AFRL-RH-WP-TR-2013-0130

DNA MICROARRAYS FOR APTAMER IDENTIFICATION AND STRUCTURAL CHARACTERIZATION

Jennifer A. Martin
National Research Council

Yaroslav Chushak
The Henry M. Jackson Foundation

Jorge Chavez Benavides
UES, Inc.

Joshua Hagen
Nancy Kelley-Loughnane
Human –Centered ISR Division
Human Signatures Branch

SEPTEMBER 2012

Interim Report

AIR FORCE RESEARCH LABORATORY
711TH HUMAN PERFORMANCE WING
HUMAN EFFECTIVENESS DIRECTORATE,
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the 88th Air Base Wing Public Affairs Office and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-RH-WP-TR-2013-0130 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

_____//signaure//_____

Nancy Kelley-Loughnane, PhD.
Work Unit Manager
Human Signatures Branch

_____//signature//_____

Louise A. Carter, PhD.
Chief, Human Centered ISR Division
Human Effectiveness Directorate
711th Human Performance Wing
Air Force Research Laboratory

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings

REPORT DOCUMENTATION PAGE

1. REPORT DATE *(DD-MM-YY)* 31 09 12	2. REPORT TYPE Interim	3. DATES COVERED *(From - To)* September 2010 to September 2012

4. TITLE AND SUBTITLE
DNA Microarrays for Aptamer Identification and Structural Characterization

5a. CONTRACT NUMBER
FA8650-10-C-6152

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62202F

6. AUTHOR(S)
Jennifer A. Martin[¥*]
Yaroslav G. Chushak[‡]
Jorge Chavez Benavides[†*]
Joshua Hagen[*]
Nancy Kelley-Loughnane[*]

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
H05U (OMSWP004)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
[¥]National Research Council
  500 Fifth Street, NW
  Washington, DC 20001

[‡]The Henry M Jackson Foundation
  6720A Rockledge Drive
  Bethesda MD 20817

[†]UES, Inc.
4401 Dayton-Xenia Road
Dayton, OH 45432

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Materiel Command
Air Force Research Laboratory
711th Human Performance Wing
Human Effectiveness Directorate
Human-Centered ISR Division
Human Signatures Branch
Wright-Patterson Air Force Base, OH 45433

10. SPONSORING/MONITORING AGENCY ACRONYM(S)
711 HPW/RHXB

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)
AFRL-RH-WP-TR-2013-0130

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution A: Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
88ABW-2014-0458; Cleared 11 February 2014

14. ABSTRACT
Aptamers are ideal recognition elements, but integrating aptamers onto a sensor platform has two main challenges: (1) aptamers are selected in solution and, as such, their sequences are not optimized to function when immobilized onto a surface, and (2) the initial aptamer selection process is time consuming. Alternatively, oligonucleotide microarrays are attractive tools for aptamer selection and sensor applications and offer several advantages over solution-based selection. First, selection is performed with oligonucleotides immobilized on a substrate, thus binding is not affected by proximity to the sensor surface. Another advantage is that arrays are commercially available and contain upwards of a million unique sequences which can be tested in a single experiment, greatly reducing the time-frame of selection. While the initial library is composed of fewer sequences than solution-based methods, *in silico* selection provides an optimization step on the starting pool to increase the probability of identifying sequences with strong ligand binding. Therefore, we will improve the efficiency of aptamer selection in a microarray experimental setting using oligonucleotide libraries enriched by pre-screening the potential landscape of binders *in silico*. This work focuses on identification of specific, high-affinity binders for target molecules including biomarkers and chemical warfare agents for future applications in sensors.

15. SUBJECT TERMS
DNA, Immunoglobulin E, Microarray, Aptamer, G-quartet, Thrombin, Cy3

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT:	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON (Monitor)
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U	SAR	55	Nancy Kelley-Loughnane 19b. TELEPHONE NUMBER *(Include Area Code)* N/A

THIS PAGE IS INTENTIONALLY LEFT BLANK.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.0   SUMMARY

Rapid, accurate detection of biological and chemical threats is a critical need of today's military. The key component to such detection is identification of recognition elements specific to the target of interest, which will later be integrated onto a sensor platform. Aptamers have been successfully used as recognition elements for targets ranging from ions to whole cells. However, aptamer-based sensor design has two main challenges: 1) Aptamers are selected in solution and, as such, their sequences are not optimized to function when immobilized on a surface; 2) The initial aptamer selection process is time consuming, which slows sensor development.

Alternatively, oligonucleotide microarrays are attractive tools for aptamer selection and sensor applications, offering several advantages over solution-based selection. First, selection is performed with oligonucleotides immobilized on a solid substrate, thus binding is not affected by proximity to the sensor surface, and more stringent washing conditions can be utilized to increase the efficiency of the system. Another advantage is that arrays are commercially available and contain up to one million unique sequences which can be tested in a single experiment, greatly reducing the time-frame of selection. Microarrays also exclude a requirement for multiple round cycling, polymerase chain reaction (PCR), cloning, and sequencing necessary in traditional aptamer selection methods.

While the initial library is composed of fewer sequences than solution-based methods, *in silico* screening provides an optimization step on the starting pool to increase the probability of identifying sequences with strong ligand binding. Therefore, we will improve the efficiency of aptamer selection in a microarray experimental setting using oligonucleotide libraries enriched by pre-screening the potential landscape of binders *in silico*. In our customized patterning, a library of oligonucleotides was generated by applying sequence constraints based on various parameters such as total length and minimum/maximum number of paired bases, thus resulting in an increased probability of target binding. Two hundred thousand sequences fitting the constraints were synthesized onto custom microarray chips in replicate in order to experimentally test the affinities of the oligonucleotides for specific targets. From the work, we were able to identify potential protein-binding sequences, several possible small molecule-binding sequences, and develop a method for testing oligonucleotides for G-quartet formation by microarray. Generally applicable findings include:

1. Linker length increases will increase fluorescence intensity by extending the probes from the surface, and some minimal length is needed to obtain reliable data.
2. The linker identity (A, C, G, or T) affects fluorescence intensity, with T base linkers proving reliable across all linker lengths and target concentrations.
3. A direct (Cy3-protein) and indirect (biotin-protein followed by Cy3-streptavidin) target labeling method were compared, and it was shown that the indirect method both enhances S/N by orders of magnitude and decreases nonspecific binding to controls.
4. The planar and/or aromatic reporter tags interacted extensively with the DNA.
5. The surface density of the arrays is too low to expect binding from a random library, so starting library optimization must be performed.

6. Purification of free dye from dye-protein conjugates is essential, and the manufacturer's method is insufficient for aptamer identification experiments.

The ideal course of action for those wishing to pursue microarray technology for aptamer identification is to entirely eliminate the need for a tag to be conjugated to the target. Otherwise, it is extremely difficult to determine whether lower-than-control level binding is real target binding or nonspecific tag interaction. Biotin target conjugation is preferable to Cy3-conjugation if necessary. Control experiments where sequences displaying binding to the tag alone can help to alleviate this concern as well. Future researchers should take into consideration how to design the initial library, because this is likely the main determinant of the success of the experiment. Studying the sequences and structures of existing aptamers to the target or similar targets could aid in this pursuit, but does not necessarily guarantee increased success, especially considering that the aptamer need is centered on targets that do not currently have aptamers available.

This work aims at rapid identification and characterization of specific, high-affinity binders for different target molecules including biomarkers and chemical warfare agents for future applications in biosensors, beginning with proof-of-concept studies with a well-known protein aptamer/target pair. The work presented has potential to create technologies that can expedite the identification of aptamer recognition elements for targets including chemical warfare agents, explosives, and biomarkers. In turn, these technologies will enable faster integration into detection devices.

2.0   INTRODUCTION

2.1   Systematic Evolution of Ligands for EXponential Enrichment (SELEX) Aptamer Selection

Aptamers are selected using a method known as SELEX[1,2] depicted in Figure 1[3]. The SELEX procedure begins with 10^{13}-10^{15} *unique* sequences from a chemically synthesized, randomized oligonucleotide library competing for binding to the target. This library consists of sequences designed with two PCR primer regions flanking a random region of 30-50 nucleotides. The sequence space, or number of different base permutations, of this library is defined as 4^N; for a 30-base random region $1X10^{18}$ base combinations are possible. This results in an under-represented unique library where each sequence in the initial 10^{15} appears only once in the starting library pool, and a portion of the sequence space is excluded. Sequences not binding the target are partitioned from binding oligonucleotides, and binders are eluted from the target. This partitioning step (separating the binding sequence from binders) is the main determinant of the efficiency of the selection. The binding sequences are PCR-amplified and converted to single-stranded deoxyribonucleic acid (ssDNA) for the next round of selection. Typically after the first round, researchers may institute a counter-selection step in which sequences binding to a control (such as support matrix or a molecule similar in structure to the target) are removed from solution, and those that do not bind to the control are retained for future SELEX rounds. The process is repeated in cyclical fashion until the final pool is enriched for sequences binding to the target. Following enrichment, the oligonucleotide pool is cloned/sequenced, individual sequences are synthesized to test for target binding, and successful candidates are characterized in terms of binding affinity, target specificity, etc. Frequently, the aptamers are truncated after the preliminary *in vitro* studies to remove regions which do not contribute to binding. This reduces the cost of aptamer synthesis, increases the yield, and may actually act to stabilize the aptamer, resulting in an increased affinity for the target[4]. Researchers may also perform mutations and insertions/deletions to a truncated aptamer to determine consensus regions necessary for binding or identify higher affinity aptamers[5].
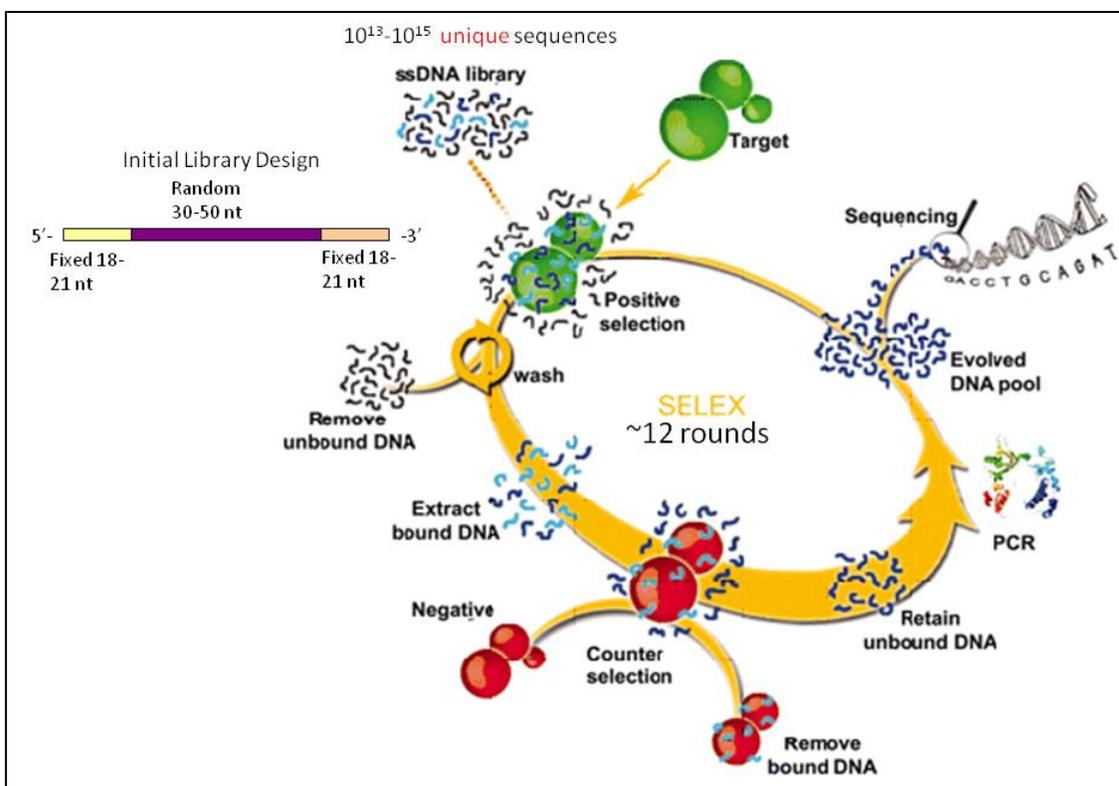
Figure 1: The SELEX Method of Aptamer Selection[3]

SELEX has been applied to targets ranging from ions and small molecules to proteins and whole cells. Our main interests are to select aptamers that bind to biomarkers of physiological states such as stress, fatigue, and vigilance. Many of these biomarkers are small molecule targets and thus pose a challenge for aptamer selection and subsequent biosensor integration. Aside from the abundance of biomarker options, one downstream application of a selected aptamer would be integration into a riboswitch construct which is found in bacteria and other microorganisms. Riboswitches are able to regulate a selected activity by changes induced by binding a small molecule target. (Although riboswitches are ribonucleic acid (RNA) level constructs, once the described methods are successful with DNA, they can later be transitioned to RNA work required specifically for this application.) However, SELEX on small molecule targets is more challenging than standard protein targets because of the difficulty of applying traditional partitioning methods. Most methods rely on a large difference in molecular weight or mobility characteristics between the bound aptamer/target complex and free non-binding sequences[6]. For example, nitrocellulose membranes employ electrostatic interactions between the negatively charged polymer and positively charged proteins to separate binders from non binders. When a mixture of the target and DNA is passed through the filter, the oligonucleotides associated with the protein will be retained by the membrane while the free sequences will be rejected due to their negative charges. Similarly, gel electrophoresis and capillary electrophoresis separate samples based on a change in electrophoretic mobility where the difference may not be enough for separation with small molecule targets depending on the properties of the system. Immobilizing the small molecule can be chemically difficult and can change the binding properties such that binding is reduced or destroyed in free solution.

4

Therefore, there are fewer small molecule aptamers reported in the literature, and the need to rapidly develop aptamers to targets relevant to biosensor work is apparent.

While SELEX has been tremendously successful in attaining high affinity aptamers for a variety of targets, the time and labor investments make high-throughput applications prohibitive. For example, the number of cycles required varies depending on the conditions used in the selection; however, a typical selection requires an average of 12 cycles and a timeline of two-six months, without including initial optimization processes or validation of aptamer candidates[6,7]. This is mainly due to an inherently low partitioning efficiency of SELEX, in which 1 in $1X10^9$ to $1X10^{13}$ sequences are binders[6], but 10^{12} nonspecific binders will be selected in the first round of the selection due to limitations of the separation method, and an inability to differentiate the few high affinity binders from the multitude of nonspecific binders with affinities several orders of magnitude smaller than specific interactions[8]. Therefore, the number of nonspecific binders selected is initially similar or in some cases larger than the amount of specific binders. This is why several rounds of SELEX are required to subtract the nonspecific "noise" from the specific binders through increased stringency of conditions. Additionally, the stringency is often intentionally very low in the first round so the rare binding sequences are not lost during the partitioning process. As described above, each sequence in the initial library is present as an individual in the pool, so loss of a binding sequence in the first round will remove it from contention as a binding candidate at the end of the selection.

A further issue with SELEX lies in the bias introduced into the selection by PCR in each round of selection. PCR has been reported to amplify different oligonucleotides unequally, resulting in an inaccuracy of comparative representation within a pool as the selection progresses[9]. The stability of structures formed by aptamers binding the target may also decrease the efficiency by which the tightest binders are amplified, and the polymerase introduces mutations into the sequences as well. A second form of bias is introduced by the cloning and Sanger sequencing method typical for aptamer identification. The pool is cloned into an appropriate vector (which has a unique set of factors affecting cloning efficiency) and transformed into competent bacterial cells to spatially deconvolute the sequences into individual colonies representing one sequence each, then typically ~100 colonies are randomly chosen from thousands of viable colonies for sequencing[10]. We will use the results of a cortisol selection (previous work performed in this lab) as an example to illustrate the concern with covering such a small portion of the sequence space. At the end of selection, the final cortisol pool contained $1.37X10^{-11}$ moles ($8.25X10^{12}$ molecules) of DNA. This cloning protocol would represent only 1 in $8.25X10^{10}$ total sequences for the cortisol selection. It would likely reflect the most abundant sequence present, but may not report those that have artificially lower numbers due to factors such as PCR bias or cloning efficiency. Cho et. al. have shown that removing the cloning bias and using next-generation sequencing methods that sample a larger percentage of the pool (10^5-10^7 sequences) produced aptamers with affinities 3-8-fold higher than a similar selection performed with cloning and Sanger sequencing.[10] Thus, reducing or eliminating the reliance on PCR and cloning would give a higher probability of identifying the true highest affinity sequences.

2.2 Microarray-Based Aptamer Identification

We propose to use microarrays as a faster, more efficient method of performing aptamer identification and characterization studies. DNA microarrays function by identifying locations of fluorescence (from fluorescently labeled target) and correlating them with the position of known sequences covalently synthesized on the array (Fig 2). A fluorescently labeled target is added and an image of the array is captured using a microarray scanner (Agilent High-resolution Microarray Scanner currently on loan from Agilent Technologies). Alternatively, the target can be biotinylated and allowed to interact with the array, followed by a short Cy3-streptavidin incubation. Our initial studies were performed with protein targets, but we are investigating options to transition the technology to small molecule biomarkers.
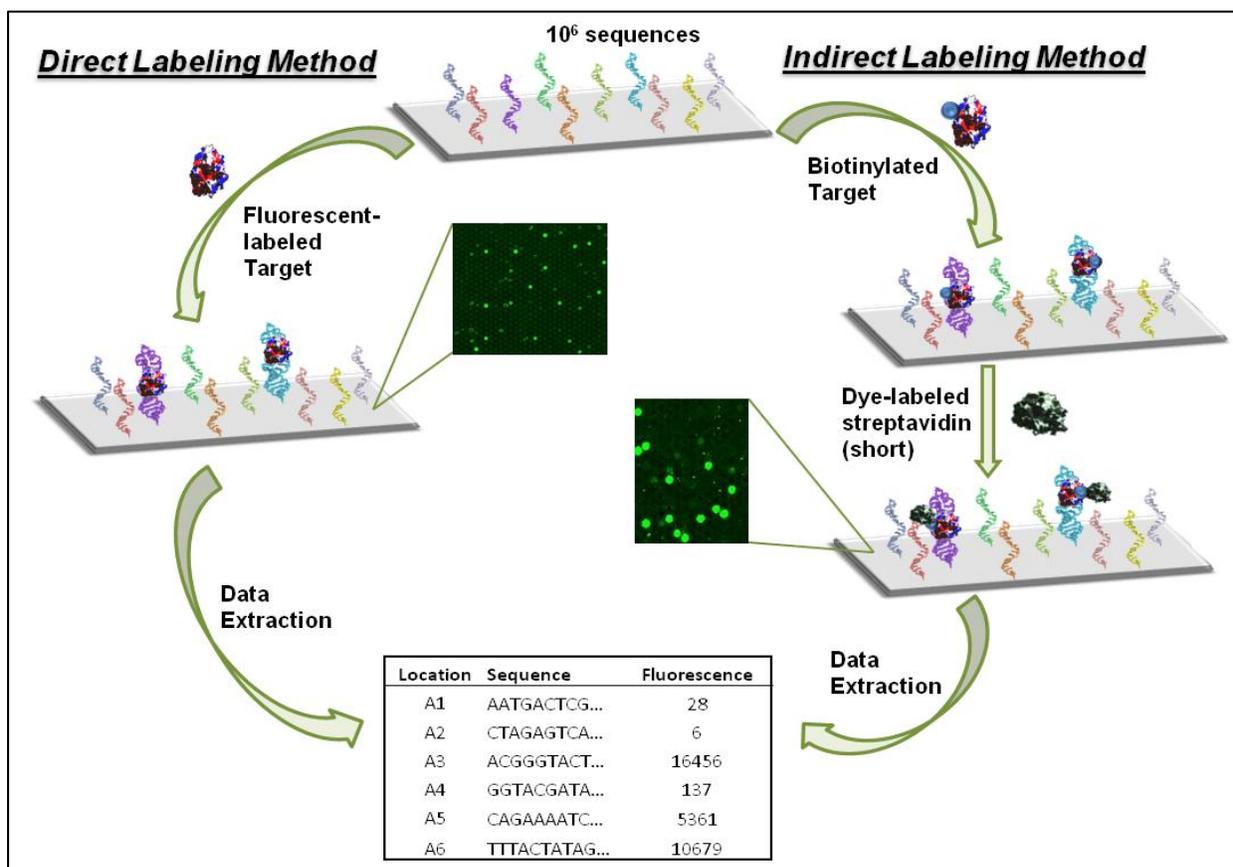
Figure 2: Methods of Microarray Aptamer Identification

The main advantage of using microarrays for this purpose is the time saved by the ability to complete studies of many truncates and mutations in parallel rather than individually. One microarray experiment can be completed in less than a day, with data analysis requiring another 1-2 days. This characteristic is a function of the improved partitioning efficiency available by covalently linking the sequences to the surface. Higher stringency conditions can be applied to identify sequences with better binding properties. Also, no PCR, cloning, or sequencing is necessary due to design of the microarray with sequences in known locations. The arrays are fully customizable so the user can define the exact sequences of interest produced on the arrays. Cost savings are also significant in that one array at a price of ~$600

(Agilent 1X1M array) can test up to a million sequences simultaneously, as opposed to ~$135 for one sequence HPLC purified from Integrated DNA Technologies (IDT) for traditional analysis.

Prior to today's array technology, some of the issues related to the use of microarrays in aptamer studies were associated with the techniques used to deposit the oligonucleotides[11]. First, the cost of synthesizing thousands of sequences for experiments was prohibitive. Second, spot size reproducibility, specifically related to the conditions used for spotting the arrays (humidity, buffer used, etc), was variable. To avoid these issues, recent developments in microarray technologies provide the option of synthesizing the oligonucleotides from the surface of the slides[12]. This significantly reduces the cost of manufacturing the chip and offers exquisite control over the oligonucleotide synthesis. Several commercial companies (MYcroarray, Agilent, Nimblegen) have essentially eliminated these issues and offer fully customizable array product options up to 80 bases long. However, this technology has been mostly developed for gene expression studies that function by direct hybridization of the fluorescent-tagged genomic fragments to cDNA on the microarray. Therefore, protocols and methods for aptamers must be developed and optimized by the individual researcher with challenges specific to this application.

Significant work has been performed by different groups in the use of aptamers for target detection in custom-made microarray settings. Many parameters will affect the aptamer response that are not as significant in cDNA hybridizations, including: probe density, proximity of the binding site to the surface, immobilization orientation (5' or 3' end) and even the buffer used in the studies[13,14]. It has been confirmed that known DNA aptamers can be immobilized on a microarray and conditions to preserve their activity can be found[11], and the conditions used for these studies should mimic the conditions used during the initial solution-based selection process[15]. Additionally, the response of an aptamer in solution can be significantly different than when it is attached to an array surface[16]. Several studies report differences in aptamer affinity when comparing microarray and solution-based dissociation constant measurements[14,17]. In most cases the affinities of these aptamers were significantly lower on the microarray than when tested in the solution-based conditions used during their selection[15]. So applying a microarray modified aptamer to riboswitch conditions would likely *improve* the aptamer function relative to the array, and when applied to more traditional biosensor platforms, it would ensure binding activity is retained.

In the initial SELEX introduction, we touched on the fact that post-SELEX modifications are often performed to truncate then introduce mutations to improve function or study some of the structural properties of aptamer/target binding. These studies give rise to a number of advantages, including a reduced cost of synthesis, improved yield, and potential improvement in the affinity of the aptamer for the target. Standard methods of aptamer truncation include radioactive labeling and fragmentation, followed by synthesis and purification of active sequences[5]. A less labor intensive method involves analysis of the individual sequences of a homologous family obtained after pool sequencing and alignment. However, this requires the synthesis, purification, and affinity characterization of many sequences in a family, also a time-consuming process. Similarly, study of mutations requires a second complete SELEX to identify higher affinity aptamers. Biesecker et. al. (1999) selected aptamers to human complement C5 with dissociation constants 20-40 nM after 12 rounds[18]. They then used the predicted secondary structure of the chosen best aptamer as a template to perform a second 8-

round "biased" SELEX to improve the dissociation constant of the highest affinity sequences to 2-5 nM with full functional activity.

Fischer et. al. has shown that microarrays are capable of exploring truncations of the well-studied immunoglobulin E (IgE) aptamer[19]. They were able to confirm the importance of a 21-base consensus sequence and document the minimum truncated sequence demonstrating fluorescence intensity similar to the original structure. They found that the consensus sequence was generally confined to the loop in a step:loop structure, and the stem was more amenable to manipulation than the consensus loop. Katilius et. al. introduced a series of single, double, and triple mutations to the IgE aptamer by microarray[17]. They were able to improve binding relative to the initial aptamer by manipulating the bases in the stem region, and showed that single base mutations can completely destroy target binding. Collectively these two studies concluded that extending the aptamers from the array surface enhances target binding, the fluorescence intensity is proportional to the concentration of target added, and higher relative fluorescence values between probes on the same array indicates higher affinity.

Three main works in the literature have actually used arrays to perform aptamer selection rather than simply employing them for detection or structural examination[20-22]. The starting libraries consisted of 10^2-10^4 initial sequences and evolved aptamers through *in silico* genetic algorithms in multiple chip generations (rounds). Each of them either used naturally fluorescent targets, utilized a target with well-studied aptamer binding motifs, or took into account the characteristics of known aptamers for library design.

As a whole, this body of work from the literature makes several important points advocating the proposed applications: 1) It is possible to detect targets using arrays; 2) Array performance will vary with the conditions used, and should be as close to the selection conditions as possible; 3) The immobilization of the aptamers in an array setting likely will not negatively impact sequences applied to riboswitches and would benefit traditional biosensor detection platforms; 4) Microarrays can be used for truncation/mutation studies and provide valuable insights into aptamer structure; 5) Probes can be evaluated in massively parallel format, and their relative fluorescence intensities can be compared to indicate higher affinity binders; 6) Aptamer selection is possible if the starting library is "designed" to contain potential binders.

Some possible technical challenges for applying this method to aptamer modifications for newly selected targets, where the behavior of the sequence is not well-documented in the literature, include nonspecific dye interactions and small molecule detection. Most dye systems used to label proteins for microarray work possess conjugated π-systems reported to interact with DNA directly [20]. This would make it difficult to differentiate between sequences that are binding the actual protein, or those that instead interact with the dye. The Gold team reported using photoactive base substitutions to covalently cross-link proteins to the arrays, followed by stringent washing steps to remove nonspecific binders[23]. However, current commercial microarray manufacturers do not offer these base substitutions in their array synthesis. A feasible method of biotinylating the protein followed by a short dye-labeled streptavidin incubation has been shown to demonstrate less nonspecific binding compared to a direct dye labeling[20,24]. Also, while methods exist to modify proteins with biotin or fluorescence tags largely without affecting their binding properties, doing so to a small molecule biomarker target may prove more difficult. These tags often alter the binding properties of the target, and their comparable size to the target molecule tends to dominate the

interaction with DNA. The chemistry to covalently link the tag to the target will vary depending on the target structure.

A major drawback is that the highest density arrays have a maximum of ~10^6 sequences. This is in stark contrast to SELEX methods, which evolve from an initial library of ~10^{15} sequences. However, the success of a selection is more dependent on the number of binders present in the initial pool than a raw total number of sequences, and combinatorial drug-screening libraries identify binders from 10^3-10^6 compounds[25]. Therefore, biasing a DNA or RNA starting library with a maximum of 10^6 sequences to have a higher probability of containing aptamers should enable microarray identification of binders.

2.3   *In Silico* Starting Library Design

In this work we designed a number of different starting library patterns in order to circumvent the microarray density problem and increase the chances of identifying aptamers binding the target. Aptamers usually show complex secondary and tertiary structures composed of a number of features like loops, stems, and multi-arm junctions that interact with each other to stabilize the structure of the aptamer-complex target[26]. It has been determined that the probability for a sequence to bind a target improves with increasing the structural complexity[27]. This means unstructured sequences or oligonucleotides that form simple structures have reduced potential to show any type of function. In a typical SELEX experiment, a random pool of oligonucleotides is used to search for binders for a particular target. Constituents of a random pool have mostly unpaired regions combined with short (low stability) stem-loop structures, and the probability of finding more complex, high-affinity aptamers in the starting random library is very low[28]. This was experimentally confirmed by Davis and Szostak who selected aptamers from a mixture of fully random and partially structured RNA libraries[29]. They identified 6 sequence families of aptamers. Four of these families came from the designed library while two families were from the random pool. It was also found that aptamers with the highest affinity were selected from the designed library of sequences with an engineered multi-base (high stability) stem-loop.

Recently, there have been several attempts to design enhanced initial pools of sequences for selection of high-affinity aptamers. Luo *et a*l., have developed Random Filtering and Genetic Filtering methods to increase the number of five-way junctions or to design a uniform structure distribution in the starting RNA/DNA pools[30]. One of the aptamers selected by SELEX from the designed structural libraries displayed higher affinity for ATP than a previously selected low-complexity aptamer while other aptamers showed weaker affinity. Thus higher complexity itself does not always lead to better affinity. In another approach Ruff *et a*l., designed a patterned library to enhance the formation of stem-loop structures[31]. For the three performed SELEX selections, the patterned library substantially outperformed the unpatterned library in two cases and performed at least as well as the unpatterned in the third case.

Here, we present a set of related libraries that have been compared by their abilities to produce structurally enriched sequences by applying a series of constraints (see Section 3.4). This set is composed of four different patterned libraries of 50nt DNA sequences used as an initial pool for aptamer selection:

Liu - $(RY)_3$-N_4-$(RY)_4$-N_3-$(RY)_4$-N_4-$(RY)_4$-N_3-$(RY)_3$

PT1 - $(RRYY)_2$-N_4-$(RRYY)$-N_3-$(RRYY)$-N_4-$(RRYY)$-N_3-$(RRYY)$-N_4-$(RRYY)_2$

PT2 - $(RRYY)_2$-N_4-$(RRRYYY)$-N_4-$(RRRYYY)$-N_4-$(RRRYYY)$-N_4-$(RRYY)_2$

PT3 - $(RRYY)_2$-N_4-$(RY)_3$-N_4-$(RY)_3$-N_4-$(RY)_3$-N_4-$(RRYY)_2$

R=(A,G), Y=(T,C), N=(A,C,G,T)

The first library (modified from Liu) has a pattern similar to one proposed by Ruff *et al*[31] from the Liu group. It consists of alternating purines (R=A,G) and pyrimidines (Y=T,C) separated by completely random regions of N_4 and N_3 . The second pattern (PT1) was designed to maximize the number four-way junctions as depicted in Figure 3[30]. Pattern Liu has 14 completely random bases while pattern PT1 has 18 random bases. Pattern PT2 and PT3 were designed with the same number of random bases, 16, but PT2 allows for three of the same purine or pyrimidine in a row while PT3 forces alternation of internal purines and pyrimidines.
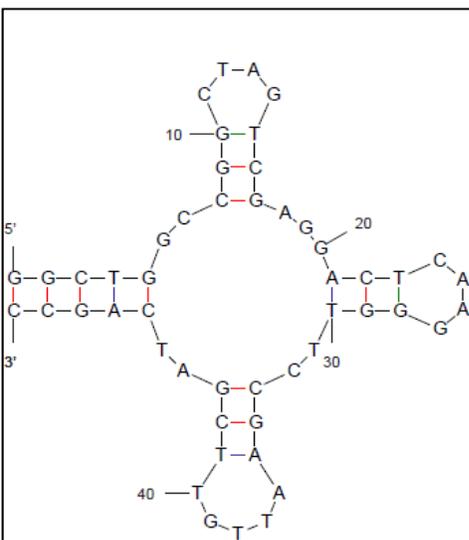
Figure 3: Example of Four-Way Junction Structure Generated with Pattern PT1

In this work, we used *in silico* patterned libraries to "design" the starting library to possess a higher probability of containing binding sequences on a microarray in order to discover new IgE aptamers. IgE is part of a well-characterized aptamer/target pair with positive and negative controls ideal for initial studies. Once validated, methods will be modified to target biomarkers of interest. The utility of a direct- and indirect- labeling method were compared. We were also able to gather information indicating that microarrays can be used to probe aptamer tertiary structures and can even be applied to small molecule aptamer identification. ***The aim of the work is to serve as proof-of-principle studies that microarrays can speed up the aptamer identification and characterization process to rapidly provide recognition elements to various biomarkers of interest for "plug-and-play" integration into riboswitch or traditional biosensor detection platforms.***

3.0 METHODS, ASSUMPTIONS, AND PROCEDURES

3.1 Chemicals and Equipment

IgE (Fitzgerald), Illustra NAP-25 desalting columns and Cy3 Mono-Reactive Dye Pack (GE Healthcare), NanoDrop (Thermo Scientific), nuclease free water (Gibco). Microarray equipment: custom 8X15k and 1X1M DNA microarrays, 8X15k and 1X1M gasket slides, ozone barrier slides, hybridization chambers, scanner cassettes, hybridization oven, and High-resolution Microarray Scanner (all Agilent). Slide rack and wash dishes (Shandon), Kimtech polypropylene wipes (Kimberly-Clark). FluoReporter Mini-biotin-XX Protein Labeling Kit (Invitrogen), Streptavidin-Cy3 Conjugate and HABA/Avidin Reagent (Sigma), dialysis cassettes (3.5k MWCO, Thermo Scientific).

Buffers: Binding [PBSMTB]- 1X PBS (8.1 mM Na_2HPO_4, 1.1 mM KH_2PO_4, 2.7 mM KCl, 137 mM NaCl, pH 7.4) + 1 mM $MgCl_2$ + 0.1% Tween-20 and 1% BSA; Washing [PBSM]- 1X PBS (8.1 mM Na_2HPO_4, 1.1 mM KH_2PO_4, 2.7 mM KCl, 137 mM NaCl, pH 7.4) + 1 mM $MgCl_2$; Rinse [R]- 1/4 dilution of PBSM and nuclease free water. RB Wash Buffer: 20 mM HEPES, 300 mM KCl, 5 mM $MgCl_2$; RB Binding Buffer: RB wash buffer + 0.1% Tween-20 + 1% BSA. IgE binding buffer without K^+: 8.1 mM Na_2HPO_4, 137 mM NaCl, pH 7.4) + 1 mM $MgCl_2$ + 0.1% Tween-20 and 1% BSA; Washing- 8.1 mM Na_2HPO_4, 137 mM NaCl, pH 7.4 + 1 mM $MgCl_2$

3.2 Direct Labeling Method- Immunoglobulin E (IgE) Cy3-Labeling

IgE was diluted with 0.1 M sodium carbonate buffer (pH=9.3) to 1 mg/mL. One mL IgE dilution was added to one Cy3 dye pack and incubated 30 min. Free dye was separated from IgE-bound dye by purification with a desalting column. Dye to protein ratio (D/P) was calculated using the manufacturers' instructions with the aid of Nanodrop UV/Vis detection.

3.3 8X15k Microarray Optimization

The 8X15k microarrays were employed to optimize conditions based on the response of several reported IgE binding and nonbinding sequences. Initial library patterning was reserved for 1X1M microarrays, which could accommodate all 50k sequences in replicate (although 8X15k arrays were "filled" with either 7k or 5k PT1 sequences in duplicate). Blocking with PBSMTB was performed on the DNA microarray loaded into a 50 mL conical tube for 1 h at room temperature. Slides were disassembled in PBSM buffer and rinsed for 5 min at room temperature using a stir plate. The slides were quickly edge-tapped to remove excess buffer. Seventy µL protein (variable concentrations) in PBSMTB was loaded onto the gasket slide then incubated with each of the 8 DNA arrays for 2 hrs at 20°C in a hybridization chamber/hybridization oven. Slides were disassembled in PBSMTB buffer and washed for 3 min in a separate PBSM buffer with the slide rack and stir plate, then the slide rack was transferred to 1/4 PBSM buffer/water for 1 min using a stir plate. Slides in the slide rack were then dipped in nuclease free water to remove any remaining salt and washed for 1 min with stir plate. The rack was slowly withdrawn from the water to promote a drier surface, the back of the slide was wiped with ethanol and then placed in a 50 mL conical tube with a polypropylene wipe at the bottom and centrifuged at 4150 rpm for 3 min. The microarray was loaded into a scanner cassette and covered with an ozone barrier slide before scanning.

3.4 Microarray Starting Library Design

UNAFold software was used to screen DNA from each pattern and determine which sequences were folding. Code was written in Perl to extract sequences which fit a series of constraints: the 1st base must pair with the 50th base, the total number of unpaired bases: 10<unpaired<30, there must be a minimum of two ≥4-unpaired base stretches, T=25°C, $[Na^+]$= 100 mM, $[Mg^{2+}]$= 5 mM. Fifty-thousand sequences that fit the constraints were incorporated onto the 1X1M microarray chip with a T_{10} spacer in replicates of 4.

3.5 1X1M Microarray

The basic protocol was the same for the 1X1M arrays as for the 8X15k arrays described above. The only difference in methodology was that the gasket slide held a 750 µL volume instead of 70 µL.

3.6 Indirect Labeling Method- IgE Biotinylation

Biotinylation of IgE was carried out according to the manufacturer's instructions, and purified using the provided spin column. The biotin/protein ratio was calculated to be 0.6 according to the HABA/Avidin test.

3.7 8X15k and 1X1M Microarrays (Indirect Labeling Method)

The 8X15k and 1X1M microarrays were analyzed in the same manner as the direct labeling method with some exceptions. Following incubation with biotin-IgE, slides were disassembled in PBSMTB buffer and rinsed for 3 min on a shaker. Slides were transferred to PBSM buffer for 3 min also on a shaker. Then 0.5, 10, or 500 nM Cy3-streptavidin (Cy3-SA) was loaded into the gaskets and incubated for 2 min. The washing steps were the same as the direct labeling method.

3.8 8X15k Arrays

The standard control array was performed as in Section 3.3, except the samples were variable: 100 nM samples: Cy3-IgE D/P= 2.3, 4.8, and 11, Cy5, Cy3-BSA, TurboRFP; 10 nM samples: Cy3-SA (to mimic the biotinylated experiments), Cy3-estradiol antibody (the sample was dilute following purification). The extended control array was also similar to Section 3.3, except all Cy3 only arrays were performed with 50 nM sample, and all Cy3-proteins were at 100 nM. When the buffers were varied, the arrays were incubated in the buffer of interest before sample loading, and slides were disassembled in water then quickly transferred (3 seconds) to PBSM for the subsequent washing steps so the PBSM wouldn't cause DNA folding with the samples remaining on the arrays.

For the final D/P analysis, 100 µL sample was dialyzed for 48 hours through a 3.5k MWCO dialysis cassette. Dialysate was MilliQ water exchanged after 2 and 4 hours. Dye to protein ratio (D/P) was calculated using the manufacturers' instructions with the aid of Nanodrop UV/Vis detection.

3.9 Data Analysis

The arrays were scanned using Agilent Scan Control software. Images (TIFF) were generated using 20-bit imaging at 3 µm (for 1X1M arrays) or 5 µm (8X15k arrays). Data was extracted using Agilent Feature Extraction software version 10.7.3.1. Mean fluorescence intensity and other values of interest of replicates were determined using code written in Perl.

3.10 Surface Plasmon Resonance (SPR)

Biacore studies were carried out on a CM7 chip with neutravidin custom immobilized on the surface. The surface was activated with a mixture of 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide (0.2 M) and N-hydroxysuccinimide (0.05 M) for 420 sec at 10 µL/min. Neutravidin (10 mg/mL) was dissolved in HyClone water then diluted 1/10 in 10 mM sodium acetate buffer (pH= 4.5) and added at 30 µL/min for 210 sec. The surface was blocked with 1 M ethanolmine for 600 sec at 5 µL/min. Biotinylated 4A018 aptamer was heated to 95°C, then introduced at 5 µM in 5 mM $MgSO_4$ for 180 sec at 30 µL/min. Samples were diluted to 27.3 µM in HEPES buffer from Biacore (10mM HEPES, 150mM NaCl, 0.05% Surfactant P20) and dialyzed into HEPES buffer overnight. The samples were serially diluted to appropriate concentrations, then introduced onto the chip at 30 µL/min for 30 sec. Data analysis was performed with BIAevaluation software.

4.0   RESULTS AND DISCUSSION

4.1   Starting Library Design

As described above, a major challenge of using microarrays for selection is the limited surface density of the arrays (10^6) compared to SELEX (10^{15}). Therefore, we designed a starting library populated with sequences that have a higher probability of binding than a random library. Again, the Liu pattern was adapted from literature, PT1 was designed to maximize the number of junctions, PT2 is a variation on the Liu pattern that allows for internal and peripheral base repeats, and the PT3 variation allows for base repeats at the periphery only. The selection rate, or number of sequences analyzed until one fits the constraints, was highest for the Liu pattern (Fig. 4A), but the structural diversity was highest for PT3 (Fig. 4B). This shows that the Liu pattern has the least number of sequences in the starting pool that fit the criteria, and the structures formed are less diverse than most of the less stringent patterns. This could mean that the Liu pattern will not generate as many binders, or it could mean that the stringency of the pattern actually evolves higher affinity sequences. In contrast, PT3 generated sequences that fit into the constraints more frequently, but it also generated the most diverse structures. More complex and diverse structures in an initial pool has been suggested to lead to better likelihood of identifying a high-affinity aptamer[25,27], so these results indicated that PT3 may be the "ideal" pattern of those studied under these conditions based on a simplified in silico analysis.
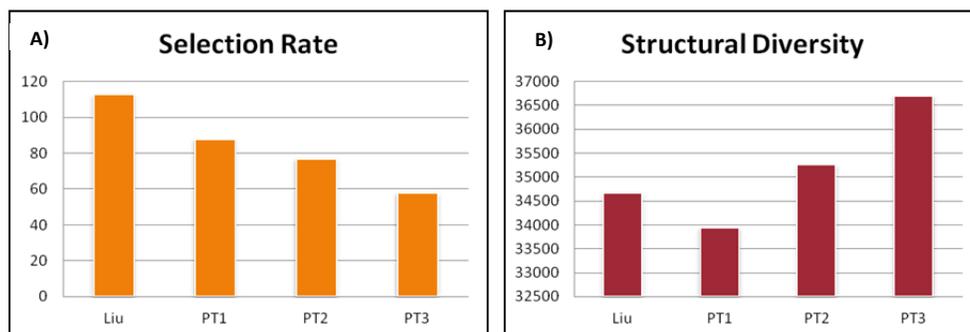
Figure 4: Starting Library Design

4.2   Direct Labeling Method

IgE was labeled with Cy3 (D/P= 11.8) according to standard methods to facilitate detection by the scanner. The IgE positive and negative controls were used to optimize the method and [Cy3-IgE] in 8X15k array format prior to aptamer identification from the library (Fig 5A). The positive controls showed a clear increasing response with increasing linker length and [Cy3-IgE] (Fig 5B), while the negative control intensity (SA= mutated streptavidin aptamer) did not increase with linker length, but showed a slight background increase with increasing [Cy3-IgE] (Fig 5C). We also saw that the identity of the linker, A, C, G, or T, will affect the fluorescence intensity (Fig 5D), with T linker the most consistent performer across various linker lengths. An interesting phenomenon occurred when K_d was plotted against linker length (Fig 5E). The K_d values are more variable at low linker lengths, then appear to plateau at a certain length. This also corresponds to an increase in the "goodness of fit" of the data (Fig 5F), suggesting

that a minimal linker length is required for accurate data. At short lengths, the Cy3-IgE may be better able to interact nonspecifically with the surface or BSA, or other components of the buffer, or the aptamer is too rigidly confined to fold into the proper binding structure. The negative control, D17-4SM17, appears to have the best K_d, but the fit of the data is variable across lengths, and never reaches acceptable limits (>0.95). The curves do not reach saturation, which affects the error associated with the curve fit of the data. Based on the results, 100 nM Cy3-IgE was applied to the 1X1M feature selection array (Fig 5G), and several probes were identified demonstrating comparable fluorescence to the positive controls. These top binding probes were sent for BIACore (Biosensor Tools, Salt Lake City, UT) testing, but did not demonstrate any binding activity (Fig 5H). A similar array protocol was then used with Cy3 dye alone. The top 100 ranked (mean fluorescence intensity) Cy3 binders were cross-referenced with the top 500 Cy3-IgE binders (Table 1). The results showed that the top Cy3-IgE binders were more likely binding the dye than the IgE, which accounts for the lack of BIACore activity (Table 1, bold probes).
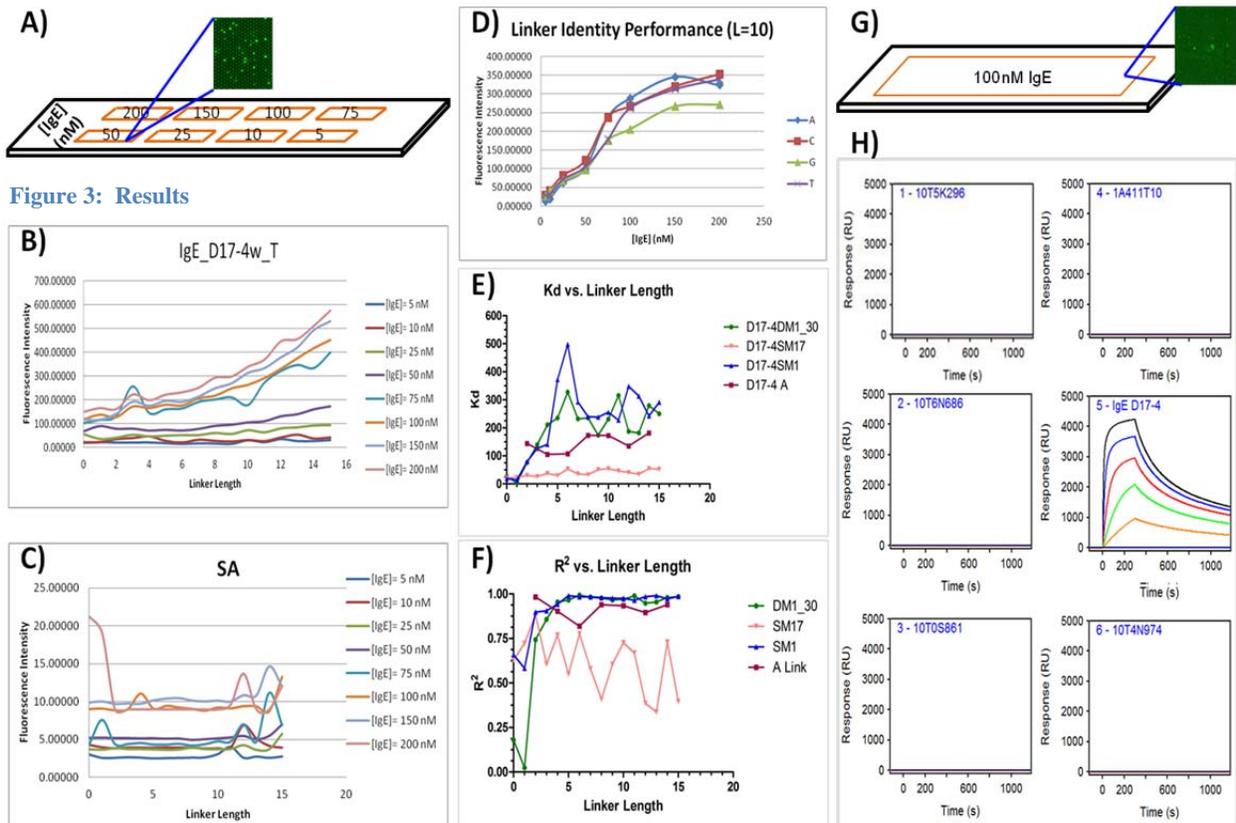
Figure 3: Results

Figure 5: Results from Direct Labeling Method

Table 1: Cy3-IgE Fluorescence Intensity Rank versus Cy3 Dye Rank

IgE Rank	Probe ID	Cy3 Rank
344	5D837	95
60	4L137	91
149	6L573	90
15	1A411	87
54	2D539	85
32	0L574	79
59	5E589	78
150	9M745	75
43	8G954	73
155	6E768	63
63	9O285	55
30	2O166	53
48	5A255	44
56	3M083	42
33	8L716	39
24	9N665	32
12	2K864	24
35	5O660	23
23	9J881	19
26	4H813	16
21	9B301	13
27	2K027	12
71	3M814	11
7	0I505	10
16	5K476	9
36	1B194	8
17	1B463	4
6	1M969	3
10	8N621	2
20	4L969	1
15	**1A411**	**87**
4	**6N686**	**69776**
3	**0S861**	**1257**
2	**4N974**	**3959**
1	**5K296**	**1894**

4.3   Indirect Labeling Method

A new scheme was developed for microarray work reported in the literature to decrease the amount of nonspecific binding. In the method, the target is biotinylated and incubated with the array as in the previous method (Fig 2). However, after this step, a short Cy3-SA incubation is performed before data extraction. Two optimization experiments (Fig 6A-B) were carried out similar to the Cy3-IgE experiments discussed above. It was clear that binding was due to the biotin-IgE instead of the Cy3-SA by visual comparison of the arrays (Fig 6C-D). Some minimal binding to the positive IgE controls was observed in the Cy3-SA only array, but this is believed to be a function of the communal washing procedure causing biotin-IgE availability rather than Cy3-SA binding. The signal was orders of magnitude higher when10 nM biotin-IgE was included, but the negative control values did not change significantly (see data in Tables A1-2; blue designates negative controls, green designates positive controls). The raw signal for similar [IgE=10 nM] was enhanced 750X in the biotin labeling method (Fig 6F) compared to the direct (Cy3-IgE) labeling method, and the LOD was improved from ~500 pM (direct) to ~4 pM (indirect). Figure 6G shows a maximum S/N enhancement of ~700X at [biotin-IgE]= 25 nM from the indirect method compared to the direct dye-IgE method (Fig 6H). We selected [biotin-IgE]= 25 nM, [Cy3-SA]= 10 nM for the aptamer identification experiment conditions due to the conditions displaying the highest S/N. These results prove that the indirect labeling method has enhanced signal and lower nonspecific binding to negative controls. We are in the process of testing lower D/P in order to rule out binding site occlusion or resonance energy transfer as the source of lower signal in the direct labeling method[32].

Two 1X1M aptamer identification experiments with the biotin-IgE were performed with different biotin-IgE concentrations to rule out experimental artifacts (Fig 7A-B). See Figure A1 (Appendix) for actual images from each location. The patterned DNA was ranked by mean fluorescence intensity in each experiment, and compared with how the same sequences ranked in the Cy3 only and Cy3-IgE trials (Fig 7C). (Blue= PT1; Red= Liu; Green= PT2; Purple= PT3; black boxes indicate sequences that appear to be good biotin and/or Cy3 binders.) Many of the top biotin-IgE binders were also top Cy3 and/or Cy3-IgE binders, indicating that these sequences are probably binding the small molecule tags rather than the IgE protein. By expanding the search parameters past the top 25 biotin-IgE binders, we were able to identify two top candidates (9D551 and 6K736) that show high ranking biotin-IgE binding, but poor Cy3 binding (Fig 7D). Although the fluorescence intensity of these probes is not nearly as high as the positive controls, they may represent new IgE binding motifs, and will be tested for binding by surface plasmon resonance (SPR).
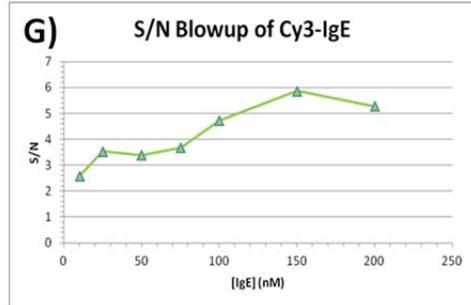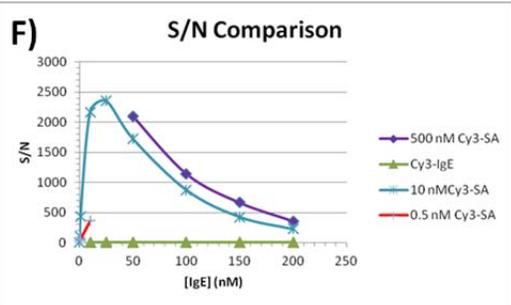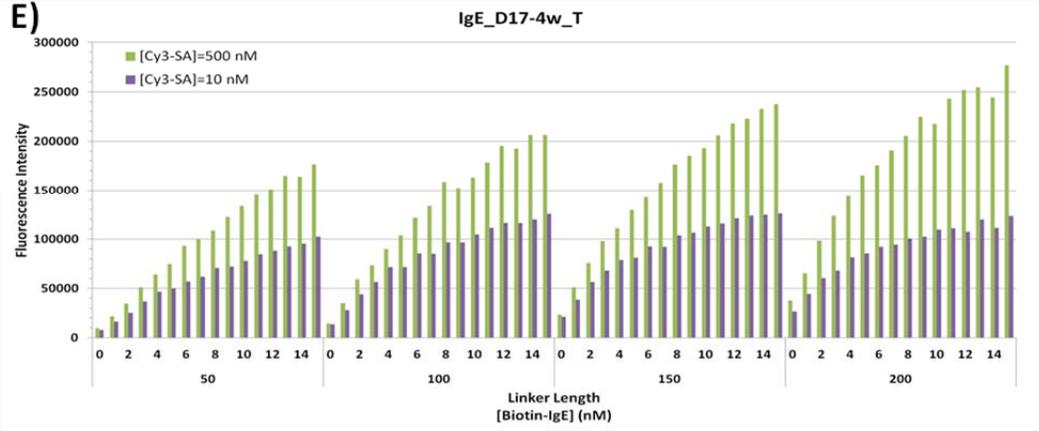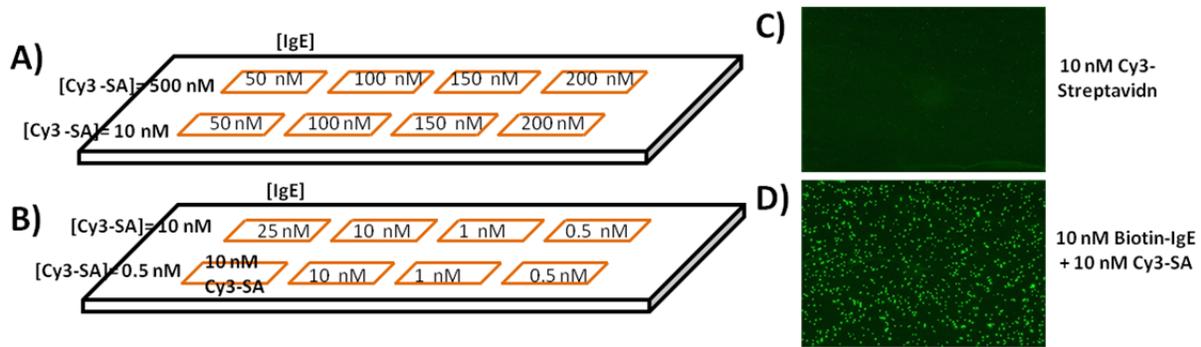
**Figure 6: Results from Indirect labeling Method Optimization**

C)

ID	B-IgE Rank	B-IgE2 Rank	Dye Rank	Dye-Ige Rank
0S861	1	1	1257	3
5K296	2	3	1894	1
1M969	3	4	3	6
2K864	4	8	24	12
0O881	5	198616	147633	112407
7A713	7	2	7137	5
0D855	8	31051	134471	192847
0L204	9	178	170511	122082
0C049	10	148254	186683	177749
9D551	11	35	168126	112104
6Q568	12	7583	90331	159883
2M867	13	2506	3405	137
8M838	14	12790	29460	193114
7M919	15	67352	171013	32841
7E554	16	2192	167054	195008
4N974	17	5	3959	2
9M832	18	121617	167206	199326
7O209	19	61	120	348
3C330	20	61980	190495	199697
2T910	21	131987	184333	199882
6K736	22	90	155014	57827
0T772	23	58590	35348	25228
3T046	24	272	22814	187752
9G129	25	2624	121537	10929

D)

ID	B-IgE Rank	B-IgE Rank2	Dye Rank	Cy3 IgE	Structure	G Pairs
9D551	10	35	168126	112104	Y	1
6N686	46	44	69776	4	Wand	5
6K736	21	90	155014	57827	Wand	2
3O274	45	94	123519	198204	Pencil	2
7L355	42	105	121026	197344	Y	3
0L204	8	178	170511	122082	Pencil	1

Figure 7: Indirect labeling Method Aptamer Identification Results

We also compared the sequences on a number of parameters to determine if any trends were present which could indicate the best options for SPR binding testing. After analyzing parameters similar to Figure 7D for a larger population of sequences (high/low ranked dye, Cy3-IgE, or biotin-IgE binders), some clear trends emerged for some of the conditions. PT2 sequences tended to be good biotin and dye binders, holding ~1/2 or 1/3 of the top 25 fluorescence intensities for the two tags. G-quartet structures are known to bind small planar or aromatic molecules, meaning PT2 sequences forming these structures may be binding the tags instead of the protein[4,25,33]. After analysis of pattern design, we observed that PT2 has a higher probability of forming G-quartet structures due to the potential to contain the highest number (7) of G repeats, three of which are R groups with a 50% chance of containing G. Also, a "wand" secondary structure was indicative of good B-IgE binders and poor dye and dye-IgE binders, making sequences that form these structure good candidates for IgE binding (Fig 8). In contrast, a "Y" secondary structure binds all categories well (biotin-IgE, dye, and dye-IgE), indicating dye/biotin interactions may be causing the signal. It agrees with prior reports that the more complex Y structure forms more pockets for small molecule binding than a stem/loop wand structure[30]. Also, PT1 appeared in all good and poor dye and B-IgE binders, P2 bound both dye and B-IgE sequences well and contained the most G repeat sets in the structures studied, and P3 tended to be good B-IgE binders and poor dye/dye-IgE binders. The Liu pattern did not have enough representatives in any category to distinguish any trends. Sequences beginning with GGTTGG(CC or TT) were frequently observed among top dye and biotin binders (similar to the first several bases of the GQ-forming TFBS). Sequence 6N686 begins with this motif and did not bind IgE in the SPR studies. Therefore, a wand structure of P3 with few G repeat sets and minimal dye binding would be an ideal candidate, although none of the sequences in Figure 7D perfectly fit this mold. The trends only serve as generalizations, and a broader picture of individual sequences should be obtained before summarily ruling sequences out for reasons other than known dye binding.

4.4   G-Quartet Aptamer Studies

An interesting observation noted in all testing formats was that the thrombin fibrinogen binding site aptamer (TFBS) was demonstrating significant fluorescence signal. In the Cy3-IgE experiments, the TFBS signal was several times higher than the signal for the positive IgE controls (Fig 9A). The TFBS was the only control to fluoresce above background in the Cy3 only trial, showing that it is not binding the protein (Fig 9B). TFBS also displayed binding in the biotin-IgE tests (Fig 9C), although the binding was reduced in comparison to the Cy3-IgE. This may be due to a higher D/P in the Cy3-IgE experiments relative to the biotin-IgE experiments, or a decreased affinity for the planar, but not aromatic biotin. The observation aids in validating the indirect labeling method as demonstrating lower nonspecific binding. Based on these results, the controls were extended in the next biotin-IgE array design (Fig 9D-F). Figure 9D shows the response of TFBS with 25 nM IgE and 10 nM Cy3-SA with different linker identities. The A and T linkers bind similarly, but the C linker increases then sharply decreases at LL=10, and the G linker does not show any change across linker length. The C linker probably is increasing in fluorescence as the probe is moved further from the array surface, but decreases when it becomes long enough to interact with the G-quartets and disrupt binding. Adding a G-linker appears to change or destroy the structure of the quartet so binding does not improve at any linker length. This further shows the importance of the G-quartet for binding small molecules. In Figure 9E, minimal binding is observed for any linker identity after a short incubation with 10 nM Cy3 only. This shows that TFBS can also bind biotin, and biotin is the cause of the enhanced fluorescence in the biotin-IgE format. The riboflavin binding aptamers (RB), also G-quartets, were incorporated into this extended design. Figure 9F shows that RB 3-tiered (RB3) binds biotin above background (10 nM biotin-IgE + 10 nM Cy3-SA), but not as well as TFBS. The 2-tiered RB (RB2) did not show any appreciable binding. The diminished binding of the RB's is due to their reported intolerance to high sodium concentration, more severe for RB2 than RB3, while the effect has not been reported for TFBS[34]. Also, these studies were performed in the IgE aptamer binding buffer, which can alter the folding of aptamers not selected under these conditions. An enhanced effect for both Cy3 and biotin binding would likely be shown for TFBS and the RB's in their respective selection buffers. ***Regardless, microarrays may serve as an initial test or even alternative to NMR and crystallography studies to confirm aptamer G-quartet structure [4]***.
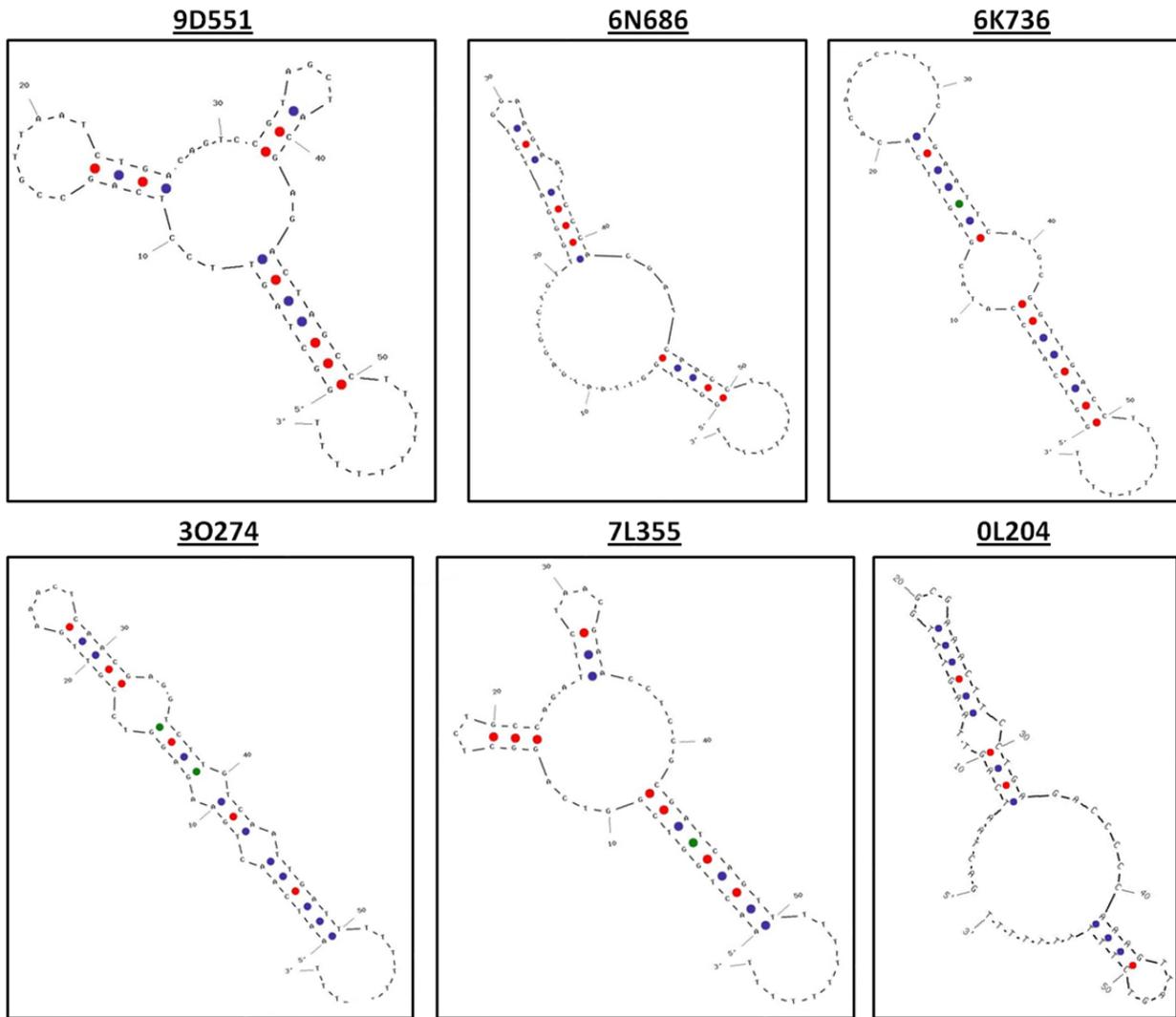
Figure 8: Secondary Structures of Potential IgE Binders
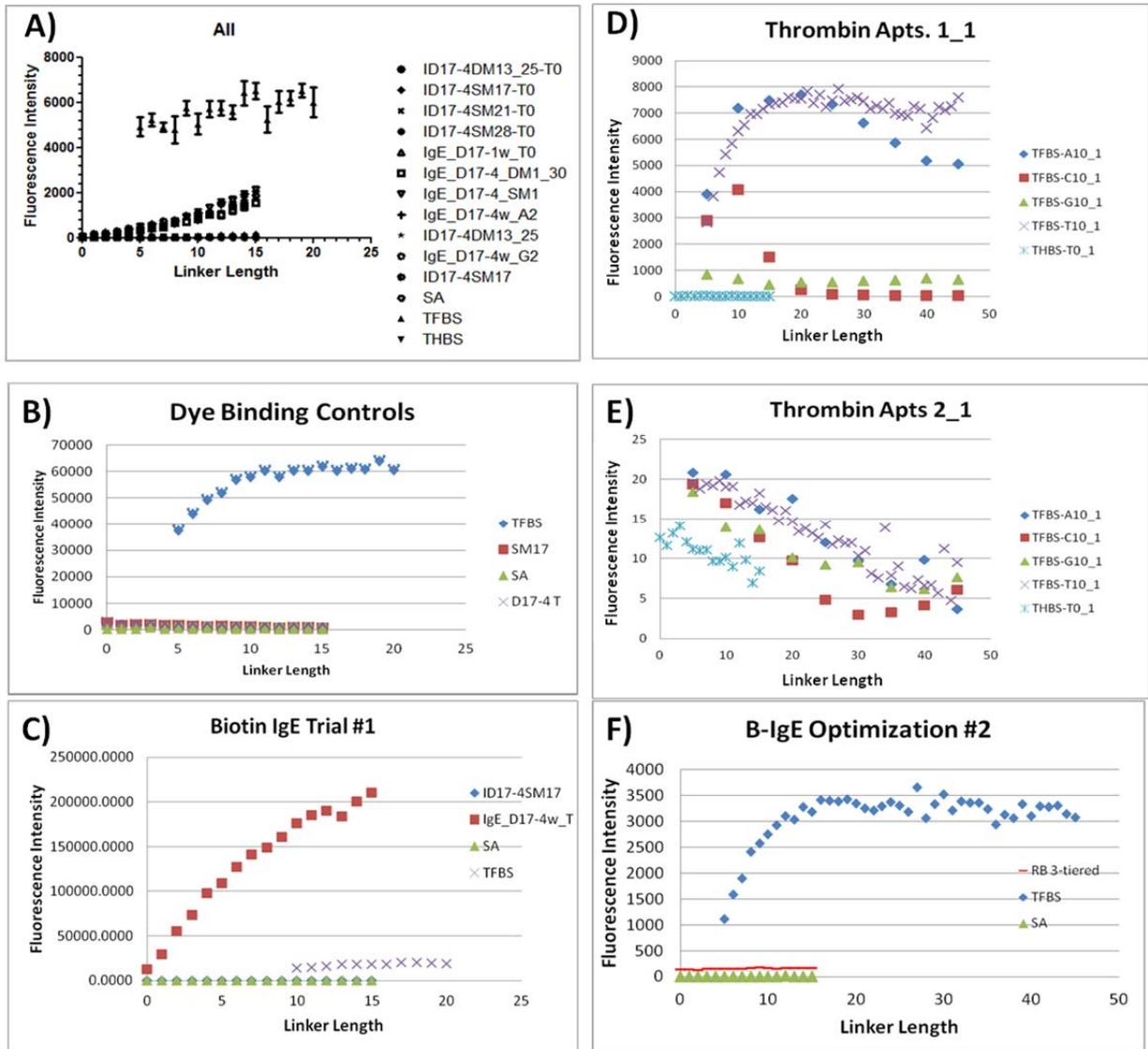
Figure 9: G-Quartet Binding Data

4.5   Elucidation of Biotin and/or Cy3 Aptamers

When comparing individual probe performance across Cy3-IgE, Cy3 only, and biotin-IgE trials similar to Figure 7, several sequences were unearthed that may warrant further study. Figure 10 depicts a number of probes that may bind to both biotin and SA, biotin only, dye only, or unbound dye only based on the mean fluorescence intensity rank for each experiment. It is undetermined whether the probes would bind other similar dyes or tags, or other dye-labeled proteins, or function in other buffers at this time. They may even be G-quartet or other higher order structures that bind small molecules somewhat indiscriminately. The probes could be utilized in any way an aptamer could be applied, including serving as positive controls in future microarray work. Aptamers could be generated for other similar small molecule targets by constructing an *in vitro* library of analogous DNA structures to act as part of a dye displacement assay for "label-free" target detection.

ID	Bio1Rank	Bio2 Rank	Dye Rank	Dye-Ige Rank
0S861	1	1	1257	3
7A713	7	2	7137	5
5K296	2	3	1894	1
1M969	3	4	3	6
4N974	17	5	3959	2
2K864	4	8	24	12
7O209	19	61	120	348
6N686	46	44	69776	4
9D551	11	35	168126	112104
6K736	22	90	155014	57827
0L204	9	178	170511	122082
3O274	45	94	123519	198204
7L355	42	105	121026	197344
8N621	14848	872	2	10
4L969	67250	7377	1	20
5K476	9796	10359	9	16
1B194	56448	77908	8	36
1B463	60479	136026	4	17
6D170	127184	85106	6	128200
5E742	47977	90705	7	42163
0E794	52587	98963	5	156522

Both
Biotin
Dye
Free Dye

Figure 10: Potential Biotin and/or Cy3 Aptamers

4.6   Interaction with G-quartets (GQ), New Thrombin Aptamers, and Other Tests

sheeThe final two microarrays were employed to answer questions related to how the samples interact with G-quartet aptamers, and whether or not the D/P (dye to protein ratio) is affecting the intensity of the Cy3-IgE/IgE aptamer interaction.   One goal, as previously stated, is to propose that micorarrays can be used as a simpler method of screening ligands for G-quartet folding by simply adding a planar, aromatic fluorescent dye such as Cy3.   Previous relevant findings are summarized in Figure 11 and include: 1) Cy3-IgE has a much lower S/N (signal-to-noise ratio) than labeling the protein with biotin, then adding Cy3-SA (streptavidin) reporter (Fig 11A, same as Fig 6F). We were not sure if this was due to an actual benefit of the method, or because of binding site occlusion and/or FRET (fluorescence resonance energy transfer) from the high D/P (~11). 2) Probe 17-1, although reported as a positive control, demonstrated ~10X lower fluorescence intensity overall, and was basically unresponsive at linker length <10 at 100 nM Cy3-IgE (Fig 11B). 3) Cy3 alone, Cy3-IgE, and biotin-IgE (lower extent) all bound to the TFBS (thrombin fibrinogen binding site) G-quartet (GQ) aptamer. Cy3-SA did exhibit ~7X above background levels TFBS, but it did not increase dramatically (over 25X background for others) across increasing linker length (enhanced availability) like the other samples (Figs 11C-E). (Cy3-SA also had very low backgrounds compared to the other samples.) Since IgE has been confirmed in the literature to not interact with TFBS, we assumed the Cy3 and biotin were the cause of binding since they are small and planar and/or aromatic. If this was the case, why wasn't Cy3-SA also binding the TFBS?

4) Cy3 alone and Cy3-IgE showed an increase in intensity with linker length for THBS which was much smaller in magnitude than TFBS increase. Biotin-IgE and Cy3-SA had a higher THBS value than background (~2X), but it did not increase across linker length (Figs 11F-G).
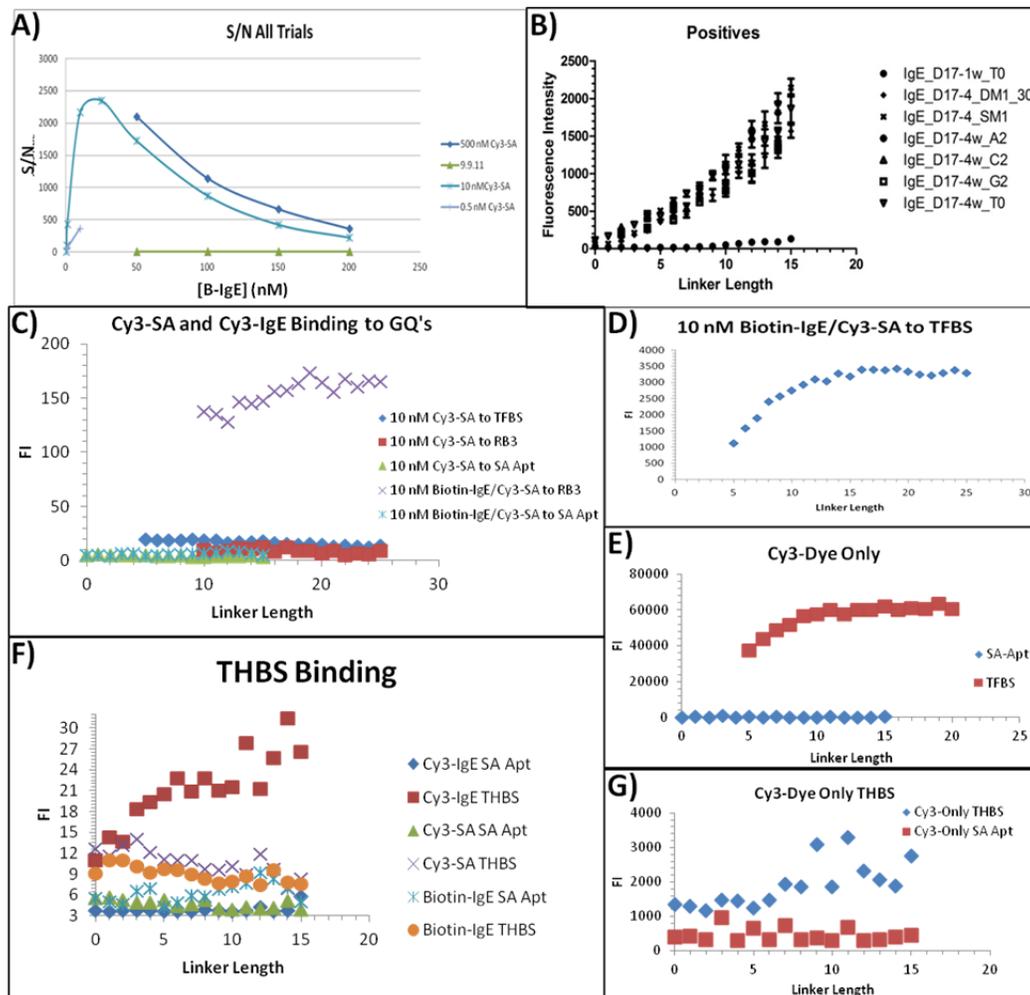
Figure 11: Summary of Pertinent Previous Work

Figure 12A-C shows the reference key for the arrays, and the questions asked by the samples of the standard control array and the extended control array. The extended control array had some additional GQ-forming aptamers added to it, specifically the 2 and 3-tiered riboflavin aptamers (RB2 and RB3). If Cy3 can interact with GQ's, we should see a positive result for the RB's, and a negative result for non-GQ forming structures. Figure 12D is a table representative of the different reasons we felt streptavidin may not be binding the TFBS in the same manner as Cy3 and Cy3-IgE. Other proteins, BSA, thrombin, IgE, Estradiol antibody, and a scrambled neuropeptide Y (NPY) sequence were evaluated to confirm or rule out factors such as MW, pI, D/P, glycosylation, and NaAz (sodium azide) interference. The two microarray experiments will need to be compared to each other for a full picture of the relevant parameters.

The most straightforward set of comparisons involve evaluating different D/P on IgE to determine whether binding site occlusion/FRET/protein quenching is resulting in the decreased S/N for the IgE aptamers compared to the biotin-IgE/Cy3-SA method. High protein labeling has

been shown to both interfere with binding sites and cause FRET which lowers fluorescence signal[32]. Cy3 and biotin have not been seen to interact with the IgE aptamers (non GQ's) above background, so binding is dictated by the protein properties (FRET, binding site availability) rather than more dye reporting more intense binding. We produced Cy3-IgE D/P= 2.3 and 4.8 in addition to 11 prepared before. The D/P was modulated by changing the amount of dye added to a constant amount of IgE. D/P= 2.3 gave higher fluorescence intensity values than D/P=11 and 4.8 for the positive controls (Fig 13A-C, compare arrays (2.3) 1_1, (11) 1_2, and (4.8) 2_1 in legends). Interestingly, clear binding was observed with probe D17-1 at D/P= 2.3, which only showed minimal binding in past experiments with D/P 11. Binding is more easily observed due to some relation between decreased FRET and/or increased site availability. D/P=2.3 also displayed lower negative control signal than D/P=11 (Figs 13D-F). The S/N values for D/P 2.3, 4.8, and 11 were 186.6, 4.9, and 5.3 vs. over 900 for the biotin-IgE/Cy3-SA experiments at the same [IgE]. Figure 13F validates mutated streptavidin aptamer shown in previous experiments to display very low background. In conclusion to point 1, *we were able to improve S/N by manipulating the D/P of Cy3-SA.* The enhancement for the biotin-IgE method may be due to the multiple Cy3 present on the Cy3-SA (~6) reporter, or the low biotin/protein (0.8) providing more binding site availability. Although we could further decrease the D/P of Cy3-IgE to ~0.8, it is unlikely the enhancement would be as much the biotin-IgE/Cy3-SA because any increased binding site availability would be counteracted by more of the binding proteins without a Cy3 reporter. Regardless, the S/N enhancement from the biotin-IgE/Cy3 IgE method makes this method superior to the Cy3-IgE method.

At the completion of the experiment, after ruling out all of the possibilities in Figure 12D as the cause of the 4 questions addressed above, sample purity from free dye was identified as a culprit. The three Cy3-IgE D/P's were dialyzed through a 3.5k membrane, and absorbance values were taken following dialysis. A large amount of free dye was initially present in the D/P 2.3 and 11 samples, because the D/P after dialysis was 0.5 (2.3) and 2.5 (11). D/P 4.8 seemed more pure than the other two samples, with a final D/P= 4. This explains why initial D/P= 4.8 showed the lowest fluorescence intensity values, as it actually had the highest D/P once the samples were extensively purified. This correlates with work by Hahn on IgG that concludes that higher D/P will reduce the fluorescence yield due to resonance energy transfer >6 D/P[32]. Possibly the threshold is slightly lower for IgE protein. The conclusion is that *the manufacturer's protocol for free dye separation is not sufficient for aptamer microarray work because DNA (especially GQ's) can interact with free dye present, resulting in false positives for binding.* Previous aptamer array work was less concerned with this because they were either selecting for an intrinsically fluorescent molecule, or because the main aim was for detection of fluorescent-labeled targets rather than selecting new aptamers[14,16,17,19,22]. New aptamers were selected for Cy3-thrombin[20] but thrombin binds GQ structures, and probably with a much higher affinity than that of the dye for the GQ. So the specific thrombin interaction would likely take precedence over the lower affinity dye binding.

A)

KEY

| 1_1 | 1_2 | 1_3 | 1_4 |
| 2_1 | 2_2 | 2_3 | 2_4 |

B) STANDARD CONTROLS

| Cy3-IgE D/P 2.3 | Cy3-IgE D/P 11 | RFP- Will a fluorescent protein bind TFBS? (No Cy3) | Dialyzed Cy3-SA- Is NaAz the problem? |
| Cy3-IgE D/P 4.8 | Cy5- Will other dyes bind TFBS? | Cy3-BSA- Similar pI to Cy3-SA | Cy3- estradiol antibody- Similar MW to IgE |

C) EXTENDED CONTROLS

| Cy3 (Pack)- Will it bind RF aptamers? | Reduced Cy3- Will it bind w/o NHS? | Cy3-NPY- Will a peptide bind? D/P? | Cy3 (Pack) in RB buffer- RF2 binding? TFBS in new buffer? |
| Biotin/Cy3 Pack mix- Competition? | Cy3-IgE- Will it bind RB aptamers? | Cy3-thrombin- Another protein bind? D/P? S v ns-binding? | Cy3-IgE in IgE buffer w/o KCl- Destroy Gq to prove cause of binding? |

D)

| Sample	MW	pI	D/P	Glycosylated?	NaAz?
Cy3-Streptavidin	60000	5-6	3-9	No	15 mM
Cy3-BSA	66430	4.7	2.6	No	N/A
Cy3-Thrombin	36700	7.0-7.6	0.8	Yes (5%)	N/A
Cy3 IgE	200000	5.2-5.8	11/4.8/2.3	Yes (12%)	N/A
Cy3-Estradiol Ant.	150000	?	9.9	Probable	N/A
Cy3 NPY (Scrambled)	4253.6	6.76	5.8	No	N/A

Figure 12: Microarray Experimental Setup

The remainder of the work for the first array will examine how the various samples interact with TFBS, THBS, RB2, and RB3 to answer the questions posed in Figure 12, either confirming that GQ's can interact with dyes, or attempting to explain why the Cy3-SA does not interact with TFBS in the same manner as Cy3-IgE. First, we tested whether Cy5 dye would bind to the G-quartet forming TFBS aptamer, since prior work showed that Cy3 dye (both alone and including Cy3-IgE) could bind TFBS (Fig 11C-E). Cy5 requires scanning in the red channel, where other arrays were scanned using the green channel specific for Cy3. Figure 14A clearly shows binding of the Cy5 dye, ***confirming that G-quartet structures can bind Cy5 as well as Cy3 dye and biotin under the given conditions.*** Figure 14B displays the green channel binding of the samples to TFBS (note Cy5 displays minimal signal in green channel, 2_2). All arrays were above the SA background (Fig 13F), but the IgE's and Cy3-BSA show an increasing trend similar to the free dye in Figure 14A. The Cy3-streptavidin sample was dialyzed because previous results showed that Cy3-streptavidin did bind TFBS above background, but did not increase with increasing linker length like the free dye and Cy3-protein samples. If the dye was the result of TFBS signal, Cy3-streptavidin should bind as well, and we speculated it may have been because of sodium azide in the sample. The dialyzed Cy3-SA (commercially ordered) has elevated fluorescence (similar to previous results), but not increasing intensity across linker length. Thus removal of NaAz did not improve binding.

Cy3-BSA did bind TFBS and increase across linker length, and it is similar in MW and pI to Cy3-SA and not glycosylated, effectively ruling these out as the cause of minimal Cy3-SA interaction. The binding profile of Cy3-Estradiol antibody (Fig 14D) is remarkably similar to dialyzed Cy3-SA. These were the only two proteins to be analyzed at 10 nM rather than 100 nM: 1) in order to rule out 10 nM Cy3-SA as the binding agent in the biotin-IgE/Cy3-SA experiments, and 2) Purification of Cy3-Estradiol antibody yielded an extremely dilute sample. The low concentration of the two samples may be the cause of the abnormal binding effects; however it also exposed purity from free dye as the explanation (as discussed above). Cy3-SA is highly purified by the manufacturer, and Cy3-estradiol antibody (and Cy3-IgE D/P= 2.3) was purified on a longer purification column than Cy3-BSA, Cy3-IgE D/P= 11, and Cy3-IgE D/P= 4.8. A longer separation column would be expected to produce better resolved Cy3-protein and Cy3 dye bands. Cy3-IgE D/P= 4.8 may potentially be purified better than Cy3-IgE D/P=11 despite using a similar column, otherwise we would expect the free dye from the former to bind TFBS in a similar intensity if free dye were the only factor. More likely, there is probably some interplay going on between free dye affinity v. protein-bound dye affinity v. D/P v. quenching. ***It appears that TFBS binds free dye with a higher affinity than protein-bound dye***, likely because of steric crowding or accessibility issues when the protein is involved. Finally, it is unclear why TurboRFP demonstrates enhanced TFBS and THBS signal. It could be related to a real response of the protein for the GQ binder or, it may be due to sample bleed over from Cy3-IgE D/P= 11 array during the washing process.
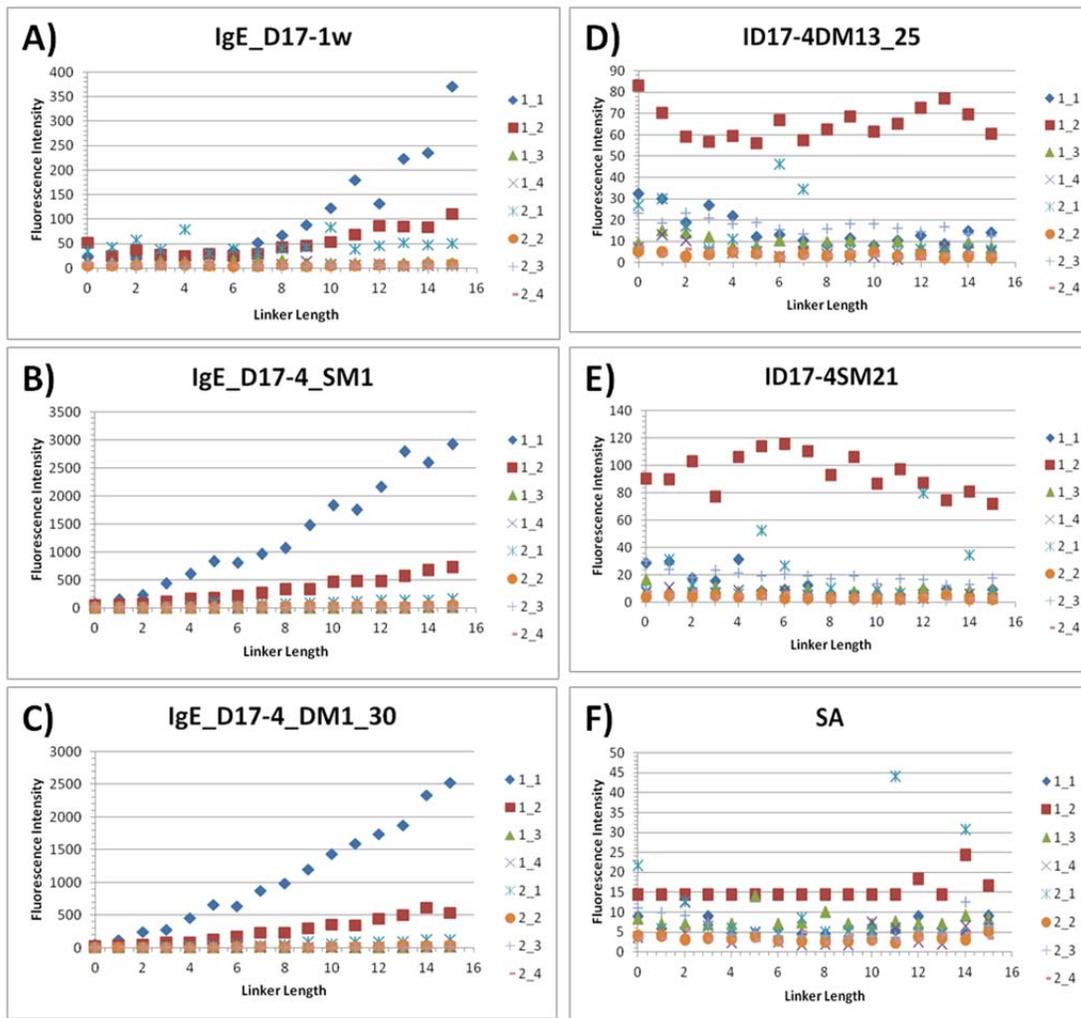
Figure 13: Interaction of the Positive (left) and Negative (right) Controls with Cy3-IgE at Different D/P

Figure 15 summarizes the conclusions from the questions we aimed to answer with the Standard Control experiments. We did find that lower Cy3-IgE D/P gave fluorescence enhancement, but still significantly lower than the biotin-IgE/Cy3-SA method. Low D/P enhanced the signal observed from the weaker binder 17-1. Cy5 dye can bind TFBS in addition to Cy3 dye, strengthening the case that dyes can be used to probe GQ structure on microarrays. Finally, we ruled out MW, pI, glycosylation, and NaAz as the cause of the decreased Cy3-SA binding profile, while adding purity from free dye as the main contributing factor. Future work aims at further characterization of these interactions, and of interactions with other known aptamers.
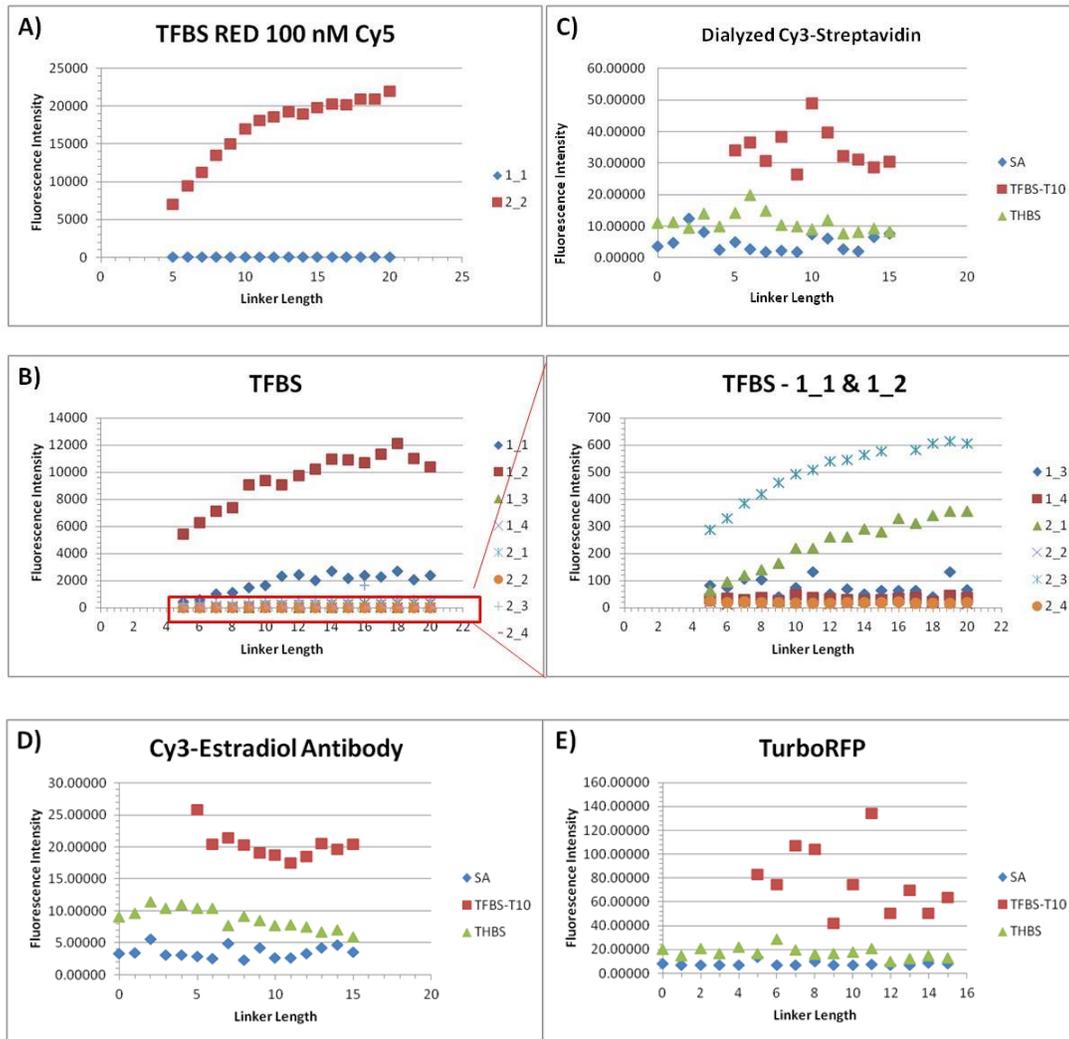
Figure 14: Interaction of the Samples with TFBS and/or THBS
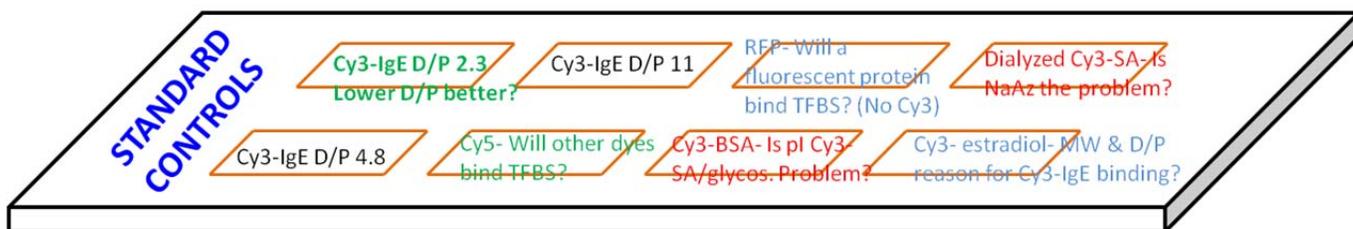
Figure 15: Pictorial Summary of Results
(Green= Yes, Red= No, Light blue= Inconclusive)

In the extended controls, the riboflavin 2- or 3-tiered GQ aptamers were included to see if the dye could bind other GQ's, and to probe some different aspects of GQ folding/function. In Figure 16A we can see **Cy3 can bind RB3 similar to (slightly better than) TFBS, confirming that the GQ structure is needed for dye binding.** Note that the RB2 aptamer does not bind, which is logical because it was reported[34] that RB2 is less stable than RB3 in a Na$^+$ based buffer, and requires a higher K$^+$ content for structural stability. In Figure 16B, the data in Figure 16A for TFBS binding of a new dye pack was compared to the binding of a dye pack stored in water for several months. The "old" dye pack was expected to hydrolyze rapidly, wherein the N-Hydroxysuccinimide (NHS) moiety of the dye would be removed and the dye would now terminate in a carboxylic acid group. **Interestingly, the "old" dye pack did not bind the TFBS GQ as strongly as the new pack, indicating that the NHS moiety may aid in binding.** RB3 binding (not shown) was also similarly diminished but above background. The samples were diluted to different concentrations and analyzed in a fluorescence plate reader for Figure 16C. The fluorescence binding curves were similar for the two samples, showing that dye photobleaching was not the cause of the lower apparent binding. We are looking into methods to evaluate the structure of the old dye pack to confirm whether the NHS moiety has been hydrolyzed. Another experiment involved mixed Cy3 and biotin together to see if biotin could compete with Cy3 for binding of the GQ's (Fig 16D). The fluorescence intensity values were not significantly diminished by biotin, but the Cy3 may have much higher affinity for the DNA since it is both a planar and aromatic compound.
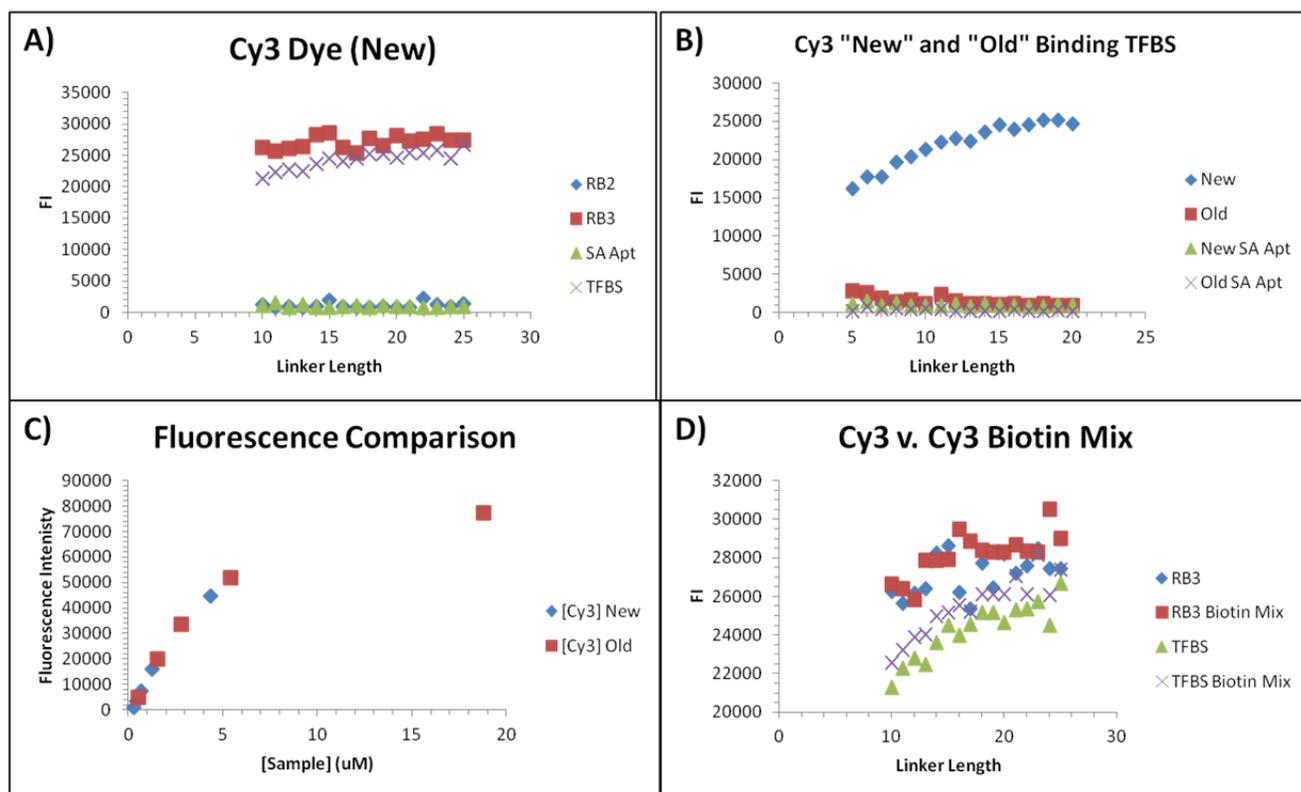

Figure 16: Cy3 Binding to Multiple GQs

Cy3-IgE, Cy3-NPYsc (scrambled neuropeptide Y), and Cy3-thrombin were also evaluated on the extended control chip. Cy3-IgE and Cy3-NPYsc exhibited enhanced TFBS binding, but also displayed elevated RB3 binding, both likely due to the free dye component (Fig 17A-B). IgE has been reported to NOT demonstrate binding to TFBS in previous studies with PBS buffers containing different components than this work[35,36], but the Ellington group has reported cross reactivity of TFBS with IgE in a HEPES buffer[14]. So there may be some potential for cross-reactivity depending on the buffer composition. Cy3-thrombin was a relatively pure sample, which displayed GQ binding for both TFBS and RB3 which exceeded that of the dye alone (Fig 17C). *This indicates that thrombin could specifically bind the RB3 GQ aptamer in addition to TFBS.* It is difficult to say that the higher fluorescence intensity of RB3 is due to a higher affinity without the benefit of a binding curve. *Clearly the added flexibility of a longer linker length seems to aid more with the much smaller dye binding the TFBS (or the suspected free dye in Fig 17A) rather than the specific thrombin interaction appearing saturated across all lengths.* The THBS binding of thrombin is also far more intense than that of dye only binding (Fig 11F), differentiating specific thrombin interactions from Cy3 interaction (Fig 17D). *The added flexibility for THBS sequences with longer linkers drastically increased signal.*
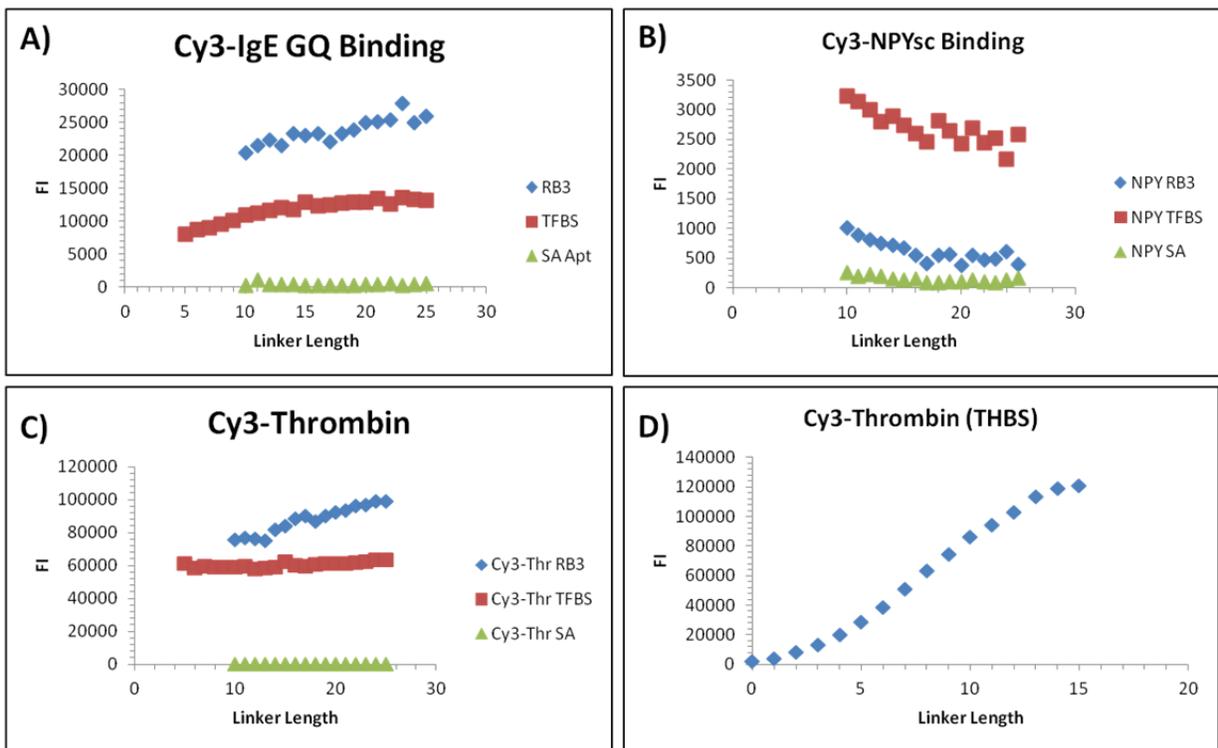
Figure 17: Protein Interactions with Multiple GQs

The final two arrays on the chip were meant to probe the GQ aspect of dye binding. By using the original buffer the RB2 and RB3 aptamers were selected in (20 mM HEPES, 300 mM KCl, 5 mM MgCl$_2$, + 0.1% Tween-20 and 1% BSA pH= 7.5), if GQ formation is the cause of dye binding, inducing the RB2 to fold into the GQ structure would promote binding to Cy3 (Fig 18A). *Binding to RB2 was restored in the RB aptamer binding buffer, validating the hypothesis that GQ formation is the cause of dye binding.* Interestingly, the RB3 aptamer displayed *decreased* fluorescence in the RB buffer compared to the normal chip conditions ((8.1

mM Na_2HPO_4, 1.1 mM KH_2PO_4, 2.7 mM KCl, 137 mM NaCl, pH 7.4) + 1 mM $MgCl_2$ +0.1% Tween-20 and 1% BSA) (see Fig 16A). When the dye was added in buffer without K^+ (8.1 mM Na_2HPO_4, 137 mM NaCl, pH 7.4) + 1 mM $MgCl_2$ +0.1% Tween-20 and 1% BSA), RB2 binding was destroyed as expected, but RB3 and TFBS performed at least as well as when K^+ ions were present (Fig 18B).

GQ formation in DNA has been reported to be highly dependent on monovalent cations[34]. The stability of the structures is claimed to be high in the presence of K^+ because the metal ion has an ionic radius between 1.3-1.5 Å, an ideal size to fit within the two GQ's of the complex[37] (see Fig 18). The metal ion is believed to shift the equilibrium of a distribution of random and GQ folded TFBS to the GQ structure, aiding in thrombin binding[38]. Some reports state that the TFBS GQ is not detected without K^+ in the system[39], while others believe K^+ is more beneficial near the melting temperature of TFBS[38]. Baldrich and coworkers found that increasing the KCl concentration from 10-100 mM increased binding to thrombin, while binding was reduced to the level of PBS buffer alone with the KCl concentration was increased to 5 µM. In contrast to the Baldrich work, Tang and coworkers[40] observed a decrease in binding from 10-50 mM KCl for TFBS and thrombin, but they were using a HEPES-based buffer with 10 mM $MgCl_2$. Na^+ and Mg^{2+} have been shown to form weaker complexes with the GQ and lower the melting temperature due to the larger ionic radii. Hianik and coworkers[41] saw that increasing the Na^+ concentration reduced thrombin binding in a Tris binding buffer, and another group reported that Na^+ reduced the stability of GQ's irrespective of thrombin[42]. Baldrich likewise reports that adding Na^+ or Mg^+ to a solution with KCl will lower the binding[38]. Baldrich concludes overall that at low temperature, K^+ is not critical for thrombin binding. This body of research explains why TFBS does not appear to be affected by the riboflavin buffer, or the buffer lacking potassium.

It may also give insight into why RB3 shows lower binding in the RB selection buffer than in the normal chip conditions. Recall that two different groups reported a decrease in binding at high KCl concentrations. While the K^+ may stabilize the GQ structure, it may actually cause more difficulty to bind a target due to either a competition effect, or by creating such a rigid structure that it cannot adapt to interact strongly with the target. Song and coworkers observed that while K^+ might stabilize the structure of THBS[43], it actually reduces the ability of THBS to form a complex with thrombin. So, in essence, RB3 may have been overloaded with KCl and "locked" into a rigid conformation that could not interact with the Cy3 dye as much. Results for both riboflavin/RB3 aptamer[34], and thrombin/TFBS[38] show that the target itself can stabilize the respective aptamer, and Baldrich also claims that the presence of thrombin itself can promote GQ formation in the TFBS. When the buffer without K^+ was involved for the Cy3 studies, the more flexible structure slightly stabilized by Na^+ can recognize the dye and change conformation to accommodate the target.

Tasset and coworkers contend that the THBS aptamer consists of a GQ that requires extra spatial mobility to form due to the longer duplex between the 5' and 3' terminus[44]. They also report that the substitution of T for A at position 4 (Fig 18) destroys the T4:T13 base pair present in the TFBS, reducing the stability of the GQ compared to TFBS. This explains why we see binding of Cy3 dye to THBS (Fig 11G) that increases substantially with linker length, and that is lower affinity than binding to TFBS. (Note the drastic difference across linker lengths for Cy3-thrombin binding THBS in Fig 17D.) Tasset also reports that THBS is not dependent on K^+, and we see improved binding in the buffer lacking K^+ (Fig 18C). The RB buffer produces the next

highest level of fluorescence, and also has the highest K^+ concentration. Potentially, while the THBS doesn't depend on K+, the increased $MgCl_2$ concentration is known to stabilize non-GQ aptamers and may aid in stabilizing the stem portion which may indirectly stabilize the GQ, or the sequence may prefer HEPES buffer to PBS[45]. T*he GQ structure of THBS has been proposed but no evidence has been presented to validate the structure before the microarray work.* Microarrays may provide an avenue for studying structures of proposed aptamers, and characterizing their behavior under changing conditions.
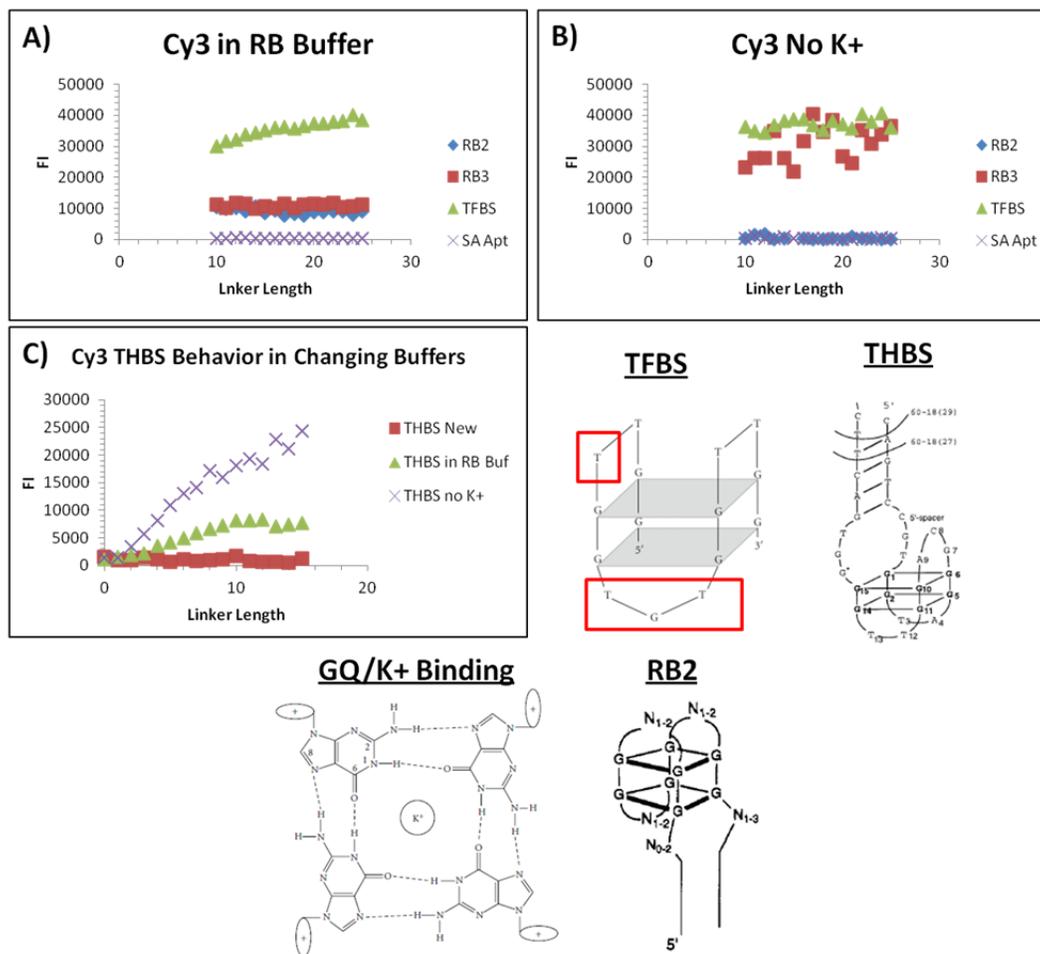
Figure 18: Cy3 Interactions in Changing Buffer Conditions[34,42,44]

Finally, the arrays were analyzed to determine if new binders to any of the Cy3 proteins were apparent. In addition to the control sequences, either 7,000 (standard controls) or 5,000 (extended controls) different sequences from PT1 were present in duplicate. As described previously, we calculated the mean and the range to evaluate potential binders. It is not ideal to use the 8X15k arrays to identify new binding sequences due to the relatively low library representation, few replicates, and generally lower diversity of different sequences compared to the 1X1M arrays. However, strong binders demonstrating fluorescence intensity values high above background values with reasonably small ranges may still be identified.

Of all the proteins tested, only thrombin bound to sequences with intensities significantly above background. The top ranked sequence, 4A018 (Fig 19A) demonstrated a mean fluorescence

intensity value >100X greater than background (SA-apt), and the range is <15% of the mean. The top 5 ranked sequences (disregarding the T_{10} linker) were aligned and analyzed for sequence homology using MAFFT (Fig 19B1). The majority of the perfect sequence homology appears in the external stem region of the DNA, but internal homology is observed when two sequences are directly compared. All binding sequences have 5 or 6 G runs. We also added the top 5 thrombin binders into the WebLogo program to generate the consensus sequence shown in Figure 19B2. The height of the bases corresponds with how strong the homology is at a certain position. Note that the 5 sequences contain matches even in the N region of PT1, designated with a green box. This consensus sequence was aligned with the TFBS in MAFFT (Fig 19B3), showing that the consensus sequence was a perfect match for 12 of 15 bases in TFBS. Cy3-only binding to the top ranked DNA was not considered since we know that thrombin binds GQ DNA, and we just showed dyes can bind GQ's as well.

Next, we wanted to determine if the PT1 library was biased in favor of the binding sequences by a high representation of potential thrombin binders present in the 5k library. This was accomplished by counting the on-chip abundance of some of the different fragments possible within the context of the pattern compared to the TFBS sequence (which we consider the first several bases to be "required" for binding, for simplicity). PT1 allows for:

PT1- (RRYYRRYY)-N_4-(RRYY)-N_3-(RRYY)-N_4-(RRYY)-N_3-(RRYY)-N_4-(RRYYRRYY); R=(A,G) - purines, Y=(T,C) - pyrimidines, N=(A,C,G,T) – random.

Some relevant potential fragments of the first several bases within this context and the abundance (of 5k total):

GGTTGGTT - 84   GGCCGGCC- 89

GGTTGGT - 194   GGCCGGC- 159

GGTTGG-   406   GGCCGG-   386

GGT**G**TGG-   0 (Because 4 and 8 positions are locked as T or C, not G)

AATTAATT - 16

AATTAAT- 59

AACCAACC- 43

AACCAAC- 78

TFBS: GGTTGGT**G**TGGTTGG.

The 6 base GGTTGG present in the TFBS accounts for 406 of the PT1 sequences, or 8.1% of the 5,000 total oligonucleotides. Expanding this to the first 7 bases of TFBS (GGTTGGT) means that 3.9% (194) of the library present meet the criteria for binding. When compared to the sequences with C's substituted for T's (non-binding), GGCCGG and GGCCGGC constitute 7.7% (386) and 3.2% (159), respectively. Oligonucleotides beginning with A were less abundant than G counterparts of similar length, but they would still be grouped into the nonbinding category. So the potential binders (considered to begin with GGTTGG as in the TFBS) do not dominate the library, and the fact that the top 5 ranked sequences all contain this motif (with the exception of a C in position 3 for 4A202) is a strong indicator of to thrombin binding. Sequences with the GGTTGGT**G** or GGT**G**TGG motif will not be present because the PT1 pattern does not allow for G at positions 4 or 8. SA-Apt is provided as a range of ranks based on the highest and lowest fluorescence intensity values for comparison (Fig 19A).

Figure 19: Identification of Potential Thrombin Binding Aptamers

The next five sequential sequences following the top five binders (ex. 4A018 is the top ranked binder, 4A019 is analyzed as a "random" sequence) were also evaluated (Fig 19C). These sequences demonstrated low fluorescence intensity values (Fig 19C1) as well as minimal structural homology with each other (Fig 19C2) since there is only one perfect match. In the order shown in Figure 19C1, the oligonucleotides have 3, 1, 6, 4, and 4 G pairs, respectively. Thus, higher amounts of G pairs will only promote binding if the surrounding sequence is correct and has potential to fold into a functional GQ structure. Sequence 1A176 is ranked 369 in mean fluorescence intensity, but the range is higher than the mean because the mean is calculated from one low point (FI= 103, close to the mean of the background SA-Apt) and one relatively high point (FI= 336). Thus, the higher point is likely more a function of experimental error from having only two replicates rather than a real binding event. The consensus sequence evaluation of the random sequences (Fig 19C3) shows fewer base matches (smaller height of bases compared to Figure 19B2) in addition to a smaller degree of alignment in the random regions (green boxes). Alignment with the TFBS showed several bases at the 3'-end matching (Fig 19C4), although the two G runs in this homologous region is not sufficient to form the GQ structure reportedly necessary for thrombin binding. ***The conclusions are that top binders display a concrete sequence homology with each other and with the TFBS aptamer, and a random sampling of non-binders has minimal homology.***

Testing of binding was performed via SPR with 4A018 and thrombin utilizing the schematic depicted in Figure 20A. Figure 20B shows a summary of the amounts of the components immobilized in each flow channel (FC), and Figures 20C-E are representative sensorgrams for neutravidin immobilization on FC4 (Fig 20C), and 4A018 immobilization on FC4 (Fig 20D) and FC3 (Fig 20E). Figure 21 shows the reference (FC1) subtracted results from adding a range of thrombin concentrations to the chip. Figures 21A-B are the reference subtracted sensorgrams for FC3-1 and FC4-1, displaying a trend of increasing response as the thrombin concentration is increased. FC2-1 did not show binding to thrombin (data not shown). The triplicate data from the sensorgrams was used to construct the affinity curves in Figures 21C-D. FC3-1 reported a $K_d = 1.82 \pm 0.16$ µM (Fig 21C), while FC4-1 $K_d = 2.88 \pm 0.10$ µM. The sensorgrams were also used to create kinetics curves (Fig 21E-F) that were fit to obtain k_a and k_d, and these rate constants were used to independently calculate K_d (Fig 21E inset).
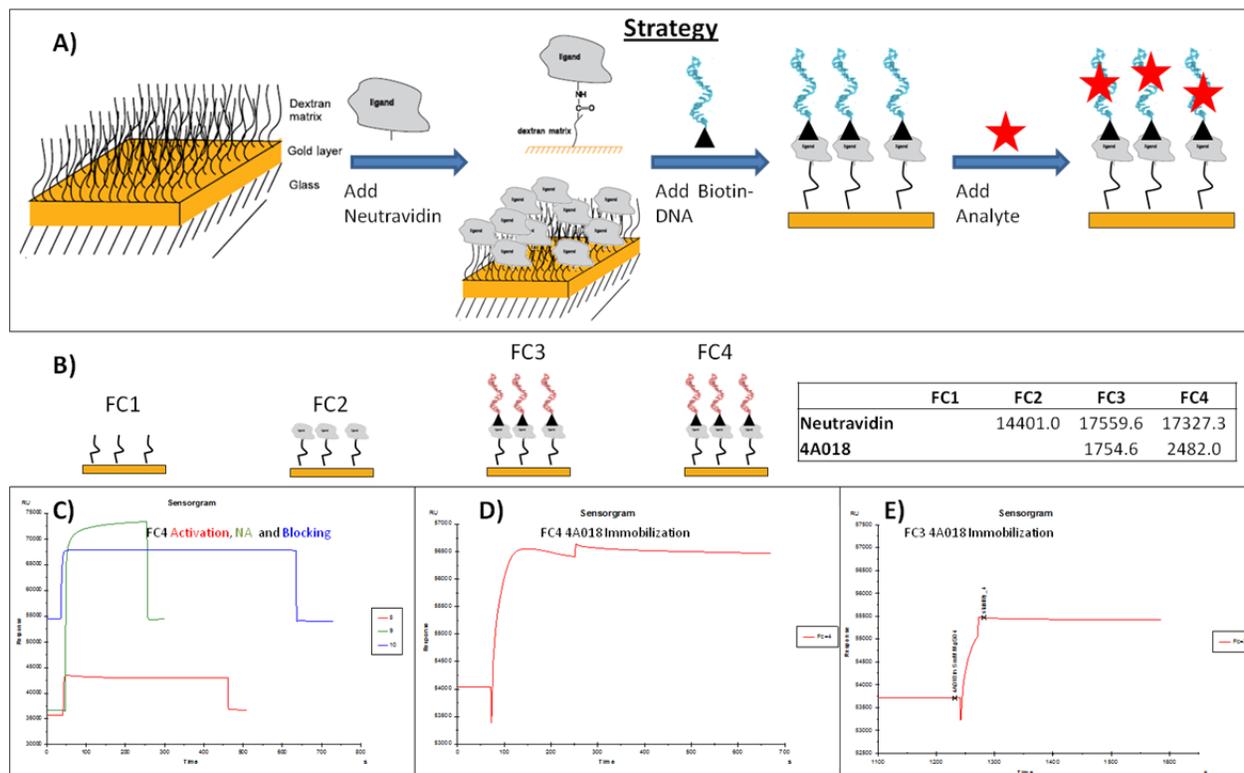

Figure 20: SPR Strategy and Immobilization of 4A018

Finally, 4A018 response was also tested in triplicate versus three other proteins as a control experiment. Proteins also with abundance in the body and similar in molecular weight to thrombin, bovine serum albumin (BSA) and human serum albumin (HSA), as well as a protein with a pI close to that of thrombin, neuropeptide Y (NPY) were chosen as controls. Figures 22A-C show the sensorgrams demonstrating minimal response for BSA, HSA, and NPY on FC4-1. The summary of protein response for FC3-1 and FC4-1 are depicted in Figure 22D-E. This shows the specificity of the 4A018 interaction for thrombin over other physiological proteins of similar size or pI. Ongoing work is being carried out to determine the structural properties of 4A018 as well as to characterize the affinity of some of the other array-discovered probes for thrombin.
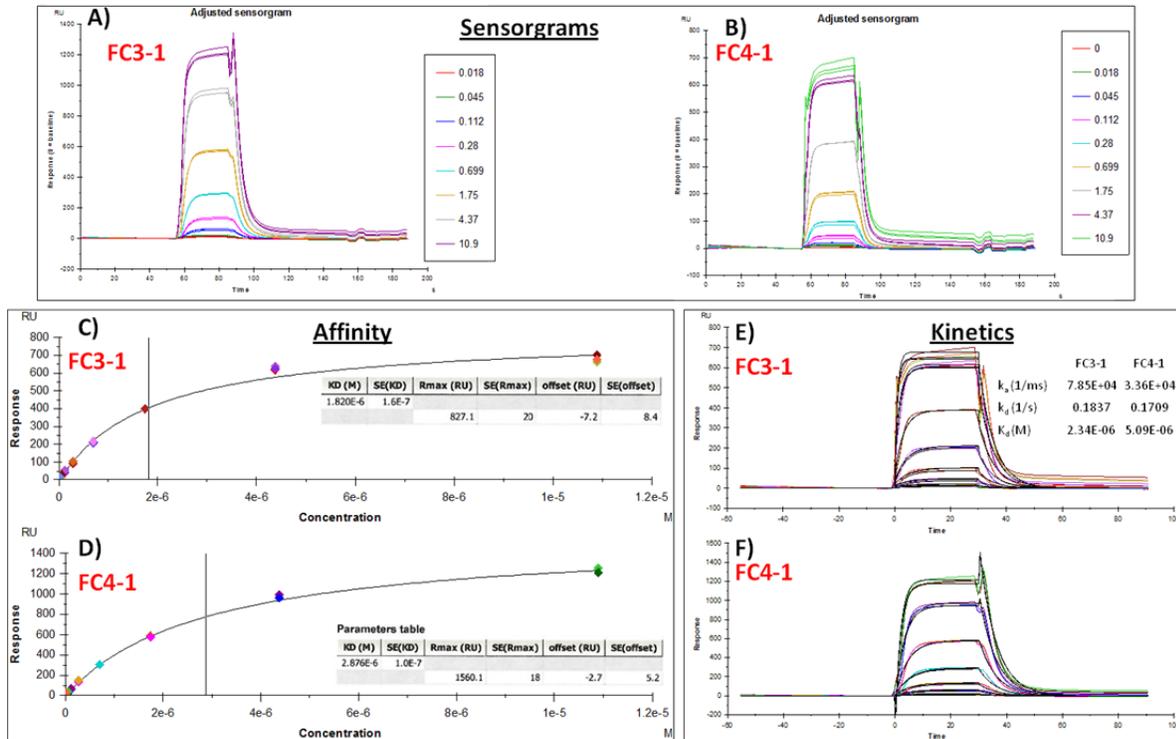
**Figure 21: Thrombin Interaction with 4A018**
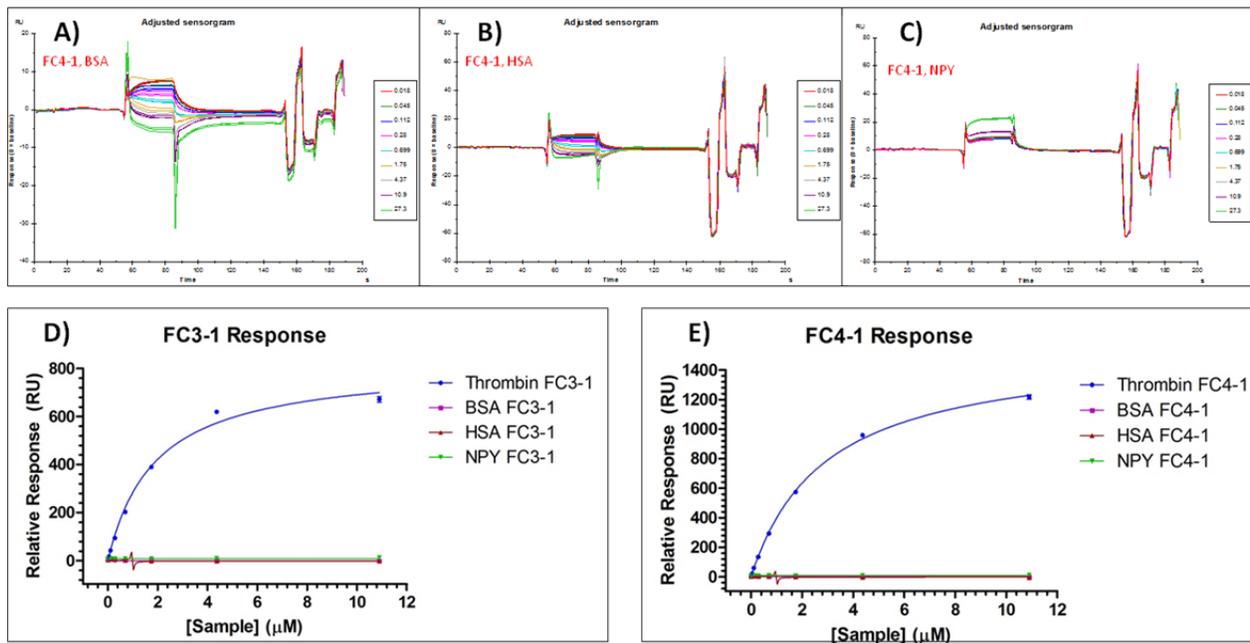
Figure 22: 4A018 SPR Response to Control Proteins and Overall Binding Summary

5.0 CONCLUSIONS

A Summary of significant findings includes:

1. Linker length increases will increase fluorescence intensity by extending the probes from the surface, and some minimal length is needed to obtain reliable data.

2. The linker identity (A, C, G, or T) affects fluorescence intensity, with T base linkers proving reliable across all linker lengths and target concentrations.

3. A direct (Cy3-protein) and indirect (biotin-protein followed by Cy3-streptavidin) target labeling method were compared, and it was shown that the indirect method both enhances S/N by orders of magnitude and decreases nonspecific binding to controls.

4. Trends of probe behavior were recognized, with the *in vitro* library designed to have more G-rich stretches appearing to display enhanced binding to the biotin and Cy3 tags.

5. Known G-quartet aptamers to non-protein targets bound the labeled protein with variable intensities depending on the target composition, purity, concentration, and stability of the DNA under experimental conditions. The small molecule tags were validated as the source of binding because neither the protein target alone nor the Cy3-SA interacted with the controls.

6. Potential aptamers to biotin and Cy3 were unearthed by using a mean fluorescence intensity ranking system to cross-reference the performance of probes with their presentation in various experiments including comparing Cy3-protein, Cy3 dye alone, and biotin-protein binding.

7. The planar and/or aromatic reporter tags interacted extensively with the DNA.

8. The surface density of the arrays is too low to expect binding from a random library, so starting library optimization must be performed. A thrombin binder was identified and characterized from the patterned library.

9. Microarray experiments generate massive amounts of data that can be problematic to organize and analyze.

10. Purification of free dye from dye-protein conjugates is essential, and the manufacturer's method is insufficient for aptamer identification experiments.

The DNA microarray work in this paper was designed to support three different initiatives: 1) To show that microarrays can be utilized to identify protein aptamer candidates; 2) Non-canonical G-quartet DNA can be detected by Cy3 (or biotin) binding on the microarray; 3) Aptamers to the small molecule biotin and Cy3 tags can be discovered by protein conjugation. Towards the first goal, we observed that increasing linker length will increase fluorescence intensity by extending the probes from the surface, and some minimal length is needed to obtain reliable data. The linker identity also affects fluorescence intensity, with T base linkers proving reliable across all linker lengths and target concentrations. We compared a direct and indirect target labeling method, and found that the indirect method both enhances S/N by orders of magnitude and decreases nonspecific binding to controls. Trends of probe behavior were recognized, with the *in vitro* library designed to have more G-rich stretches appearing to

display enhanced binding to the biotin and Cy3 tags. Several probes were identified with potential for IgE binding with structures different from those previously reported in the literature. Most likely, strong candidates were not observed because the patterns used in the libraries does not allow for "good" IgE binders. IgE aptamers may require a fairly rigid base sequence that was not obtainable. If strong binders were present, they would likely preferentially bind IgE over the free dye, and present with enhanced signals similar to thrombin v. Cy3 binding.

We also studied the behavior of the TFBS, THBS, and RB aptamers to examine G-quartet properties on the arrays. G-quartets bound to targets with variable intensities depending on the target composition, concentration, and stability of the DNA under experimental conditions. The small molecule tags (dye or biotin) were validated as the source of binding because neither the IgE alone nor the Cy3-SA interacted with the controls. We were able to use the tags to study known GQ aptamers under changing conditions, and elucidated interesting structural information on the probes. Literature reports on THBS structure and behavior was validated using microarrays. Therefore, DNA microarrays can be used to assay new and existing aptamers to screen for G-quartet formation by simply adding dye. Interesting studies of aptamer stability under various buffer, temperature, etc. conditions could also be performed in a multiplexed format using microarrays. Potentially, microarrays could be used to quickly screen probes for function that will be immobilized on surfaces such as nanoparticles.

Potential aptamers to biotin and Cy3 were unearthed by using a ranking system to cross-reference the performance of probes across the various experiments. At the least, these aptamers will enrich the field of small molecule aptamers in a high-throughput platform. The aptamers could be utilized as positive controls in future Cy3/biotin microarray work. It shows the feasibility of conjugating small molecule targets to a larger reporting system in which binding occurs between the small molecule rather than the reporter. This could be used to develop small molecule aptamer discovery protocols, or a dye displacement strategy can be employed from an in vivo-constructed library of DNA molecules with similar structures. Eventually we plan to use microarrays as a screening tool for patterned libraries. We will screen for libraries with a low selection rate and high structural diversity *in silico*, then evaluate them on the arrays for small molecule binding. "Winning" patterns will be applied to variations of SELEX to increase the selection efficiency and decrease required labor.

Finally, we were able to identify and confirm thrombin aptamers using only 5k sequences from a partially structured library. The library was not specifically designed to contain thrombin-binders, but motifs similar to TFBS demonstrated the highest signal. Applying thrombin to a 1X1M array would likely provide even higher affinity aptamers. Thrombin was a "better" target candidate than IgE due to the potential to bind a range of similar structures that form GQs.

There are several major drawbacks associated with using DNA microarrays for aptamer selection. Most notably, the planar and/or aromatic tags interacted extensively with the DNA. It is extremely difficult to determine whether lower-than-positive control level binding is real target binding or nonspecific tag interaction. The purity of targets from free dye after conjugation should be a priority, and the manufacturer's recommendation of a desalting column is insufficient when new binders are being pursued. The indirect labeling method was able to lower suspected nonspecific binding, but an ideal strategy would not require a label at all. Secondly, the surface density of the arrays when compared to SELEX is too low to expect

binding from a random library. If only 1 in 1×10^9 sequences are binders in a random SELEX library[6], extraordinary fortune would be required for 1 sequence to show any binding on an array. Design of the starting library has shown promise, but the methods from the literature largely take into account binding behavior from previously existing aptamers for the target. The most interesting targets will be those which do not currently have an aptamer available. Also, gene expression is the marketed application for the microarrays, and evolving a protocol to suit aptamer studies can require significant effort. For example, the Agilent arrays were designed for a small air space to populate each array so rotation of the air space will enhance the mass transfer of the target to the surface. However, this action requires an additive that changes the surface tension of the binding buffer, which would also affect the binding of the DNA. Instead, we overfilled the arrays to take up the space otherwise occupied by the air bubble, which causes a reliance purely on target diffusion. The array manufacturing time takes an average of 4-5 weeks to possess the chips in-hand. This tends to impede progress when relying on the results of initial experiments to change or improve the next generation of chips. Another challenge many may not consider is how to organize and analyze the massive amounts of data generated by one 1X1M experiment. A simple Excel spreadsheet is insufficient to work with the data, especially when different amounts of replicates are evaluated and parameters such as mean fluorescence intensity and standard deviation of the replicates must be calculated. Therefore, code to carry out these functions had to be written and tested before we could begin to assess the results. Basic data analysis accounted for 5-10X additional time consumption compared to one physical experiment alone. More complex functions such as cross referencing between experiments and analysis of multiple trends also required separate programming, although software such as GeneSpring does exist that can perform a number of the necessary processes. Finally, current technology is only capable of manufacturing DNA microarrays, but the company MYcroarray is willing to collaborate with us to develop RNA arrays. Initial work with DNA microarrays would serve as proof-of-concept studies to proceed with RNA microarray development compatible with riboswitch biosensor studies.

Overall, we have shown that DNA microarrays have great potential in rapid aptamer identification due to elimination of the need for SELEX cycling, PCR, cloning or sequencing, reducing the theoretical timeline to one day of experimental work. A novel thrombin aptamer has been identified, and the binding affinity of the probe to thrombin has been quantified. Additionally, an aptamer initially selected on a surface provides knowledge that the sequence will function immobilized on a biosensor platform setting. Microarrays are capable of highly multiplexed studies, limited only by the initial design of the array. For example, the IgE binders, biotin/Cy3 binding sequences, and G-quartet results were all gleaned from different probes and trends from the same physical experiments. Once some of the challenges of the technology are addressed, microarrays are a viable option for accelerated aptamer identification and structural study provided the target and initial library are a good match.

6.0 REFERENCES

1) Ellington, A. and Szostak, J., "*In Vitro* Selection of RNA Molecules that Bind Specific Ligands," *Nature,* **346**, 1990 pp. 818-822.

2) Tuerk, C. and Gold, L., "Systematic Evolution of Ligands by Exponential Enrichment: RNA Ligands to Bacteriophage T4 DNA Polymerase," *Science*, **249**, 1990 pp. 505–510.

3) Shangguan, D., Li, Y., Tang, Z., Cao, Z. C., Chen, H. W., Mallikaratchy, P., Sefah, K., Yang, C. J., Tan, W., "Aptamers Evolved from Live Cells as Effective Molecular Probes for Cancer Study," *Proc Natl Acad Sci.*, **103**, 2006 pp. 11838-11843.

4) Wilson, C. and Szostak, J. W., "Isolation of a Fluorophore-specific DNA Aptamer with Weak Redox Activity," *Chem. Biol.*, **5**, 1998 pp. 609-617.

5) Shangguan, D., Tang, Z. W., Mallikaratchy, P., Xiao, Z. Y., Tan, W., "Optimization and Modifications of Aptamers Selected from Live Cancer Cell Lines," *Chembiochem*, **8**, 2007 pp. 603-606.

6) Gopinath, S. C. B., "Methods Developed for SELEX," *Anal Bioanal Chem*, **387**, 2007 pp. 171-182.

7) Jayasena, J. D., "Aptamers: An Emerging Class of Molecules that Rival Antibodies in Diagnostics," *Clin. Chem.*, **45**, 1999 pp. 1628-1650.

8) Djordjevic, M., "SELEX experiments: New prospects, applications and data analysis in inferring regulatory pathways," *Biomolecular Engineering*, **24**, 2007 pp. 179-189.

9) Savory, N., Abe, K., Sode, K., Ikebukuro, K., "Selection of DNA Sptamer Against Prostate Specific Antigen Using a Genetic Algorithm and Application to Sensing," *Biosensors and Bioelectronics*, **26**, 2010 pp. 1386-1391.

10) Cho, M., Xiao, Y., Nie, J., Stewart, R., Csordas, A. T., Oh, S. S., Thomson, J. A., Soh, H. T., "Quantitative Selection of DNA Aptamers Through Microfluidic Selection and High-throughput Sequencing," *Proc Natl Acad Sci.*, **107**, 2010 pp. 15373-15378.

11) Walter, J-G. Kokpinar, O., Friehs, K., Stahl, F., Scheper, T., "Systematic Investigation of Optimal Aptamer Immobilization for Protein-Microarray Applications," *Anal. Chem.*, **80**, 2008 pp. 7372-7378.

12) Srivannavit, O., Gulari, M., Hua, Z., Gao, X., Zhou, X., Hong, A., Zhou, T., Gulari, E., "Microfluidic Reactor Array Device for Massively Parallel in Situ Synthesis of Oligonucleotides," *Sensors and Actuators B: Chemical*, 2009, **140**, 473-481.

13) Lao, Y-H., Peck, K., Chem, L-C., "Enhancement of Aptamer Microarray Sensitivity through Spacer Optimization and Avidity Effect," *Anal. Chem.*, **81**, 2009 pp. 1747-1754.

14) Cho, E. J., Collett, J. R., Szafranska, A. E., Ellington, A. D., "Optimization of aptamer microarray technology for multiple protein targets," *Analytical Chimica Acta*, **564**, 2006 pp. 82-90.

15) Collett, J. R., Cho, E. J., Lee, J. F., Levy, M., Hood, A. J., Wan, C., "Functional RNA Microarrays for High-throughput Screening of Antiprotein Aptamers," Ellington, A. D. *Anal. Biochem.*, **338**, 2005 pp. 113-123. (a)

16) Collet, J. R., Cho, E. J., Ellington, A. D., "Production and Processing of Aptamer Microarrays," *Methods*, **37**, 2005 pp. 4-15. (b)

17) Katilius, E., Flores, C., Woodbury, N. W., "Exploring the Sequence Space of a DNA Aptamer Using Microarrays," *Nucleic Acids Research*, **35**, 2007 pp. 7626-7635.

18) Biesecker, G., Dihel, L., Enney, K., Bendele, R. A., "Derivation of RNA aptamer inhibitors of human complement C5," *Immunopharmacology*, **42**, 1999 pp. 219–230.

19) Fischer, N. O., Tok, J. B.-H., Tarasow, T. M., "Massively Parallel Interrogation of Aptamer Sequence, Structure and Function," *Plos One*, **3**, 2008 pp. e2720.

20) Platt, M., Rowe, W., Wedge, D. C., Kell, D. B., Knowles, J., Day, P. J., "Aptamer Evolution for Array-based Diagnostics," *Analytical Biochemistry*, **390**, 2009 pp. 203-205. (a)

21) Knight, C. G, Platt, M., Rowe, W., Wedge, D. C., Khan, F., Day, P. J. R., McShea, A., Knowles, J., Kell, D. B., "Array-based Evolution of DNA Aptamers Allows Modeling of an Explicit Sequence-fitness Landscape," *Nucleic Acids Res.*, **37**, 2009 pp. e6.

22) Asai, R., Nishimura, S. I., Aita, T., Takahashi, K., "In Vitro Selection of DNA Aptamers on Chips Using a Method for Generating Point Mutations," *Analytical Letters*, **37**, 2004 pp. 645-656.

23) Bock, C., Coleman, M., Collins, B., Davis, J., Foulds, G., Gold, L., Greef, C. et al., "Photoaptamer Arrays Applied to Multiplexed Proteomic Analysis," *Proteomics*, **4**, 2004 pp. 609-618.

24) Platt, M., Rowe, W., Knowles, J., Day, P. J., Kell, D. B., "Analysis of Aptamer Sequence Activity Relationships," *Integrative Biology*, **1**, 2009 pp. 116-122. (b)

25) Osborne, S. E., Ellington, A. D., "Nucleic Acid Selection and the Challenge of Combinatorial Chemistry," *Chem. Rev.*, **97**, 1997 pp. 349-370.

26) Hermann, T., Patel, D J., "Adaptive Recognition by Nucleic Acid Aptamers," *Science*, **287**, 2000 pp. 820-825.

27) Carothers, J. M., Oestreich, S. C., Davis, J. H. and Szostak, J. W. "Informational Complexity and Functional Activity of RNA Structures," *J Am Chem Soc*, **126**, 2004 pp. 5130-5135.

28) Chushak, Y. and Stone, M. O. "*In Silico* Selection of RNA Aptamers," *Nucleic Acids Research*, **37**, 2009 pp. e87.

29) Davis, J. and Szostak, J., "Isolation of High-affinity GTP Aptamers from Partially Structured RNA Libraries," *Proc Natl Acad Sci.*, **99**, 2002 pp. 11616-11621.

30) Luo, X., McKeague, M., Pitre, S., Dumontier, M., Green, J., Golshani, A., Derosa, M. C., Dehne, F., "Computational approaches toward the design of pools for the in vitro selection of complex aptamers," *RNA*, **16**, 2010 pp. 2252-2262.

31) 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,* **132**, 2010 pp. 9453– 9464.

32) Hahn, C. D., Riener, C. K., Gruber, H. J., "Labeling of Antibodies with Cy3-, Cy3.5-, Cy5-, and Cy5.5-monofunctional Dyes at Defined Dye/Protein Ratios," *Histochemie*, **43**, 2000 pp. 1-5.

33) Li, Y., Geyer, C. R., Sen, D., "Recognition of Anionic Porphyrins by DNA Aptamers," *Biochemistry*, **35**, 1996 pp. 6911-6922.

34) Lauhon, C. T., Szostak, J. W., "RNA Aptamers that Bind Flavin and Nicotinamide Redox Cofactors," *J. Am. Chem. Soc.*, **117**, 1995 pp. 1246-1257.

35) Stadtherr, K., Wolf, H., Lindner, P., "An Aptamer-Based Protein Biochip," *Anal. Chem.*, **77**, 2005 pp. 3437-3443.

36) German, I., Buchanan, D. D., Kennedy, R. T., "Aptamers as Ligands in Affinity Probe Capillary Electrophoresis," *Anal. Chem.*, **70**, 1998 pp. 4540-4545.

37) Kankia, B. I., Marky, L. A., "Folding of the Thrombin Aptamer into a G-Quadruplex with Sr^{2+}: Stability, Heat, and Hydration," *J. Am Chem. Soc.*, **123**, 2001 pp. 10799-10804.

38) Baldrich, E., Restrepo, A., O'Sullivan, C. K., "Aptasensor Development: Elucidation of Critical Parameters for Optimal Aptamer Performance," *Anal. Chem.*, **76**, 2004 pp. 7053-7063.

39) Arthanari, H., Basu, S., Kawano, T. L., Bolton, P. H., "Fluorescent dyes specific for quadruplex DNA," *Nucleic Acids Research*, **26**, 1998 pp. 3724-3728.

40) Tang, Q., Su, X., Loh, K. P., "Surface plasmon resonance spectroscopy study of interfacial binding of thrombin to antithrombin DNA aptamers," **315**, 2007 pp. 99-106.

41) Hianik, T., Ostatna, V. Sonlajtnerova, M., German, I., "Influence of ionic strength, pH and aptamer configuration for binding affinity to thrombin," *Bioelectrochemistry*, **70**, 2007 pp. 127-133.

42) Diculescu, V. C., Chiorcea-Paquim, A.-M., Eritja, R., Oliveira-Brett, A. M., "Thrombin-Binding Aptamer Quadruplex Formation: AFM and Voltammetric Characterization," *J. Nucl. Acids*, 2010 pp. 1-8.

43) Song, M., Zhang, Y., Li, T., Want, Z., Yin, J., Wang, H., "Highly sensitive detection of human thrombin in serum by affinity capillary electrophoresis/laser-induced fluorescence polarization," *J. Chromatogr. A*., **1216**, 2009 pp. 873-878.

44) Tasset, D. M., Kubik, M. F., Steiner, W., "Oligonucleotide Inhibitors of Human Thrombin that Bind Distinct Epitopes," *J. Mol. Biol.*, **272**, 1997 pp 688-698.

45) Taylor, J. N., Darugar, Q., Kourentzi, K., Willson, R. C., Landes, C. F., "Dynamics of an anti-VEGF DNA aptamer: A single-molecule study," *Biochem. and Biophys. Res. Comm.*, **373**, 2008 pp. 213-218.

APPENDIX A - DNA Microarrays for Aptamer Identification and Structural Characterization
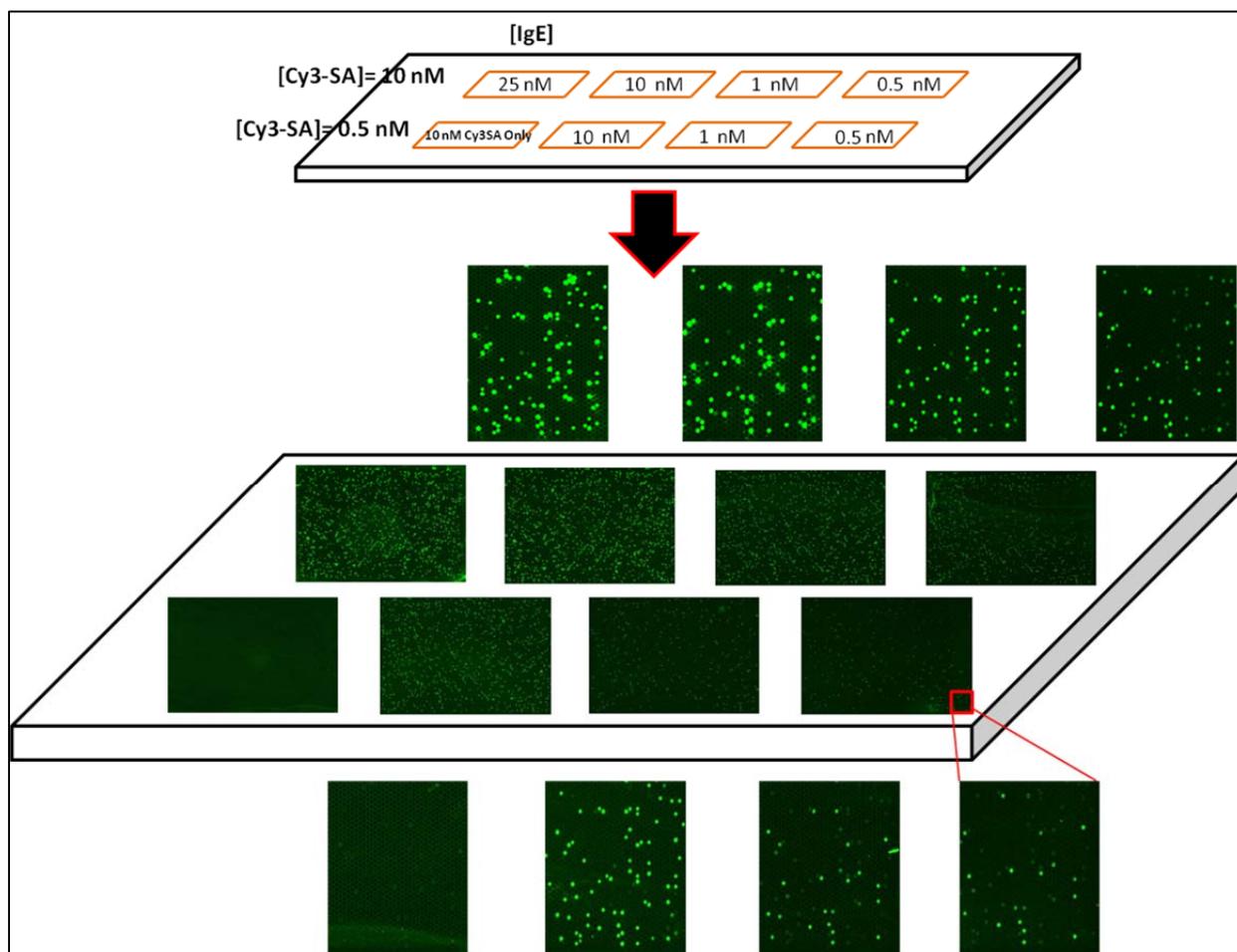

Figure A1: Images from Experiment in Figure 6B

Cutouts in Figure A1 are all blowups of the lower right corner for each array. Note that similar locations fluoresce for the different conditions, corresponding to the positive controls.

Table A1: Control Values from 10 nM Cy3-SA Only

Linker Length	ID17-4DM13_25	ID17-4SM17	ID17-4SM21	ID17-4SM28	IgE_D17-1	IgE_D17-4_DM1_30	IgE_D17-4_SM1	IgE_D17-4A	IgE_D17-4C	IgE_D17-4G	IgE_D17-4T
0	5.44	6.13	7.62	7.79	11.82	9.24	6.03				7.16
1	7.14	6.48	6.65	5.60	10.30	9.09	9.82				9.65
2	6.09	6.10	5.78	5.84	10.23	9.42	11.95	9.64	11.45	10.37	11.79
3	4.86	5.60	4.34	5.54	10.38	12.11	12.49				12.15
4	5.43	7.71	4.76	4.42	9.31	14.63	14.50	16.00	16.52	13.72	15.74
5	4.85	4.71	3.58	3.89	9.68	16.02	18.30				19.58
6	6.36	5.46	4.98	4.83	13.24	20.42	23.90	19.94	21.12	18.84	17.96
7	4.22	4.22	3.47	12.57	11.66	21.50	25.24				22.08
8	4.21	4.97	5.65	3.32	15.18	28.64	35.80	19.88	26.76	23.48	25.90
9	3.87	2.95	3.81	4.67	9.97	28.82	32.70				32.48
10	3.55	3.20	4.73	4.02	11.92	37.24	37.24	28.40	32.90	33.22	28.82
11	3.50	3.63	3.07	8.93	9.36	44.38	46.34				41.26
12	4.60	3.90	2.95	2.95	12.98	48.64	41.64	31.30	33.44	37.02	47.24
13	3.61	7.13	2.99	2.95	11.80	38.92	49.16				38.78
14	3.39	4.48	3.11	2.95	10.48	49.86	71.46	32.68	39.12	43.80	52.42
15	3.07	3.22	2.95	2.95	11.12	58.42	66.80				55.10

Table A2: Control Values from 10 nM Cy3-SA + 10 nM Biotin-IgE

Linker Length	ID17-4DM13_25	ID17-4SM17	ID17-4SM21	ID17-4SM28	IgE_D17-1	IgE_D17-4_DM1_30	IgE_D17-4_SM1	IgE_D17-4A	IgE_D17-4C	IgE_D17-4G	IgE_D17-4T
0	6.83	8.51	8.36	11.16	19.10	854.8	1033.2				1096
1	5.17	8.93	7.38	5.62	25.16	2162	2486				2470
2	5.60	7.53	8.78	5.23	48.04	3366	3854	4504	4524	3926	4084
3	4.99	8.80	6.36	5.54	126.4	5164	6080				6126
4	5.58	10.38	6.92	6.24	380	7480	8418	7558	7518	6602	7764
5	3.32	11.40	5.24	5.99	724	8506	9334				8712
6	4.75	8.45	6.68	5.19	1135	9564	11160	10146	11140	9490	10564
7	4.26	10.79	5.60	4.99	1566	10680	13180				11920
8	3.53	10.46	3.79	4.05	2262	13460	14160	12780	13020	13380	13820
9	5.54	11.73	5.75	4.90	2856	14280	16040				14140
10	4.50	12.80	5.20	3.78	3286	15880	17980	14860	14900	16600	16000
11	3.61	12.14	5.83	3.94	4138	18760	20680				18300
12	5.46	13.24	4.78	3.18	5028	19500	20260	16020	16640	19240	19500
13	3.44	18.66	4.14	4.13	5634	21500	23260				21000
14	4.56	16.34	3.25	4.90	6534	22460	23260	18540	18640	20140	21940
15	3.92	17.36	4.69	3.18	6778	22680	25900				23880
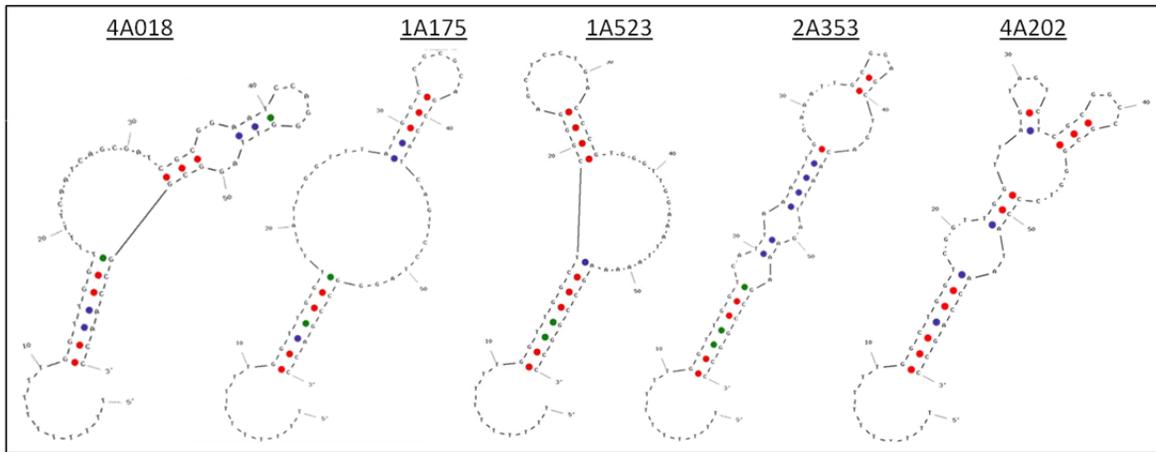
Figure A2: Structures of Potential Thrombin Aptamers

LIST OF ACRONYMS/GLOSSARY

A, C, G, T, R, Y	adenine, cytosine, guanine, thymine, R= (A/G), Y= (T/C)
ATP	adenosine triphosphate
BSA	bovine serum albumin
Cy3	Cyanine dye3
Cy5	Cyanine dye5
Cy3-BSA	Cyanine dye3 conjugated to bovine serum albumin
Cy3-Estradiol	Cyanine dye3 conjugated to estradiol antibody
Cy3-IgE	Cyanine dye3 conjugated to Immunoglobulin E
Cy3-NPYsc	Cyanine dye3 conjugated to scrambled neuropeptide Y sequence
Cy3-SA	Cyanine dye3 conjugated to streptavidin
Cy3-thrombin	Cyanine dye3 conjugated to thrombin
DNA	deoxyribonucleic acid
D/P	dye-to-protein ratio
FC	flow channel
FRET	fluorescence resonance energy transfer
G-quartet (GQ)	guanine quartet
HABA	4'-hydroxyazobenzene-2-carboxylic acid
HPLC	high performance liquid chromatography
HSA	human serum albumin
IgE	Immunoglobulin E
K_d	dissociation constant
LOD	limit of detection
nt	nucleotide
NMR	nuclear magnetic resonance
NPY	neuropeptide Y
PCR	polymerase chain reaction
PT1,2,3	DNA patterned library #1, 2, or 3
RB2	2-tiered g-quartet riboflavin aptamer
RB3	3-tiered g-quartet riboflavin aptamer
RNA	ribonucleic acid
S/N	signal-to-noise ratio
SA	mutated non-binding streptavidin aptamer (background)
SELEX	Systematic Evolution of Ligands by Exponential Enrichment

SPR   surface plasmon resonance
ssDNA   single-stranded DNA
UV/Vis   ultraviolet/visible light absorption spectroscopy
TFBS   thrombin fibrinogen binding site aptamer
THBS   thrombin heparin binding site aptamer