/10.1126/science.1215063.

(415) Ren, A.; Rajashankar, K. R.; Patel, D. J. Fluoride Ion Encapsulation by Mg2+ Ions and Phosphates in a Fluoride Riboswitch. *Nature* **2012**, *486* (7401), 85–89.

https://doi.org/10.1038/nature11152.

(416) Merino, E. J.; Wilkinson, K. A.; Coughlan, J. L.; Weeks, K. M. RNA Structure Analysis at Single Nucleotide Resolution by Selective 2'-Hydroxyl Acylation and Primer Extension (SHAPE). *J. Am. Chem. Soc.* **2005**, *127* (12), 4223–4231. https://doi.org/10.1021/ja043822v.

(417) Wilkinson, K. A.; Merino, E. J.; Weeks, K. M. Selective 2'-Hydroxyl Acylation Analyzed by Primer Extension (SHAPE): Quantitative RNA Structure Analysis at Single Nucleotide Resolution. *Nat. Protoc.* **2006**, *1* (3), 1610–1616. https://doi.org/10.1038/nprot.2006.249.

(418) Boniecki, M. J.; Lach, G.; Dawson, W. K.; Tomala, K.; Lukasz, P.; Soltysinski, T.; Rother, K. M.; Bujnicki, J. M. SimRNA: A Coarse-Grained Method for RNA Folding Simulations and 3D Structure Prediction. *Nucleic Acids Res.* **2016**, *44* (7), e63–e63. https://doi.org/10.1093/nar/gkv1479.

(419) Yesselman, J. D.; Eiler, D.; Carlson, E. D.; Gotrik, M. R.; d'Aquino, A. E.; Ooms, A. N.; Kladwang, W.; Carlson, P. D.; Shi, X.; Costantino, D. A.; Herschlag, D.; Lucks, J. B.; Jewett, M. C.; Kieft, J. S.; Das, R. Computational Design of Three-Dimensional RNA Structure and Function. *Nat. Nanotechnol.* **2019**, *14* (9), 866–873. https://doi.org/10.1038/s41565-019-0517-8.

(420) Li, B.; Cao, Y.; Westhof, E.; Miao, Z. Advances in RNA 3D Structure Modeling Using Experimental Data. *Frontiers in Genetics*. Frontiers Media S.A. October 26, **2020**, p 574485. https://doi.org/10.3389/fgene.2020.574485.

(421) Schlick, T.; Pyle, A. M. Opportunities and Challenges in RNA Structural Modeling and Design. *Biophysical Journal*. Biophysical Society July 25, 2017, pp 225–234. https://doi.org/10.1016/j.bpj.2016.12.037.

(422) Nussinov, R.; Jacobson, A. B. Fast Algorithm for Predicting the Secondary Structure of Single-Stranded RNA. *Proc. Natl. Acad. Sci. U. S. A.* **1980**, *77* (11), 6309–6313. https://doi.org/10.1073/pnas.77.11.6309.

(423) Das, R.; Karanicolas, J.; Baker, D. Atomic Accuracy in Predicting and Designing Noncanonical RNA Structure. *Nat. Methods* **2010**, *7* (4), 291–294. https://doi.org/10.1038/nmeth.1433.

(424) Callaway, E. "It Will Change Everything": DeepMind's AI Makes Gigantic Leap in Solving Protein Structures. *Nature*. NLM (Medline) December 1, **2020**, pp 203–204. https://doi.org/10.1038/d41586-020-03348-4.

(425) Kryshtafovych, A.; Schwede, T.; Topf, M.; Fidelis, K.; Moult, J. Critical Assessment of Methods of Protein Structure Prediction (CASP)—Round XIII. *Proteins Struct. Funct. Bioinforma.* **2019**, *87* (12), 1011–1020. https://doi.org/10.1002/prot.25823.

(426) Senior, A. W.; Evans, R.; Jumper, J.; Kirkpatrick, J.; Sifre, L.; Green, T.; Qin, C.;

Žídek, A.; Nelson, A. W. R.; Bridgland, A.; Penedones, H.; Petersen, S.; Simonyan, K.; Crossan, S.; Kohli, P.; Jones, D. T.; Silver, D.; Kavukcuoglu, K.; Hassabis, D. Improved Protein Structure Prediction Using Potentials from Deep Learning. *Nature* **2020**, *577* (7792), 706–710. https://doi.org/10.1038/s41586-019-1923-7.

(427) Singh, J.; Hanson, J.; Paliwal, K.; Zhou, Y. RNA Secondary Structure Prediction Using an Ensemble of Two-Dimensional Deep Neural Networks and Transfer Learning. *Nat. Commun.* **2019**, *10* (1), 1–13. https://doi.org/10.1038/s41467-019-13395-9.

(428) Emami, N.; Ferdousi, R. AptaNet as a Deep Learning Approach for Aptamer–Protein Interaction Prediction. *Sci. Rep.* **2021**, *11* (1), 6074. https://doi.org/10.1038/s41598-021-85629-0.

(429) Sato, K.; Akiyama, M.; Sakakibara, Y. RNA Secondary Structure Prediction Using Deep Learning with Thermodynamic Integration. *Nat. Commun.* **2021**, *12* (1), 1–9. https://doi.org/10.1038/s41467-021-21194-4.

(430) Baker, M. Blame It on the Antibodies. *Nature* **2015**, *521* (7552), 274–276. https://doi.org/10.1038/521274a.

(431) Baird, G. S. Where Are All the Aptamers? *Am. J. Clin. Pathol.* **2010**, *134* (4), 529–531. https://doi.org/10.1309/AJCPFU4CG2WGJJKS.