A Multipronged Approach in Targeting *Clostridium difficile*: Multiple Domain Selection for Aptamer Isolation

A Multipronged Approach in Targeting *Clostridium difficile*: Multiple Domain Selection for Aptamer Isolation

By

AMJAD ARRABI

B.Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree Master of Science

McMaster University © Copyright by Amjad Arrabi, September 2017

MASTER OF SCIENCE (2017) McMaster University

(Biochemistry & Biomedical Sciences) Hamilton, Ontario

TITLE: A Multipronged Approach in Targeting *Clostridium difficile*: Multiple Domain Selection for Aptamer Isolation

AUTHOR: Amjad Arrabi, B.Sc. (McMaster University)

SUPERVISOR: Professor Yingfu Li

NUMBER OF PAGES: x, 92

ABSTRACT

Clostridium difficile, the causative agent of *C. difficile* infection (CDI), causes hundreds of thousands of hospital-acquired infections in the United States and Canada annually. Furthermore, the prevalence and severity of CDI has been on the rise in developed countries, especially with the appearance of "hypervirulent" strains. Detection of CDI is thus of great importance. Traditional detection methods can be time consuming or lack the desired sensitivity. On the other hand, aptamers pose great prospects as diagnostic and therapeutic agents. Aptamers are nucleic acid ligands with molecular recognition capabilities rivaling those of antibodies. They are obtained by a process of *in vitro* selection known as systematic evolution of ligands by exponential enrichment (SELEX). However, the current approach may result in aptamers that experience non-specific binding in complex or biological samples. Here, we propose a multiple domain selection (MDS) method for aptamer isolation. This method utilizes independent selections on separate components of a single target in order to obtain uniquely specific aptamers. The aptamers can then be unified into a heterobivalent construct able to recognize two sites on one target. We hypothesize the combined aptamer would result in greater affinity and specificity for the target, resulting in greatly increased aptamer utility in current and future applications. In the current study, we have cloned and purified full length *C. difficile* DnaK as well as the N-terminal domain (NTD) and C-terminal domain (CTD) of the protein. MDS was performed on each target and the resulting aptamers were combined into a heterobivalent construct. The construct resulted in an approximately

100-fold affinity increase relative to the single aptamer for DnaK, and could detect much smaller quantities of target. Although it experienced low level recognition of high concentrations of purified *E. coli* DnaK, there was no detectable non-specific binding in several biological samples.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Dr. Yingfu Li for the opportunity to have worked under his supervision. He has always been a source of inspiration and encouragement, and truly enjoys thoughtful, intellectual discussions. His passion in the pursuit of knowledge and functional nucleic acids research has fueled my own interests in this field and beyond.

My graduate experience would not be complete without the valuable insights of my committee members, Dr. John Brennan and Dr. Alba Guarné. I would like to thank them both for their expertise, which were an asset to my work. Furthermore, I would like to thank my colleagues in the Li lab who constantly created a positive environment of collaboration and thoughtful discussion.

Last but not least, I would like to thank my wife and family for their support over the years, especially my unconditionally loving mother.

Table of Contents

List of Figures

List of Abbreviations and Symbols

ATP	Adenosine 5'-Triphosphate
BLAST	Basic Local Alignment Search Tool
BSA	Bovine Serum Albumin
CDI	*C. difficile* Infection
CTD	C-Terminal Domain
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide Triphosphate
EIA	Enzyme Immunoassay
EMSA	Electrophoretic Mobility Shift Assay
F	Forward
FNA	Functional Nucleic Acid
MDS	Multiple Domain Selection
MOBIX	McMaster Institute for Molecular Biology and Biotechnology
NAAT	Nucleic Acid Amplification Test
NCBI	National Centre for Biotechnology Information
NTD	N-Terminal Domain
NGS	Next-Generation Sequencing
OD	Optical Density
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction
PMSF	Phenylmethylsulfonyl fluoride
R	Reverse
RNA	Ribonucleic Acid
ssDNA	Single-stranded DNA
SELEX	Systematic Evolution of Ligands by Exponential Enrichment

Chapter 1 – Introduction

1.1 Functional Nucleic Acids (FNAs)

For most of the history of modern biology and biochemistry, and even in current scientific education, nucleic acids are mostly seen as either carriers of genetic information (DNA) or transmitters of genetic information (RNA). This viewpoint made sense intuitively based on our initial understanding of genetics and the central dogma of molecular biology proposed by Francis Crick (Crick, 1970). The central dogma states that the flow of genetic information (within a biological system) begins with DNA and ends with proteins, with RNA being the intermediary. Furthermore, this flow of genetic information is unidirectional and cannot be reversed (from protein to DNA). In essence, this describes the transformation of a genotype into a phenotype. Since then, however, we have learned much more about nucleic acid structure and function. Although the central tenants of the dogma still hold true, the specifics certainly have changed. We now know that proteins can be modified post-translationally (Lodish, 1981; Uy & Wold, 1977); they can also exist in alternate, infectious, forms known as prions, with the ability to transmit their alternate structure onto native proteins (Bolton, McKinley, & Prusiner, 1982; Prusiner, 1982); they can even exist as inteins – self-catalytically splicing protein elements within other proteins (Kane et al., 1990; Perler, 2002). Furthermore, DNA itself can be modified in several ways so as to alter gene expression without altering gene sequence (Jaenisch & Bird, 2003; Rideout Ill et al., 2001). This is all to say that our

knowledge of molecular biology is continually changing and expanding. Nucleic acids are no different in this regard, and their role outside of genetics is ever expanding.

1.1.1 Nucleic Acid Enzymes (NAEs)

The first evidence that nucleic acids had a functional role in biological systems was the discovery of RNA molecules that could catalyze reactions (Guerrier-Takada et al., 1983; Kruger et al., 1982). Thomas Cech's group (Kruger et al.) discovered a 413 bp self-splicing RNA sequence within the 26S rRNA coding region of *Tetrahymena thermophila.* This intervening sequence could excise itself from the larger strand, covalently cyclize itself, and ligate the remaining strands together. Sidney Altman's group, on the other hand, showed that it was the RNA portion of *E. coli* and *B. subtilis* ribonuclease P which was responsible for the catalysis of tRNA cleavage. These catalytic RNA molecules were termed ribozymes and formed the foundation of the RNA world hypothesis whereby current translation mechanisms involving proteins could have evolved from RNA molecules capable of performing both transcriptional (informational) and translational (catalytic) roles in cellular organisms.

Since then, many other natural ribozymes have been discovered and are classified mainly by size. The larger ribozymes are made up of the self-splicing (group I and II) introns as well as ribonuclease P, which were mentioned earlier. The other class of ribozymes is made up of a smaller group with a distinct catalytic mechanism. These smaller ribozymes perform a transesterification reaction that is self-cleaving, producing a

cyclic phosphate terminus and hydroxyl group. This class of ribozymes includes the hammerhead (Pley et al., 1994), Varkud Satellite (VS) (Lafontaine et al., 2001), glmS (Klein & Ferré-D'Amaré, 2006), hairpin (Rupert & Ferré-D'Amaré, 2001), and hepatitis delta virus (HDV) (Ferré-D'Amaré et al., 1998) ribozymes. Interestingly, this subset of ribozymes can function in the absence of divalent metal ions, unlike their larger counterparts, possibly through utilization of their nucleobases (Bevilacqua & Yajima, 2006). Natural ribozymes are not only limited to the above reactions as evidenced by the peptidyl transferase ability of the ribosome, which can be considered another example of a ribozyme (Ban et al., 2000; Nissen et al., 2000). Although the ribosome exists as a vast conglomerate of RNA and protein, it is only the RNA component present at the active site which catalyzes peptide bond formation.

Natural ribozymes are not the only source of nucleic acid enzymes; there has been tremendous work done in identifying synthetic NAEs, both from RNA (ribozymes) and DNA (deoxyribozymes, DNAzymes). In fact, synthetic NAEs comprise a very versatile group of catalytic molecules which can perform many more reactions than naturally identified NAEs. The "catalyst" to the identification of these molecules was the independent work done by three groups to identify functional nucleic acids through an *in vitro* selection process (Ellington & Szostak, 1990; Robertson & Joyce, 1990; Tuerk & Gold, 1990). This method which can be described as "evolution in a test tube" has resulted in the identification of synthetic ribozymes capable of an extensive number of reactions including, but not limited to: RNA cleavage (Williams et al., 1995), ligation

(Ekland et al., 1995), phosphorylation (Lorsch & Szostak, 1994), peptide bond formation (Zhang & Cech, 1997), acyl transfer (Suga et al., 1998), urea synthesis (Nieuwlandt et al., 2003), amide synthesis (Wiegand et al., 1997), pyrimidine synthesis (Chapple et al., 2003), purine synthesis (Lau et al., 2004) and many more.

Although, initially, RNA has been favoured due to its implicit role in the RNA world hypothesis and its inherent advantage (reactivity) due to the 2'-hydroxyl group, many synthetic deoxyribozymes have also been discovered. DNAzymes have been found to catalyze RNA cleavage (Breaker & Joyce, 1995), RNA ligation (Purtha et al., 2005), DNA adenylation (Li et al., 2000), DNA ligation (Sreedhara et al., 2004), and DNA phosphorylation (Wang et al., 2002), among other reactions. The repertoire of DNAzymes may currently be limited in comparison to their RNA counterparts, but they certainly are not catalytically inferior. The DNAzyme identified as 10-23 has a k_{cat} of 10 min^{-1} and a catalytic efficiency of ~10^9 M^{-1} min^{-1}, rivaling similar RNA and protein enzymes (Santoro & Joyce, 1997). Because investigations into deoxyribozymes began later than ribozymes, it is likely that even more diverse species will be found in the years to come.

1.1.2 Aptamers (Nucleic Acid Ligands)

Nucleic acid enzymes are not the only kind of functional nucleic acids. The same *in vitro* selection method which lead to the rise of NAEs, also lead to the rise of nucleic acid ligands, termed aptamers. The term itself was coined by Szostak and Ellington from

the Latin *aptus* (to fit) because they are viewed as three dimensional structures that "fit" their target (Ellington & Szostak, 1990). Aptamers are able to fold into specific shapes, and through molecular interactions they can dock inside a large target or wholly encompass a small one (Figure 1.1). These interactions include hydrogen bonding, electrostatic interactions, hydrophobic contacts, and stacking of π systems (Hermann & Patel, 2000). It is also believed that aptamers are mostly unstructured until contact with their target, whereby they undergo conformational changes for specific target recognition; in some instances the target itself may undergo conformational changes when bound by an aptamer (Nguyen et al., 2002). It is because of these properties that aptamers have been generated for targets ranging from small molecules (Sassanfar & Szostak, 1993) , to proteins (Wang et al., 1993), viruses (Dey et al., 2005), and even whole cells (Shangguan et al., 2006).

As with NAEs, there are two main types of aptamers: DNA (ssDNA) and RNA. The main difference between the two moieties is the presence of a 2'-hydroxyl on the ribose sugar of RNA. This simple but critical difference differentiates the two in an important way – chemical stability. The presence of the 2'-hydroxyl makes RNA more susceptible to hydrolysis and enzymatic digestion. Furthermore, ribonucleases are much more prevalent in biological samples, making RNA more vulnerable. Thus, it can be a challenge for assays using RNA aptamers, which can degrade quickly under biological conditions.

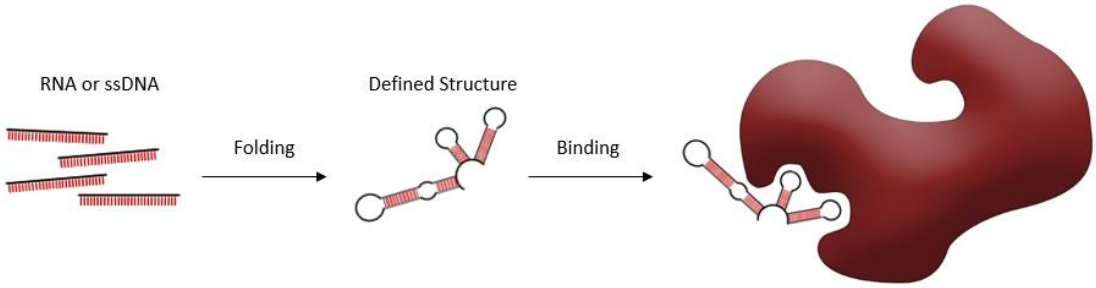

Figure 1.1. Schematic representation of aptamer folding and molecular recognition. RNA or single-stranded DNA aptamers are isolated through *in vitro* selection and fold into three dimensional structures that define their molecular recognition ability.

This can be overcome by using RNA libraries with modified bases or sugars, such as replacing the 2'-hydroxyl with 2'-fluoro or 2'-amino groups, but this requires using mutant T7 RNA polymerase during selection (Keefe & Cload, 2008). Easy chemical modifications, however, give RNA an advantage by increasing the total diversity and structure of potential aptamers, which can lead to better target binding (Vaish et al., 2003). DNA can also be chemically modified to increase stability (Shoji et al., 2007), and both RNA and DNA aptamers can be made using L-(deoxy)ribose, making them extremely resistant to cellular nucleases (Vater & Klussmann, 2015). These aptamers, called Spiegelmers (from the German "Spiegel," mirror), work by using D-(deoxy)ribose selections against the mirror image of the intended target, which allows the L-configuration to bind the intended enantiomer. Unlike RNA, DNA is inherently more stable. It is also easier to work with during selection and does not need to be reverse transcribed. Generally, DNA is easier and cheaper to produce than RNA, which if

modified, is especially costly. Each type of aptamer has its own strengths and drawbacks, but both are equally useful for target selection and can even be interchanged (Walsh & DeRosa, 2009).

Just as NAEs can be thought of as the nucleic acid analogues of protein enzymes, aptamers can be thought of as the nucleic acid analogues of antibodies. Unlike NAEs however, the performance of aptamers relative to their protein counterparts can be considered equivalent, in terms of affinity and specificity (Jayasena, 1999). Aptamers, however, have several advantages over antibodies: they are smaller and can be readily and reliably manufactured in large quantity; they can be generated for almost any target, and unlike antibodies, they can be obtained for non-immunogenic or toxic targets due to the nature of the *in vitro* selection process; they can also be denatured and renatured without loss of function, and have a longer shelf life. Furthermore, aptamers typically do not induce an immune response when administered to a host and many readily permeate cells (Cheng et al., 2013; Thiel et al., 2012).

It has become increasingly clear that aptamers possess unique properties and are useful for many potential applications. Due to their tight binding and molecular recognition properties, aptamers are both useful for diagnostic applications (Zhou et al., 2014) and are also valuable therapeutic agents (Keefe et al., 2010). Indeed, they are not only able to specifically bind targets, but they also have the potential to disrupt or enhance a target's function by binding an active site or delivering a modifying agent.

They are also useful research tools when combined with fluorescent dyes, allowing molecular investigations within living cells to be viewed and imaged (Paige et al., 2011).

1.2 Selection of Aptamers (SELEX)

The *in vitro* selection of aptamers, or SELEX, has been modified a great many times since its first inception, for specific purposes. However, the underlying principles have remained the same (Figure 1.2). Generally, one begins with a random library of ~10^{15} unique sequences, composed of an inner random domain flanked by primer binding sites. The library is exposed to a target of interest for a period of time under select conditions. Next, the unbound sequences are washed away and those that remain bound are eluted. An amplification step (PCR) is then carried out in order to amplify the remaining sequences and produce an enriched library. The process is repeated several times before the final pool is tested for binding and/or analyzed to determine candidate aptamers.

The first SELEX protocols utilized nitrocellulose filter binding to obtain aptamers (Ellington & Szostak, 1990; Tuerk & Gold, 1990). This method works by taking advantage of nitrocellulose's ability to bind proteins nonspecifically. The selection library is incubated with the target protein and passed through a nitrocellulose filter; those sequences that bound the target protein are captured while the others pass through.

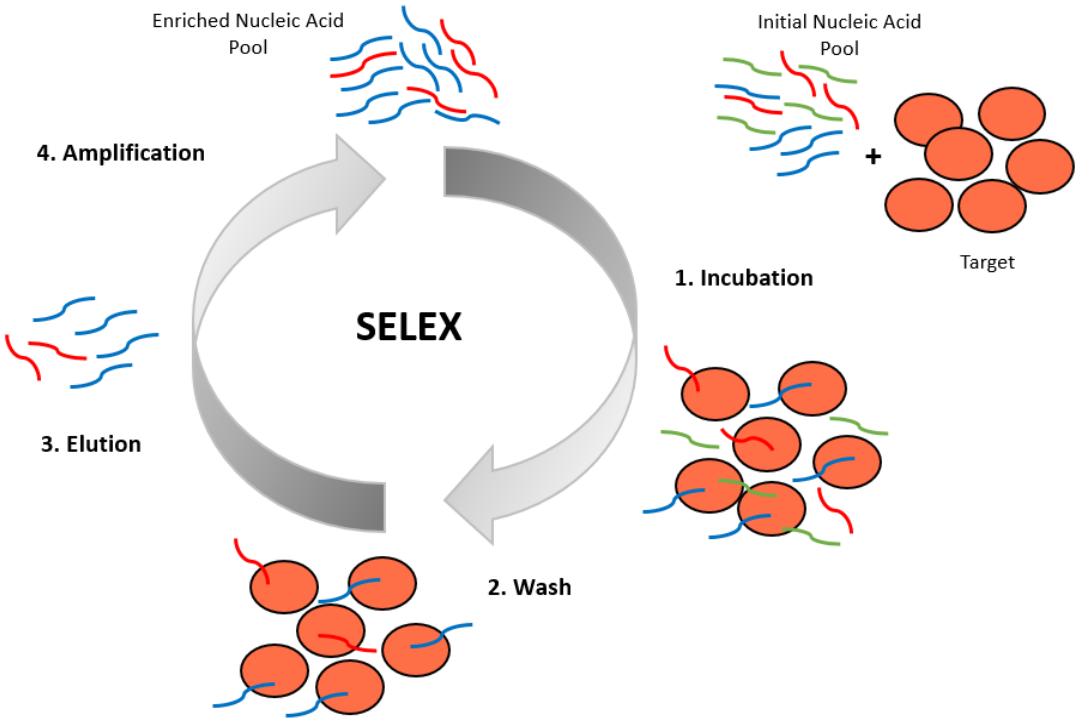

Figure 1.2. General SELEX scheme for obtaining target-specific aptamers. The selection begins with a large randomized nucleic acid library, on the order of 10^{15} molecules. This initial pool is then incubated with the target of interest (1). After incubation, the target is washed to remove any non-specific sequences (2). Those oligonucleotides that have remained bound are then eluted (3) and then PCR amplified (4). The enriched pool is then used for another round of selection.

The library is thus enriched and can be amplified and reapplied for another round of selection. The greatest advantage of this method is that it allows true equilibrium between potential aptamers and the target in solution. It is also relatively simple but does have its limitations. Although proteins bind nonspecifically to the nitrocellulose, not all proteins bind efficiently, and small target molecules pass through altogether. Even more, some aptamers have been shown to bind the nitrocellulose and propagate throughout the

selection (Shi et al., 2002). Alternatively, this method can be used as a relatively inexpensive follow up assay to quantitatively probe aptamer-protein binding (Hall & Kranz, 1999). Due to the limitations of nitrocellulose SELEX, other protocols were also developed.

Bead-based SELEX methods whereby targets are immobilized on a solid surface are also widely used (Oh et al., 2009; Stoltenburg et al., 2005; Tok & Fischer, 2008). The most common beads used are agarose and magnetic beads that take advantage of particular binding chemistries (e.g. Ni-NTA affinity to the 6xHis-tag and glutathione affinity to the GST-tag). The versatility of this method makes it an attractive option for many targets, especially tagged recombinant proteins. Generally, the target is incubated with the beads first to allow binding. After a few washes, only the targets that have bound the beads remain, which can then be exposed to the selection library. Alternatively, the target can be exposed to the library first and then captured by the functionalized beads. The latter approach may be more advantageous due to the free-floating state of the target, giving potential aptamer sequences full access to the target surface.

In addition to the aforementioned bead-based SELEX, more novel approaches have been considered. A single round selection scheme, MonoLEX, utilizes a sectioned column to retrieve high affinity aptamers in one round (Nitsche et al., 2007). In theory, it works due to the fast-on and slow-off rates of aptamers that bind near the top of a column when passing through the gel matrix (FIG. 3). Although this was not consistent, Nitsche

et al. were able to generate high affinity aptamers to whole Vaccinia virus in only one round of selection. Another novel approach used a single bead selection to obtain high affinity aptamers against botulinum neurotoxin within two rounds (Tok & Fischer, 2008). These are only but a few examples, as there are numerous variations of SELEX (Ozer et al., 2014).

Deciding when to end a selection is a critical part of the SELEX process, and has also seen drastic changes. The electrophoretic mobility shift assay (EMSA) is a common and suitable method to determine target-aptamer binding (Hellman & Fried, 2007). This method relies on the fact that aptamers that have bound a target will increase in apparent molecular weight and complexity, and thus realize a shift in their band position when run on a gel. The greater the number of sequences within a library that bind the target, the more of the library will be shifted upward on a gel. Once a significant amount of shifting is observed within a library, selection can be stopped and the sequences characterized. Other techniques take advantage of FACS (Shangguan et al., 2006) and capillary electrophoresis (Yunusov et al., 2009) to determine aptamer binding.

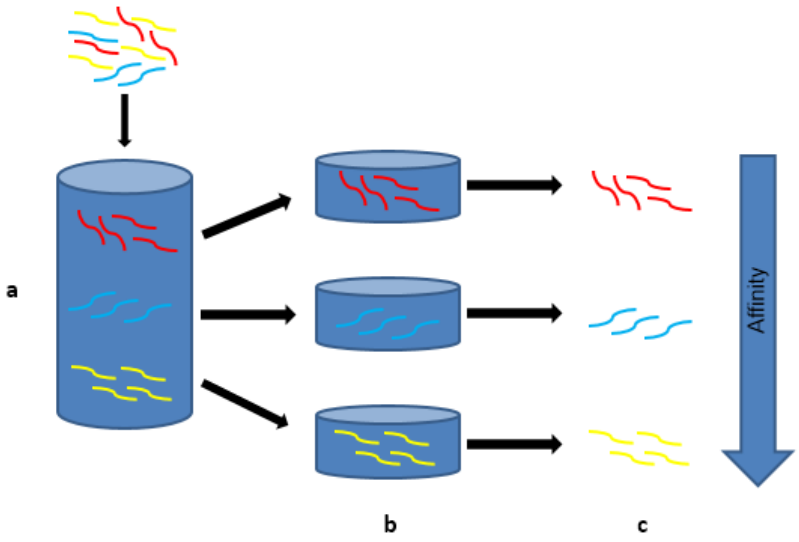

Figure 1.3. Schematic representation of MonoLEX. The initial library is applied to a sectioned column containing the target of interest (a). The columns are segmented with different affinity aptamers in separate sections (b). Aptamers are eluted from each section, amplified, and characterized (c).

Until recently, enriched aptamer libraries would have to be cloned into plasmids and then sequenced. This produced several hundred sequences to be tested and is rather laborious. However, the field has taken a giant leap forward with the advent and use of high-throughput sequencing (Cho et al., 2010). Next-generation sequencing (NGS), as the name suggests, has transformed the way aptamers are identified and characterized. NGS is able to produce many millions of reads in one go, which allows researchers to look at a larger proportion of their library and identify the most enriched sequences. It can even be used to analyze the enrichment processes itself over several rounds of selection (Zimmermann et al., 2010). This was used to show that the amplification steps during

SELEX can introduce biases (Schütze et al., 2011), highlighting the priority of obtaining high affinity sequences within a reasonable number of rounds.

1.3 Applications of Aptamers

Proteins play extremely important roles in cellular processes and are the primary targets of the majority of pharmaceuticals (Brauner-Osborne et al., 2007; Dorsam & Gutkind, 2007; Rosenbaum et al., 2009; Wise et al., 2002). This makes proteins some of the best targets for aptamers, which can be used in biotechnology (Liss et al., 2002; Murphy, 2003; Romig et al., 1999), diagnostic assays (Balamurugan et al., 2008; Levy et al., 2005; Stojanovic & Kolpashchikov, 2004), to combat infectious microorganisms (Bruno, Carrillo, & Phillips, 2008) and viruses (Held et al., 2006), and as therapeutics (Bunka et al., 2010).

Aptamer affinity towards a target can be used to target toxigenic bacteria or viruses. If an aptamer is able to bind an important protein and disrupt its function, then it can be used as an antibiotic or antiviral. Furthermore, aptamers can also be used to simply carry an antibiotic or antiviral to the target of interest (Figure 1.4). An important example of this includes work targeting HIV-1 replication by disrupting the crucial reverse transcriptase used by the virus (Held et al., 2006; Nickens, 2003; Tuerk et al., 1992; Whatley et al., 2013). This has produced aptamers able to target the protein with low nanomolar affinities that block the DNA binding region and stop DNA replication. Another important viral protein is HIV-1 protease which is necessary for cleaving viral

polyproteins into their functional form. Without it, the virus would be non-infectious. To that end, Zhu et al were able to conjugate a siRNA targeting HIV-1 protease to a DNA aptamer targeting CD4 protein on T Cells, where the virus propagates (Zhu et al., 2012). The resulting chimeric aptamer could enter $CD4^+$ T Cells and knock down the expression of HIV-1 protease. More recently, another group was able to produce high affinity aptamers to the HIV-1 protease itself, inhibiting its activity (Duclair et al., 2015). The aptamers were shown to be specific enough not to bind the related HIV-2 protease nor the cellular aspartyl protease, Cathepsin D.

Another method of targeting HIV-1 is to stop it before infecting a cell. To do so, aptamers would have to target the viral proteins responsible for internalization of viral particles, gp120 and gp41. To that end, several aptamers have been developed to target gp120 (Dey, Khati, et al., 2005; Dey, Griffiths, et al., 2005), and inhibit viral fusion by up to 10,000 fold. A crucial component of this binding revealed that the aptamers access a hidden site, which only becomes exposed when gp120 binds CD4, which is inaccessible to antibodies. Aptamers have also been employed to target other important viruses and pathogenic bacteria (Kavruk et al., 2015; Lee et al., 2013; Yang et al., 2013).

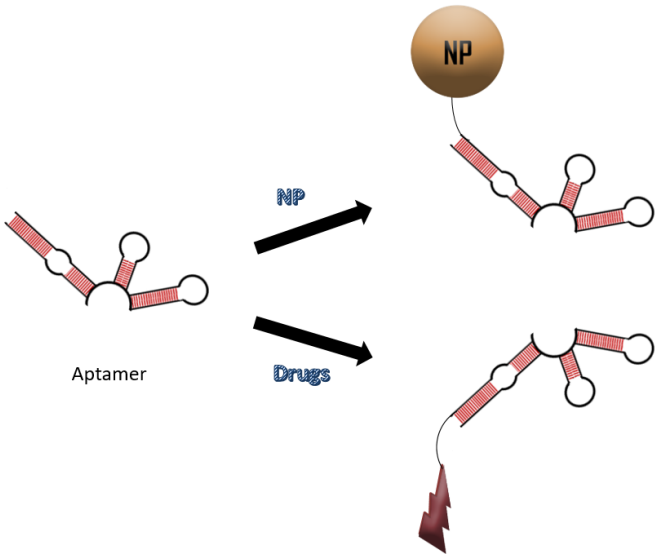

Figure 1.4. Schematic representation of aptamer-drug conjugates. Aptamers can be covalently modified to include drugs or nanoparticles that contain drugs for delivery to a target.

While many aptamers are being investigated for potential applications in the research lab, some have, or are in the process of moving into the clinic. The first aptamer to be approved by the FDA for therapeutic use was Pegaptanib, marketed as Macugen (Gragoudas et al., 2004; Ng et al., 2006) for the treatment of age-related macular degeneration (AMD). The aptamer targets VEGF, and once bound, inhibits VEGF's interaction with its receptors. This interaction is essential in the pathological angiogenesis characterized by AMD. Other important aptamers being considered for therapeutic use include AS1411 which was used by one group to target tumour cells for drug delivery (Aravind et al., 2012). AS1411 targets the nucleolin protein, which is taken up by tumours that overexpress the nucleolin receptor. Here, the group created drug-carrying

nanoparticles functionalized with AS1411 and showed they could target and kill cancerous cells. Another aptamer, RB006, that targets coagulation factor IXa, is being used as part of an anticoagulation system, REG1, as a reversible anticoagulant during medical surgery (Chan et al., 2008). The NOX-E36 aptamer has one of the best affinities for a selection target, 150 pM, and binds to the CCL2 protein (Maasch et al., 2008; Ninichuk et al., 2008). CCL2 recruits monocytes and lymphocytes into the extravascular space at sites of inflammation and NOX-E36 is being used as a possible treatment for type 2 diabetes and diabetic complications. NOX-A12, an aptamer that binds CXCL12, which plays an important role in angiogenesis and tumour metastasis, is being developed for use in hematopoietic stem cell transplants (Steurer et al., 2014). All of the mentioned therapeutic aptamers, with the exception of Pegaptanib which has already been approved, are in preclinical or clinical trials. Other aptamers are also being investigated for their therapeutic potential (Jaffe et al., 2016; Jilma et al., 2010).

1.4 *C. difficile Infection* (CDI)

Clostridium difficile is an anaerobic, Gram-positive, bacterium and the causative agent of *Clostridium difficile* infection (CDI). It was first isolated in 1935 (Hall & O'Toole, 1935), but was not associated with human disease until 1978, when it was linked to pseudomembranous colitis (Lance George et al., 1978; Larson et al., 1978). CDI is an intestinal disease mediated by toxin A and toxin B, which are produced from the bacterial pathogen during colonization of the colon. The clinical manifestations of the

disease vary but typically include diarrhea, abdominal pain, fever, and leukocytosis. More severe CDI can result in potentially fatal conditions such as bowel perforation, sepsis, and shock (Rupnik et al., 2009). Classically, those most vulnerable to the disease have been the hospitalized elderly (Kelly & LaMont, 1998), however, CDI in previously low risk populations, such as pregnant women and children, is on the rise (Khanna et al., 2013; Kim et al., 2008; Rouphael et al., 2008). Furthermore, the disease is also increasing within communities (Gupta & Khanna, 2014), and it has been suggested that upwards of 40% of cases are acquired within a community setting (Khanna et al., 2012; Lambert et al., 2009). CDI is now one of the leading nosocomial infections in the developed world with increased prevalence outside of healthcare facilities.

Data from the U.S. Centers for Disease Control and Prevention, among others, show that the CDI rate doubled from the late 90's to mid-2000's in the United States, and has continued to increase since (McDonald et al., 2006; Peery et al., 2012; Zilberberg et al., 2008). Similar or larger increases have also been reported in various developed nations (Burckhardt et al., 2008; Di Bella et al., 2013; Furuya-Kanamori et al., 2014; Lyytikäinen et al., 2009; Soler et al., 2008). In the U.S. there were fewer than 150,000 hospital discharge diagnoses of CDI in 2000, increasing to over 300,000 by 2006 (Rupnik et al., 2009). In 2011, there were an estimated half a million cases of CDI, resulting in more than 29,000 deaths (Lessa et al., 2015). The approximate monetary cost to the U.S. healthcare system is in excess of $3 billion annually (O'Brien et al., 2007). *C. difficile* infection is also a major problem in Canada (Eggertson & Sibbald, 2004; Loo et al.,

2005; Pépin et al., 2004; Valiquette et al., 2004), with the associated mortality rate increasing four-fold from 1997 to 2005 (Gravel et al., 2009). Furthermore, there is a hypervirulent strain of *C. difficile* that has been associated with outbreaks in recent years, and which is potentially epidemic (He et al., 2013; Lanis et al., 2013; McDonald et al., 2005; Pépin et al., 2005; Petrella et al., 2012). In light of the seriousness of CDI, and the evident increase in both frequency and severity, several detection methods have been developed to identify the infection.

1.5 *C. difficile* Detection

In order to be diagnosed with CDI, a suspected case, usually indicated by foul smelling diarrhea, which in itself is insufficient for a diagnosis, requires definitive laboratory findings (Bartlett & Gerding, 2008; Cohen et al., 2010). The traditional gold standard detection method is the cell culture cytotoxicity neutralization assay (CCCNA) (Burnham & Carroll, 2013; Planche & Wilcox, 2011). The assay is performed by preparing a stool filtrate that is then applied onto a cell monolayer; if cytopathic effects are observed, a neutralization assay is performed to determine whether the toxicity is mediated by *C. difficile* toxins or nonspecific toxicity. The performance of the CCCNA, however, depends on many factors: toxin degradation in the sample, the cell line used for culturing, pre-treatment of the patient with antibiotics, and the sample preparation method used (Burnham & Carroll, 2013; Stamper et al., 2009). This assay is no longer seen as the reference method as it is inadequately sensitive to be the gold standard; it is

also time consuming, taking a minimum of 24 to 48 hours to perform, and requires expertise in cell culture methods and interpretation of results.

The alternative reference method is the toxigenic culture, which is based on isolating *C. difficile* from patient samples and determining whether the strain recovered is toxin-producing (Burnham & Carroll, 2013; Delmée, Van Broeck et al., 2005; Planche & Wilcox, 2011; Stamper et al., 2009). Several protocols are used with variable success. Each employs anaerobic solid or liquid culture containing selective agents which promote the isolation of *C. difficile* but reduce the overgrowth of other fecal microbes (Hink et al., 2013). Once colonies arise, *C. difficile* is identified by using different methods including gram staining, colony morphology, odor, and biochemical testing (Burnham & Carroll, 2013). Isolates must then be tested for the ability to produce toxin, usually by culturing in broth and employing CCCNA or enzyme immunoassays (She et al., 2009). Although the toxigenic culture is very sensitive and is now seen as a reference standard, it has limited diagnostic utility. It is tedious in nature and takes several days to complete; it also detects *C. difficile* toxin production *in vitro* which may not be the case *in vivo* (Burnham & Carroll, 2013; Planche et al., 2013).

Due to the slow turnaround time of the reference methods, enzyme immunoassays (EIAs) and nucleic acid amplification tests (NAATs) were employed (Barbut et al., 2003; Gumerlock et al., 1991; Kato et al., 1991; Turgeon et al., 2003; van den Berg et al., 2005; Wren et al., 1990). EIAs utilize monoclonal or polyclonal antibodies directed against *C.*

difficile toxin(s) for detection, although those directed solely against toxin A can miss pathogenic strains that only encode toxin B (Elliott et al., 2011; Kim et al., 2008; Kim et al., 2012). Until recently, EIAs were the most frequently used diagnostic method to detect CDI due to their rapid turnaround times (hours) and familiarity of technique. However, it has since become apparent that these assays have considerable variations in sensitivity (Eastwood et al., 2009; Humphries et al., 2013; Luna et al., 2011; Peterson et al., 2011).

Given the drawbacks of EIAs, nucleic acid amplification methods were adopted. These methods utilize techniques such as PCR to amplify a variety of target genes including *tcdA* (toxin A), *tcdB* (toxin B), and 16S rRNA (Gumerlock et al., 1991; Kato et al., 1991). Initially, the procedures were time consuming and used gel electrophoresis and/or Southern blotting to detect PCR products but newer methods have improved sample extraction and utilize real-time PCR for detection (Pancholi et al., 2012; Peterson et al., 2007; Stellrecht et al., 2014). However, nucleic acid amplification techniques are not without drawbacks: they are relatively costly, require laboratory and/or molecular expertise, and cannot differentiate CDI from asymptomatic carriage because they detect the genes that encode toxins, not the toxins themselves (Carroll, 2011; de Jong et al., 2012; Johnson et al., 2013; Koo et al., 2014). Although current detection methods have helped in the fight against toxigenic *C. difficile*, it is clear that better methods are needed. In that respect, aptamers are a clear choice as potentially diagnostic molecular tools. However, aptamers are also susceptible to non-specific binding in complex, biological

samples, which may necessitate sample pre-treatment (Guthrie et al., 2006; Keefe et al., 2010; Liu et al., 2006; Zhao et al., 2009).

1.6 Multiple Domain Selection (MDS)

To get around potential non-specific binding in complex samples, and to enhance the affinity and general utility of aptamers, we propose an approach utilizing multiple domain selection (MDS) for aptamer isolation (Figure 1.5). This approach can be applied to any molecule that can potentially be divided into two or more parts, and still retain structural identity. In this regard, proteins containing multiple domains, are good targets. This strategy involves performing SELEX on individual domains of a protein in order to obtain specific, high affinity aptamers for multiple sites. We hypothesize that linking the resultant aptamers into a single heterobivalent aptamer will produce an aptamer with better recognition (specificity) and binding (affinity) capabilities than individual aptamers. This heightened specificity and affinity may allow aptamers to be used in assays without requiring pre-treatment of complex or biological samples, and should increase the overall accuracy and utility of the assays.

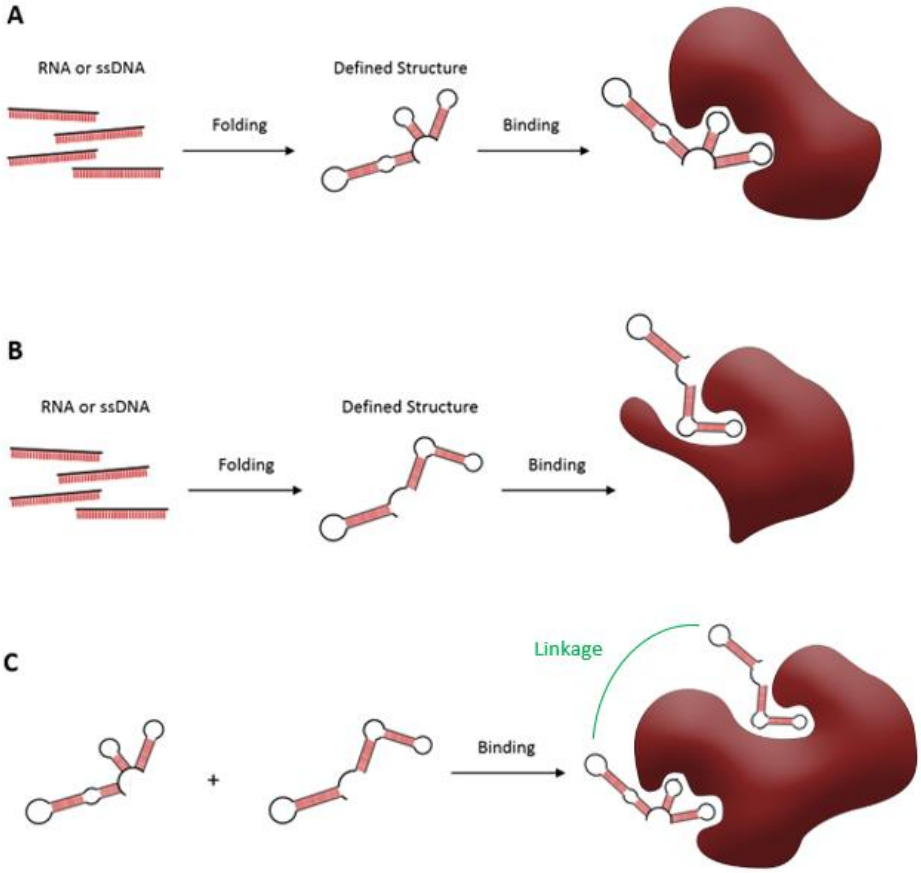

Figure 1.5. Schematic representation of multiple domain selection (MDS). (A) Selection against one domain of a protein yielding a specific aptamer and (B) selection against another, distinct, domain of the same protein yielding a separately specific aptamer. (C) Individual, high affinity, aptamers that recognize distinct parts of the same protein can be chemically linked.

In the current study, we aimed to investigate multiple domain selection (MDS) for isolating aptamers against a *C. difficile* protein which could be useful for biosensing and as a general proof of concept. In that regard, an endogenous protein that is similar in *C. difficile* as well as other bacteria, would be optimal. For this, we investigated *C. difficile* DnaK (Hsp70). DnaK is an ATP-dependent chaperone of approximately 70 kDa involved

in numerous protein folding processes and is a central component of the cellular surveillance network (Langer et al., 1992; Mayer, 2013). It also appears to have an export signal in *C. difficile*, and therefore found to some degree within the extracellular environment (Boetzkes et al., 2012). Furthermore, DnaK contains two functional domains that can be targeted for selection (Figure 1.6), and it is found to be upregulated during a clinically relevant heat stress (Mayer, 2013; Ternan et al., 2012).

The concept of combining or linking two or more aptamers has been proposed previously and has been shown to increase affinity (Hasegawa et al., 2008; Kim et al., 2014; Mallikaratchy et al., 2011; Nonaka et al., 2010; Poniková et al., 2011). However, our approach is unique in that we began with independent selections against individual domains of a single target. We then constructed a heterobivalent aptamer from single aptamers targeting different sites on that target, which had not been done previously. Furthermore, the target (DnaK) is ubiquitous and has close sequence and structural similarity to homologs from different species of bacteria. Thus, it is a good target to demonstrate the specificity and affinity of a heterobivalent aptamer.

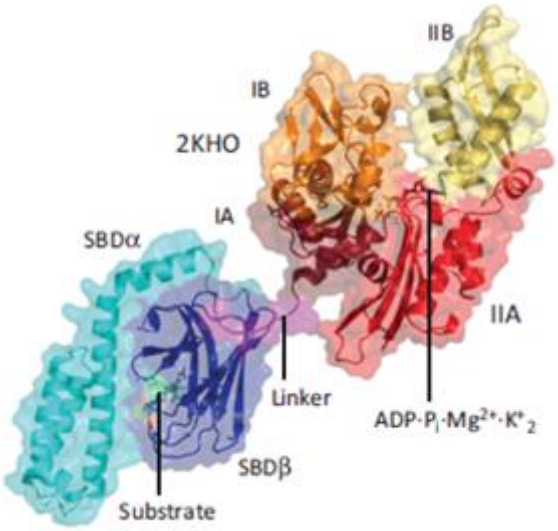

Figure 1.6. Crystal structure of *E. coli* DnaK. (Mayer, 2013).

1.7 Thesis Objectives

Since the structure of *C. difficile* DnaK had not yet been resolved, we predicted the location of the inter-domain linker as this region is highly conserved between species, and is known in *E. coli* (Bertelsen et al., 2009; Mayer, 2013). Alignment of the *C. difficile* NAP7 DnaK amino acid sequence with the *E. coli* K12 homolog revealed the likely location of the linker (Figure 1.7). We chose to divide the 18 amino acid linker in two, leaving nine attached to the N-terminal domain (NTD) and nine to the C-terminal domain (CTD).

Our objectives are to 1) successfully clone the full DnaK protein, as well as its constituent domains (NTD and CTD); 2) express and purify the proteins sufficiently for use in selection; 3) perform MDS and successfully obtain specific aptamers for each target; 4) construct and test the performance of the heterobivalent aptamer relative to the single aptamers. The results will allow us to test our hypothesis: whether we can successfully perform MDS to create a heterobivalent aptamer with better affinity and/or specificity than a single aptamer for the target of interest.

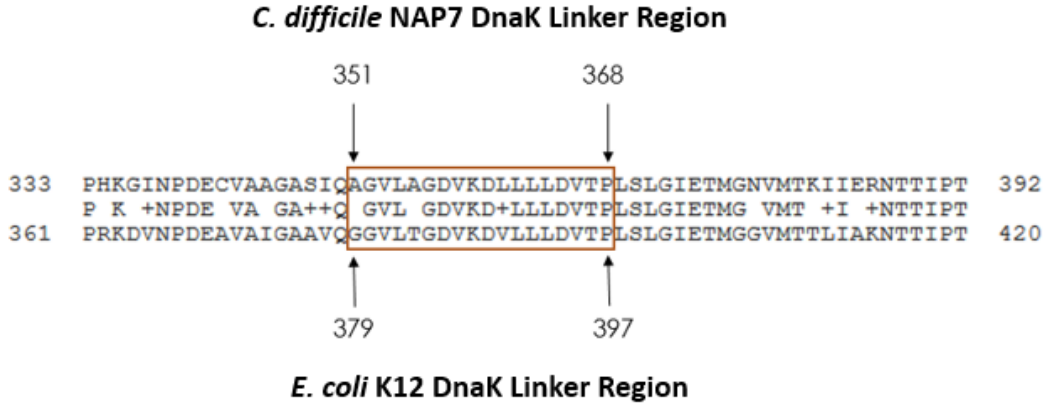

Figure 1.7. Alignment of *E. coli* K12 and *C. difficile* NAP7 DnaK linker sequences. The linker sequences in *E. coli* K12 comprises residues 379-397, which can be aligned to residues 351-368 in *C. difficile* NAP7. Alignments were obtained by utilizing the BLAST program at the NCBI website.

Chapter 2 – Materials and Methods

2.1 Cloning of *C. difficile dnaK*, NTD, and CTD

2.1.1 PCR Amplification

To clone *C. difficile* NAP7 full length *dnaK* as well as the NTD and CTD the following primers were used: F_{DnaK} 5' CCGGCTCGAGCATGGGAAAAATAATAGG 3', R_{DnaK} 5' CCCGGGGAATTCGTTTATCTTCATCTACTTC 3', F_{NTD} 5' CCGGCTCG AGCATGGGAAAAATAATAGG 3', R_{NTD} 5' CCGG GAATTCGTTTAACATCCCCT GC 3', F_{CTD} 5' CGGCTCGAGTATGGACTTA TTATTACTAG 3' and R_{CTD} 5' CCGGG GAATTCGTTTATCTTCATCTAC 3' (incorporating XhoI and EcoRI restriction sites in the F and R primers, respectively). Primers were obtained from MOBIX (McMaster University) and PAGE purified. Template DNA was obtained from lysed *C. difficile* ATCC® BAA1875™. *dnaK*, NTD, and CTD PCR reactions were performed in 50 µL reaction volumes containing: 5 µL template DNA (unquantified), 1.0 µM F primer, 1.0 µM R primer, 0.2 mM of each dNTP, 1 U Vent DNA polymerase (New England Biolabs), and 1X ThermoPol® reaction buffer in ddH_2O. The PCR temperature profile was one cycle 94 ºC for 5 min followed by 29 cycles 94 ºC for 45 sec, 63 ºC for 30 sec, 72 ºC for 1 min, and a final extension at 72 ºC for 10 min. The annealing temperature for CTD PCR amplification was 59 ºC.

2.1.2 Agarose Gel Electrophoresis and Gel Extraction

Unless otherwise noted, samples were run on 1.2% agarose gels stained with SYBR® Safe DNA Gel Stain (Invitrogen). Extractions were performed using the GeneJet Gel Extraction Kit (ThermoFisher Scientific) as per manufacturer instructions.

2.1.3 DNA Double Digestion

In order to ligate the amplified genes (*dnaK* and its domains) with the cloning vector (pBAD/*Myc*-His B) both were digested with XhoI and EcoRI-HF (High Fidelity, New England Biolabs). *dnaK*, NTD, CTD, and plasmid double digestions were carried out in 50 µL reaction volumes containing 250 ng PCR-amplified gene or 650 ng plasmid, 1X CutSmart™ buffer (New England Biolabs), 2.5U XhoI, 2.5U EcoRI in ddH_2O. Reactions were left at 37 ºC for 3 hours, followed by 90 ºC for 5 min to inactivate endonucleases.

2.1.4 DNA Ligation

Ligation of *dnaK* was performed in 20 µL reaction volume containing 16 ng digested vector, 27 ng digested *dnaK*, 1X T4 DNA Ligase buffer, 0.5U T4 DNA Ligase (Bio Basic, Canada) in ddH_2O. Ligation of digested NTD was performed as above with 65 ng digested NTD and ligation of the CTD was performed with 23 ng digested CTD. Ligations were left at room temperature for 3 hours.

2.1.5 Preparation of Electrocompetent Cells

Individual DH5α colonies grown on LB medium (25 g/L) were picked, placed into 14 mL culture tubes containing 5 mL LB broth, and grown overnight at 37 ºC, 250 RPM. Next, 200 mL LB broth was inoculated with 2 mL of the overnight culture and grown at 37 ºC, 250 RPM, until an OD_{600} of ~0.45 was reach. Cells were then placed into 50 mL BD Falcon™ tubes (BD Biosciences) and chilled on ice for 15 min. Cells were pelleted at 5200 x g for 10 min at 4 ºC, then resuspended in 200 mL ice cold 10% glycerol and pelleted at 5200 x g for 10 min at 4 ºC. This was repeated with 100 mL and 50 mL chilled 10% glycerol until pellets were resuspended in 2 mL 10% glycerol, at which point 40 µL aliquots were transferred into separate, sterile, 1.5 mL microfuge tubes on ice. Cells were immediately flash-frozen with liquid nitrogen and stored at -80 ºC. The same procedure was carried out in preparing electrocompetent Top10 *E. coli*.

2.1.6 Transformation of Electrocompetent Cells

In order to transform electrocompetent cells, ligations were purified using phenol/chloroform extraction. Briefly, 1:1 ratio phenol-chloroform of equal volume to each ligation reaction was used, briefly vortexed, and centrifuged at 13,000 RPM for 5 min. Top, aqueous, layer was removed and placed in 1.5 mL microfuge tube with equal volume pure chloroform. Samples were briefly vortexed and centrifuged at 13,000 RPM for 5 min. Top, aqueous, layer was removed and placed in 1.5 mL microfuge tube. Phenol/chloroform extractions were followed by ethanol precipitation. Concentrations

were determined using the Nanovue Plus spectrophotometer (GE Healthcare). Transformation of *E. coli* DH5α cells was conducted by placing 40 µL of electrocompetent cells into electroporation cuvettes with ligation product or an appropriate control. Addition of undigested plasmid (1 ng) was used as a positive control. Ligation products used: 18 ng vector-*dnaK*, 11.5 ng vector-NTD, 25 ng vector-CTD. Cells were then electroporated at 1650V using an Electroporator 2510 (Eppendorf), transferred to 1.5 mL microfuge tubes containing 700 µL LB and incubated at 37 ºC, 250 RPM, for 1 hour. Tubes were then centrifuged at 5000 RPM (Centrifuge 5424, Eppendorf) for 1 min and 600 µL of LB broth was then removed. Pellets were resuspended in the remaining LB broth, spread on LB-ampicillin plates (100 µg/mL), and incubated overnight at 37 ºC. Transformation of Top 10 *E. coli* cells was conducted as above except that 100 ng vector-gene was used in each instance.

2.1.7 Transformed *E. coli* DH5α Minipreps

To obtain vector-*dnaK*, -NTD, and -CTD from transformed *E. coli* DH5α, individual colonies were picked and enriched. Three vector-*dnaK* colonies were picked and placed individually into 14 mL culture tubes containing 5 mL LB-amp (100 µg/mL). Cultures were incubated at 37 ºC, 250 RPM, overnight. Five colonies each of vector-NTD and vector-CTD were picked and enriched as above. To isolate cloned vectors, enriched colonies were transferred from culture tubes into 1.5 mL microfuge tubes and centrifuged at 14,000 RPM for 1 min. Supernatants were discarded and pellets were

resuspended in 600 µL ddH_2O. Minipreps were conducted using the PureYield™ Plasmid Miniprep kit (Promega) as per manufacturer instructions.

2.1.8 Insert Identification and Sequencing

In order to check whether gene insertions were successful, PCR amplification was performed on minipreped vectors as above, except 0.5 µL of minipreped plasmid was used as template. PCR-amplified samples were run on 1.2% agarose gels and visualized with a Typhoon™ 9410 imager (GE Healthcare). To determine the nucleotide sequence of inserted genes, pertinent miniprep samples (100 ng/µL) were sent to MOBIX for sequencing using vector-specific primers (2 µM): F 5' ATGCCATAGCATTTTTATCC 3' and R 5' GATTTAATCTGTATCAGG 3'. To determine the nucleotide sequence of full length *dnaK*, a second F primer was used (F_2): 5' CAGGGCTTGATGTTAAGAG 3' in order to cover the full length of the gene. Results were analyzed manually using FinchTV software (Geospiza). Nucleotide sequences were converted to amino acid sequences using the translate tool by ExPASy (http://web.expasy.org/translate/). Amino acid sequence similarities were obtained by utilizing the BLAST program by the National Center for Biotechnology Information (NCBI) website.

2.2 Protein Expression and Solubility Assay

To check whether full length *dnaK*, NTD and CTD were expressed and whether the expressed products were soluble, a protein solubility assay was conducted. Individual colonies from Top10 *E. coli* transformation plates were picked and transferred to 14 mL

culture tubes containing 5 mL LB-ampicllin (100 µg/mL). An individual colony from a transformed culture plate containing cells with the empty vector only was cultured as above, as a negative control for expression. Cultures were incubated at 37 ºC, 250 RPM, for 16 hours. A 1/100 subculture was then performed into fresh LB-amp. Cells were grown at 37 ºC, 250 RPM, until an OD_{600} of ~0.50 was reached, at which point 700 µL aliquots were collected from each sample (uninduced pellet sample) and centrifuged at 13,000 RPM for 3 min. L-arabinose (0.2% final) was then added to the remainder of the cultures to induce expression of the pBAD/*myc*-His B vectors. Cells were incubated at 37 ºC, 250 RPM for 4 hours, at which point 350 µL aliquots were collected (induced pellet sample) and centrifuged as above. At the same time, 4.5 mL aliquots were collected (induced soluble sample) from each sample and centrifuged at 13,000 RPM. The induced-soluble pellets were resuspended in 200 µL 20 mM TRIS (1.4 mM β-mercaptoethanol, pH 8.0). Lysozyme was added to a final concentration of 1 mg/mL and samples were incubated for 30 min on ice. NaCl and lauryldimethylamine oxide (LDAO) were then added to a final concentration of 0.5 M and 0.05%, respectively, followed by incubation on ice for 30 min. Samples were then spun at 13,000 RPM for 20 min. Supernatant (20 µL) aliquots were mixed with equal volume 2X SDS loading buffer. The remaining pellets (induced and uninduced) were resuspended in 50 µL 2X SDS loading buffer. Samples were then placed on a 90 ºC heating block for 10 min. Samples (15 µL each) were run on 12% SDS-PAGE gels. Bands were visualized by using the Bio-Safe™ Coomassie G-250 Staining Solution (Bio-Rad) as per manufacturer instructions.

2.3 Large Scale Expression and Purification

Transformed *E. coli* Top10 (DnaK, NTD, and CTD) were cultured in LB-ampicllin (100 µg/mL) at 37 ºC, 250 RPM, for 16 hours. A 1/100 subculture was then performed into fresh LB-amp. Cells were grown at 37 ºC, 250 RPM, until an OD_{600} of ~0.70 was reached. L-arabinose (0.2% final) was then added to the cultures to induce expression. Cells were incubated at 37 ºC, 250 RPM for 4 h. After induction, cells were spun down at 5000 x g for 15 min, 4 ºC. The supernatant was removed, pellets were resuspended in chilled 1X PBS and pelleted once again. Remaining supernatant was removed and pellets were flash frozen and stored at -80 ºC. Cell pellets were thawed on ice and resuspended in 20 mL Ni A buffer (50 mM TRIS-Cl, pH 8.0, 500 mM NaCl, 20 mM imidazole, 1.4 mM β-mercaptoethanol, 5% glycerol) per L pellet. Resuspended cells were then lysed using a Constant Systems Cell Disruptor at 20 kpsi. PMSF was added to the collection tube to 100 mM. Lysates were spun down at 20,500 RPM for 40 min, 4 ºC. The supernatant was removed and applied to a HisTrap HP (GE Life Sciences) Ni-NTA column, equilibrated with Ni A buffer. Samples were eluted using a gradient of 30 to 300 mM imidazole and fractions were collected. Fractions containing the protein of interest (DnaK, NTD, or CTD) were pooled and diluted with Q A buffer (20 mM TRIS-Cl, pH 8.0, 5 mM β-mercaptoethanol, 5% glycerol) to 100 mM NaCl to be further purified by anion exchange chromatography. Diluted samples were applied to a HiTrap Q HP (GE Life Sciences) column calibrated with Q A buffer. Proteins were eluted using a gradient of 100 to 500 mM NaCl and fractions were collected. Fractions containing the protein of

interest were pooled and concentrated with either a 30 kDa (DnaK, NTD) or a 10 kDa (CTD) MW cut-off Amicon Ultra centrifugal filter (EMD Millipore).

2.4 *In vitro* Selection

2.4.1 Selection Libraries

DnaK, NTD, and CTD libraries all consisted of a 40 nt random domain flanked by primer binding sites. DnaK library: 5' –AACAGCCCAAGTCAGGTCTAT – N_{40} – GTA GTT CTG ACT GAC GCT TGG– 3', NTD library: 5' – AGTGTCAGCCAG TATAACCCA – N_{40} – TAACCCGGAGTGAACACCTA – 3', and CTD library: 5' – CCAAACAGCCAGTATGTCAGT – N_{40} – TGGGTTATACTGGCTGACACT – 3'. All DNA oligonucleotides were commercially synthesized by Integrated DNA Technologies (Coralville, Iowa) and provided desalted. Further purification was done by 10% denaturing PAGE followed by elution and ethanol precipitation. Concentrations were determined using the Nanovue Plus spectrophotometer (GE Healthcare).

2.4.2 Selection Setup

The initial rounds of selection began with 1 nmol of each respective library in 100 µL selection buffer (SB) (50 mM HEPES, pH 7.6, 100 mM NaCl, 10 mM KCl, 2 mM $MgCl_2$, 0.05% Tween-20). Libraries were heated at 90°C for 5 min followed by rapid cooling on ice for 10 min and then equilibration at room temp for 10 min. Proteins were prepared for *in vitro* selection by immobilization on HisPur Ni-NTA magnetic beads

(ThermoFisher Scientific) as per manufacturer instructions, using SB, at 4°C with moderate rotation. Equilibrated libraries were then incubated with ~100 nM bead-bound protein in 1 mL total SB. This selection mixture was equilibrated at room temp for 30 min with moderate rotation. Beads were then isolated using neodymium magnets (MagRack 6, GE Lice Sciences) and supernatant was discarded. Beads were then washed once with resuspension and twice without using SB. Elusion buffer (EB) (SB with 300 mM imidazole) was added and the mixture heated at 90°C for 5 min, followed by vigorous shaking for 15 min at room temp. Beads were once again isolated with magnets, the supernatant was collected and ethanol precipitated. Pellets were resuspended with 20 µL ddH_2O. Counter selection was performed using 40 µL HisPur Ni-NTA magnetic beads alone, before the 4th round of selection and every round thereafter. A separate counter selection was performed, during the positive selections starting round 6, using 1 nmol BSA, which cannot be bound to the Ni-NTA magnetic beads. Thus, non-specific sequences bound to BSA would be discarded with the supernatant.

2.4.3 Two Step PCR Amplification

Primary PCR amplification was performed using the following primers: F_{DnaK} 5' AACAGCCCAAGTCAGGTCTAT 3', $R1_{DnaK}$ 5' CCAAGCGTCAGTCAGAACTAC 3', F_{NTD} 5' AGTGTCAGCCAGTATAACCCA 3', $R1_{NTD}$ 5' TAGGTGTTCACTCCGGGT TA 3', F_{CTD} 5' CCAAACAGCCAGTATGTCAGT 3' and $R1_{CTD}$ 5' AGTGTCAGCCAG TATAACCCA 3'. DnaK, NTD, and CTD PCR reactions were all performed in 50 µL

reaction volumes containing: 2 µL template DNA (resuspended pellets), 1.0 µM F primer, 1.0 µM R primer, 0.2 mM of each dNTP, 2.5 U Taq DNA polymerase (GenScript), and 1X reaction buffer in ddH_2O. The PCR temperature profile was one cycle 94 ºC for 1 min followed by 12 cycles 94 ºC for 30 sec, 50 ºC for 45 sec, 72 ºC for 45 sec, and a final extension at 72 ºC for 5 min. Agarose gel electrophoresis (2%) was performed to check proper amplification of products. Secondary PCR amplifications were performed using primary PCR products as template and the same protocol. However, the reverse primers contained an 18 atom hexaethyleneglycol spacer linked to a 15 nt string of thymidine residues, and thus cannot be fully extended. This results in a shorter sense strand that can be isolated and used for the next round of selection. The following primers were used: $R2_{DnaK}$ 5' TTTTTTTTTTTTTTT /iSp18/CCAAGCGTCA GTCAGAACTAC 3', $R2_{NTD}$ 5' TTTTTTTTTTTTTTT /iSp18/ TAGGTGTTCACTCCG GGTTA 3', and $R2_{CTD}$ 5' TTTTTTTTTTTTTTT /iSp18/AGT GTCAGCCAGTATAAC CCA. DNA strands were separated using 10% denaturing PAGE, sense strands were excised, eluted and ethanol precipitated. Pellets were resuspended in 20 µL ddH_2O to be used in the next round of selection.

After 8 rounds of selection for NTD and 10 rounds for DnaK and CTD, samples were deep sequenced using MiSeq DNA sequencer (Illumina) at the Farncombe Metagenomics Facility, McMaster University. Paired-end reads were merged and trimmed of primers. Non-complementary sequences were discarded and the remaining sequences were ranked by copy number.

2.5 Aptamer Characterization

2.5.1 Electrophoretic Mobility Shift Assay

Aptamer DNA (100 pmol) was 5'-phoshporylated using 10U T4 polynucleotide kinase (PNK) at 37 °C with 10 μCi [γ-^{32}P]ATP for 30 min. Radiolabeled DNA was then purified by 10% denaturing PAGE and resuspended in SB after elution and ethanol precipitation. Aptamer and target protein were then incubated in SB (20 μL total volume) at room temperature for 30 min with gentle shaking. Samples were then mixed with 1X final concentration native loading dye (1X SB, 10% glycerol, 0.01% bromophenol blue) and immediate loaded onto 7% native PAGE gels. Gels were run at 425 V for 45 min at 4 °C in 0.5X HB (25 mM HEPES, pH 7.6, 50 mM boric acid, 1.5% glycerol). Gels were visualized on storage phosphor screens scanned on a Typhoon Imager (GE Life Sciences).

2.5.2 Protein-Bound Bead Binding Assay

Aptamer DNA was radiolabeled and proteins were bound to HisPur Ni-NTA magnetic beads, as described above. Aptamers were then incubated with bead-bound protein in SB (30 μL total volume) for 30 min, room temperature, with gentle shaking. Beads were separated using neodymium magnets and supernatant collected. Beads were washed three times with SB. Samples were then run on 10% denaturing PAGE gels and visualized on storage phosphor screens scanned on a Typhoon Imager (GE Life Sciences).

2.5.3 Construction of the Heterobivalent Aptamer

To construct the heterobivalent aptamer, CAp1-M1 was 5'-phosphorylated as described above, but with cold (non-radioactive) ATP. Then equimolar (500 pmol each) NAp1-M4-T10, CAp1-M1, and N4-C1 Splint were combined and heated at 90 °C then allowed to gradually cool to room temperature. Thereafter they were ligated with 5U T4 DNA ligase (50 µL total volume) at 20 °C, 2 h. Ligations were purified by 10% denaturing PAGE.

2.5.4 Bacterial Cultures and FIM Preparation

E. coli and *B. subtilis* were cultured in LB media from single colony origin glycerol stocks, 37 °C, overnight. *C. difficile* and *F. nucleatum* were cultured in cooked meat broth from single colony origin glycerol stocks, in an anaerobic chamber (5% CO_2, 10% H_2, 85% N_2), 37 °C, for 36 h. Small aliquots were taken from each culture, serially diluted, and plated on LB-agar or cooked meat broth-agar plates, grown under the same conditions as above, in order to determine colony forming units (CFU)/ mL. The remaining cultures were lysed as described above. Lysates were then taken and passed through 100 kDa MW cut-off Amicon Ultra centrifugal filters (EMD Millipore). Flow through was collected and passed through 50 kDa MW cut-off filters. Concentrates were collected and stored at -80 °C.

Chapter 3 – Results

3.1 Molecular Cloning and Sequence Analysis

In order to employ the proposed Multiple Domain Selection (MDS) strategy, the target of interest (DnaK) needed to be divided into its constituent domains. The likely linker region was identified (Figure 1.7) and used to split the protein, and thus the gene, into the NTD and CTD. Since the linker is composed of 18 amino acids, the first nine were included with the NTD and the remaining nine were to remain part of the CTD. The whole gene (*dnaK*) and the respective portions of the gene (NTD & CTD) were amplified using PCR and the appropriate primers. The F and R primers contained XhoI and EcoRI restriction sites, respectively, in order to ligate the amplified sequences into the vector of choice. The first PCR reaction was then used as template for a second reaction using a higher anneal temperature in order to get more specific products. Results were visualized on an agarose gel (Figure 3.1). Full length *dnaK* (~1800 bp) as well as the NTD (~1100 bp) and CTD (~750 bp) were successfully amplified.

In order to ligate the PCR-amplified DnaK gene and its domains into the plasmid vector (pBAD/myc-His B), each needed to be digested by the chosen restriction enzymes corresponding to the restriction sites on the plasmid. PCR amplified *dnaK*, -NTD and -CTD as well as the pBAD/myc-His B vector were digested using XhoI and EcoRI-HF (High Fidelity) restriction endonucleases. Digestion reactions were visualized on an agarose gel (Figures 3.2 and 3.3).

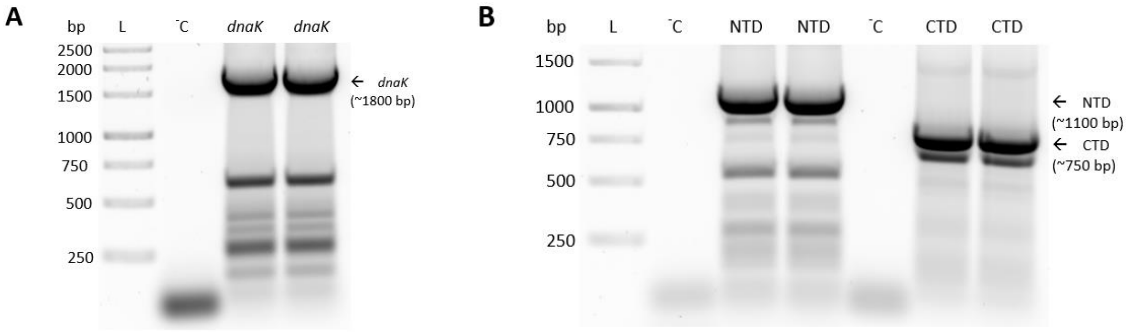

Figure 3.1. PCR amplification of *C. difficile dnaK*, NTD and CTD. (A) *dnaK* PCR reactions were performed in 50 µL reaction volume containing: 5 µL pre-amplified DNA, 1.0 µM F and R primers, 0.2 mM dNTPs, 1 U Vent DNA polymerase, and 1X ThermoPol® reaction buffer. The PCR temperature profile was one cycle 94 ºC for 5 min followed by 29 cycles 94 ºC for 45 sec, 64 ºC for 30 sec, 72 ºC for 1 min, and a final extension at 72 ºC for 8 min. (B) NTD and CTD PCR reactions were performed as above except the annealing temperatures were 62 ºC and 59 ºC, respectively. Negative controls (-C) were performed without template DNA. L represents O'GeneRuler 1 kb DNA Ladder, Ready-to-Use, 250-10,000 bp.

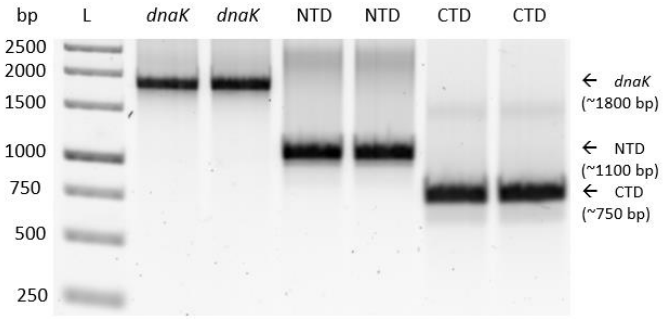

Figure 3.2. Double digestion of PCR amplified *dnaK*, NTD and CTD. *dnaK* double digestion was carried out in 50 µL reaction volume containing 16.8 ng/µL PCR-amplified *dnaK*, 1X CutSmart™ buffer, 2.5U each XhoI and EcoRI-HF. NTD and CTD double digestions were performed as above, except 18.0 ng/µL of PCR-amplified product was used. Reactions were left at 37 ºC for 3 hours, followed by inactivation at 90 ºC for 5 min. L represents O'GeneRuler 1 kb DNA Ladder, Ready-to-Use, 250-10,000 bp.

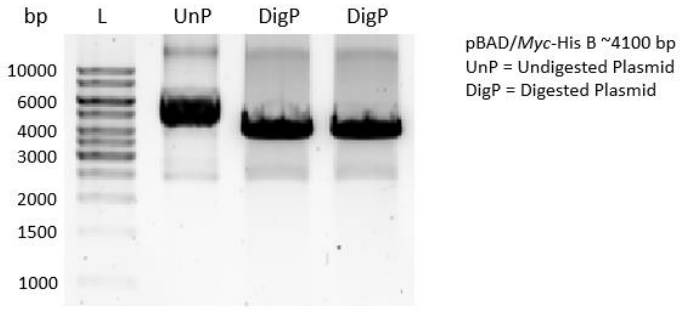

Figure 3.3. Double digestion of pBAD/*Myc*-His B plasmid. Digestions were carried out in 50 µL reaction volumes containing 41 ng/µL pBAD/*Myc*-His B, 1X CutSmart™ buffer, 2U each XhoI, and EcoRI-HF. Reactions were left at 37 ºC for 3 hours, followed by inactivation at 90 ºC for 5 min. L represents O'GeneRuler 1 kb DNA Ladder, Ready-to-Use, 250-10,000 bp.

Digested *dnaK*, NTD, and CTD products were extracted and ligated to the digested plasmid using T4 DNA ligase. Ligations were then purified using phenol/chloroform extraction followed by ethanol precipitation. In order to propagate the clones, electrocompetent *E. coli* DH5α cells were transformed with ligated vector-*dnaK* (Full, NTD, or CTD). Individual colonies were then cultured and enriched overnight in LB-ampicillin. The resulting cultures were minipreped to extract and purify the clones. Before further analysis could be conducted, the individual colonies and the clones they harbor were tested to make sure they contained the genes of interest. To do this, the purified plasmids were used as templates to PCR amplify *dnaK*, the NTD and CTD. If correctly incorporated within the plasmid, the genes will be amplified and present at the correct lengths when run on an agarose gel. The PCR reactions were run and visualized on an agarose gel (Figure 3.4). All samples were successfully amplified. In order to

confirm the exact sequences contained within each plasmid from their respective colonies, K1-K3, N1-N3, and C1, C3, and C5 were sent for Sanger sequencing. Sequences were then manually analyzed, converted from nucleotide to amino acid sequences, and BLAST was used to determine sequence similarity to the DnaK protein and its domains (Figure 3.5 A, B, and C). K1, K2, N1, N3, C3, and C5 resulted in 100% amino acid identity to their respective *C. difficile* NAP7 protein or domain. K3, N2, and C1 could not be fully analyzed due to inadequate sequencing data.

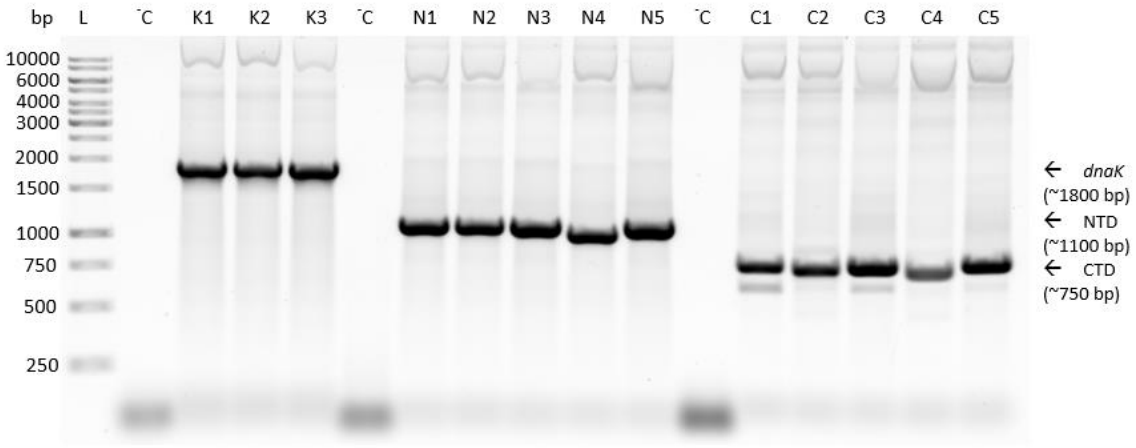

Figure 3.4. PCR amplification of *dnaK*, NTD and CTD from minipreped plasmids. Transformed *E. coli* DH5α colonies were individually picked, enriched, and minipreped to obtain plasmid DNA. *dnaK* (K1-3), NTD (N1-5), and CTD (C1-5) PCR reactions were performed in 50 µL reaction volumes containing: 5 µL template DNA, 1.0 µM F and R primers, 0.2 mM dNTPs, 1 U Vent DNA polymerase, and 1X ThermoPol® reaction buffer. The *dnaK* PCR temperature profile was one cycle 94 ºC for 5 min followed by 29 cycles 94 ºC for 45 sec, 63 ºC for 30 sec, 72 ºC for 1 min, and a final extension at 72 ºC for 8 min. The same cycling parameters were used for the NTD and CTD except the annealing temperatures were 62 ºC and 59 ºC, respectively. Negative controls (-C) were performed without template DNA. L represents O'GeneRuler 1 kb DNA Ladder, Ready-to-Use, 250-10,000 bp.

A

Score	Expect	Method	Identities	Positives	Gaps
1236 bits(3197)	0.0	Compositional matrix adjust.	615/615(100%)	615/615(100%)	0/615(0%)

B

Score	Expect	Method	Identities	Positives	Gaps
723 bits(1866)	0.0	Compositional matrix adjust.	359/359(100%)	359/359(100%)	0/359(0%)

C

Score	Expect	Method	Identities	Positives	Gaps
509 bits(1311)	2e-175	Compositional matrix adjust.	256/256(100%)	256/256(100%)	0/256(0%)

Figure 3.5. Amino acid sequence similarity of cloned DnaK, NTD and CTD samples with *C. difficile* NAP7. (A) Sequence alignment of cloned DnaK, (B) sequence alignment of cloned NTD, and (C) sequence alignment of cloned CTD. Pertinent miniprep samples were sent to MOBIX for sequencing using plasmid-specific primers. For the full *dnaK* gene-containing plasmid, a third, middle primer, was used in order to cover the full range of the sequence. Results were analyzed manually using FinchTV software. Nucleotide sequences were converted to amino acid sequences using the translate tool by ExPASy. Amino acid sequence alignments were obtained by utilizing the BLAST program at the NCBI website.

3.2 Protein Expression and Purification

In order to determine whether the cloned genes would be expressed and that their protein products were soluble, a protein solubility assay was conducted. Purified plasmids containing *dnaK*, NTD and CTD were each transformed into *E. coli* Top10, a compatible cell line for the expression of cloned genes contained within an arabinose-inducible plasmid. An individual colony from each construct was selected and grown in LB-ampicillin media (50 µg/mL). Cultures were grown to mid-log phase (OD_{600} of 0.5-0.7) to ensure a healthy and active population of bacteria ready for large production of protein. The cells were induced with 0.20% L-arabinose for 4 hours; uninduced and

induced samples were collected and run on SDS-PAGE gels (Figure 3.6). Over-expression of DnaK, NTD and CTD was evident at the appropriate band sizes in the induced fractions (pellet and soluble) but not the uninduced fraction, suggesting the expression of recombinant protein was successful. Because the expressed proteins were found in both induced fractions and not the pellet fraction alone suggests that they are soluble and did not form aggregates.

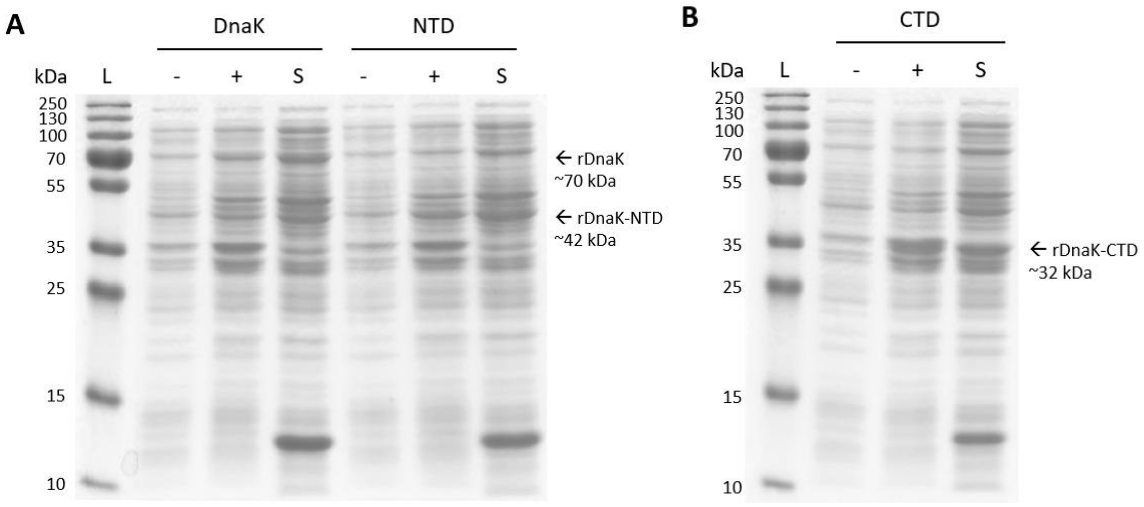

Figure 3.6. Expression and solubility of DnaK, NTD and CTD in Top10 cells. (A) Solubility assay for DnaK and NTD in Top10, (B) solubility assay for CTD in Top10. Transformed *E. coli* Top10 were cultured in LB-ampicillin and grown at 37 ºC, 250 RPM, until an OD_{600} of 0.5-0.7. Uninduced aliquots (-) were collected from each sample and pelleted. Cultures were induced with L-arabinose and grown at 37 ºC, 250 RPM for 4 hours, at which point induced aliquots (+) were collected and pelleted. At the same time, induced aliquot were also taken, chemically lysed, and centrifuged. The soluble fractions (S) were then collected. Samples were run on 12% SDS-PAGE gels. L represents PageRuler Plus Prestained Protein Ladder (ThermoFisher Scientific).

After confirmation of soluble protein expression, large scale production (250 mL cultures) was performed as with the protein solubility assay. Because the cloned proteins are fused with a His_6 tag, the purifications were performed step-wise using the Qiagen Ni-NTA Fast Start Kit. Flow through, wash, and elusion samples were run on SDS-PAGE gels (Figure 3.7). Expression and purification of DnaK, NTD, and CTD was successful, however, and with the exception of E2 for the NTD, we wanted to have samples that were more pure before starting *in vitro* selection. Purification was attempted again, this time using cobalt based columns and specially made buffers to reduce background binding. Purification was successful but still contained impurities (not shown).

To get around the purification problems, another strategy was implemented: an imidazole gradient, rather than step-wise elusion, was used to elute the recombinant proteins. In this setup, cell lysates are loaded onto a nickel column but eluted with a gradual increase in imidazole over time, while fractions are collected throughout. This way, many fractions can be identified and tested, and those pure enough can be used for downstream applications. This was done using an AKTA FPLC system (GE Life Sciences) with select fractions run on SDS-PAGE gels (Figure 3.8). Furthermore, cells were lysed using a high pressure homogenizer in order to get more efficient lysis and more product. As can be seen in the gels, there was efficient lysis and product, however, the purity, with the exception of the NTD, was still subpar, particularly for the full DnaK protein.

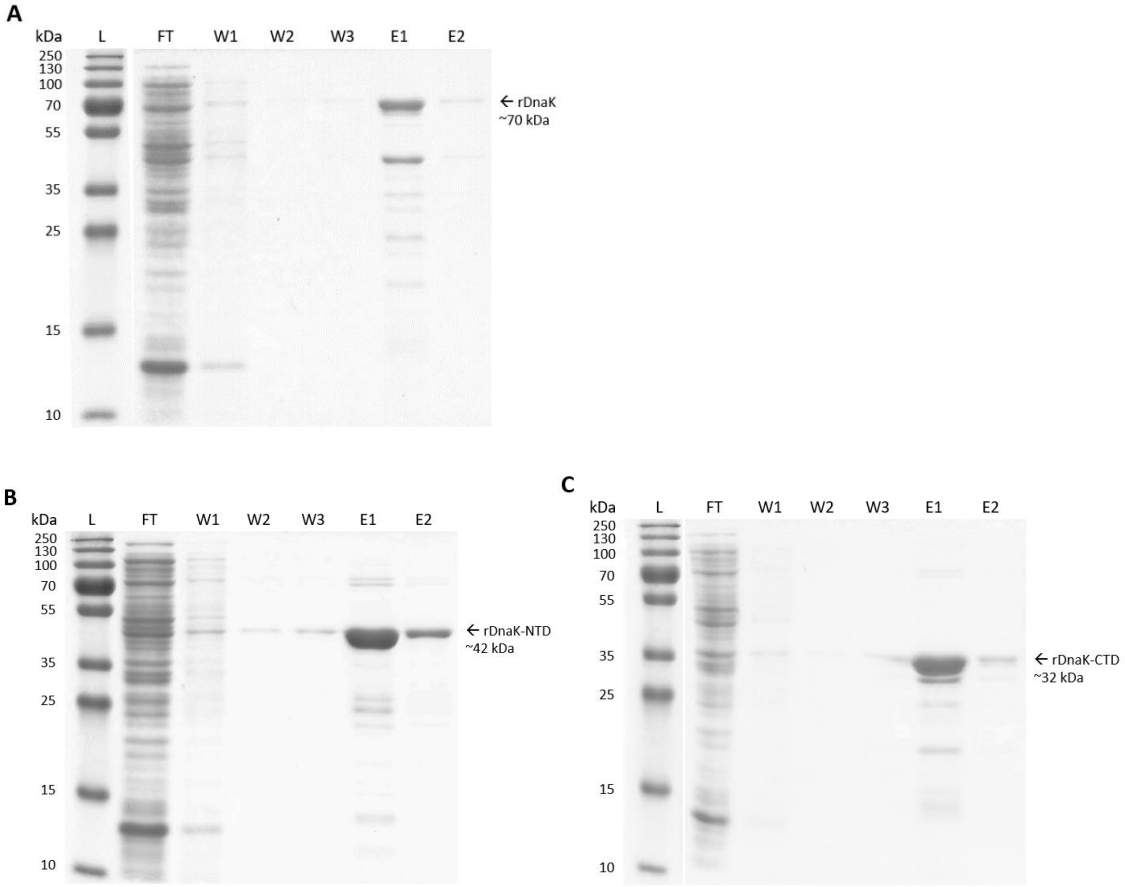

Figure 3.7. Purification of DnaK, NTD and CTD with a Ni-NTA column. (A) Purification of DnaK, (B) purification of NTD, and (C) purification of CTD visualized on SDS-PAGE gels. Expression was induced in 250 mL of culture when the OD_{600} reached 0.5-0.7. After 4 hours, 200 mL from each sample (DnaK, NTD, and CTD) were spun down and chemically lysed. Purification of samples was performed using the Qiagen Ni-NTA Fast Start Kit as per instructions provided therein for purification under native conditions. Samples were run on 12% SDS-PAGE gels. L represents the PageRuler Plus Prestained Protein Ladder (ThermoFisher Scientific); FT represents the flow through; W1-3 represent washes 1-3; E1-2 represent elusions 1-2. Gels (A) and (C) were cropped in order to position the ladder on the left hand side.

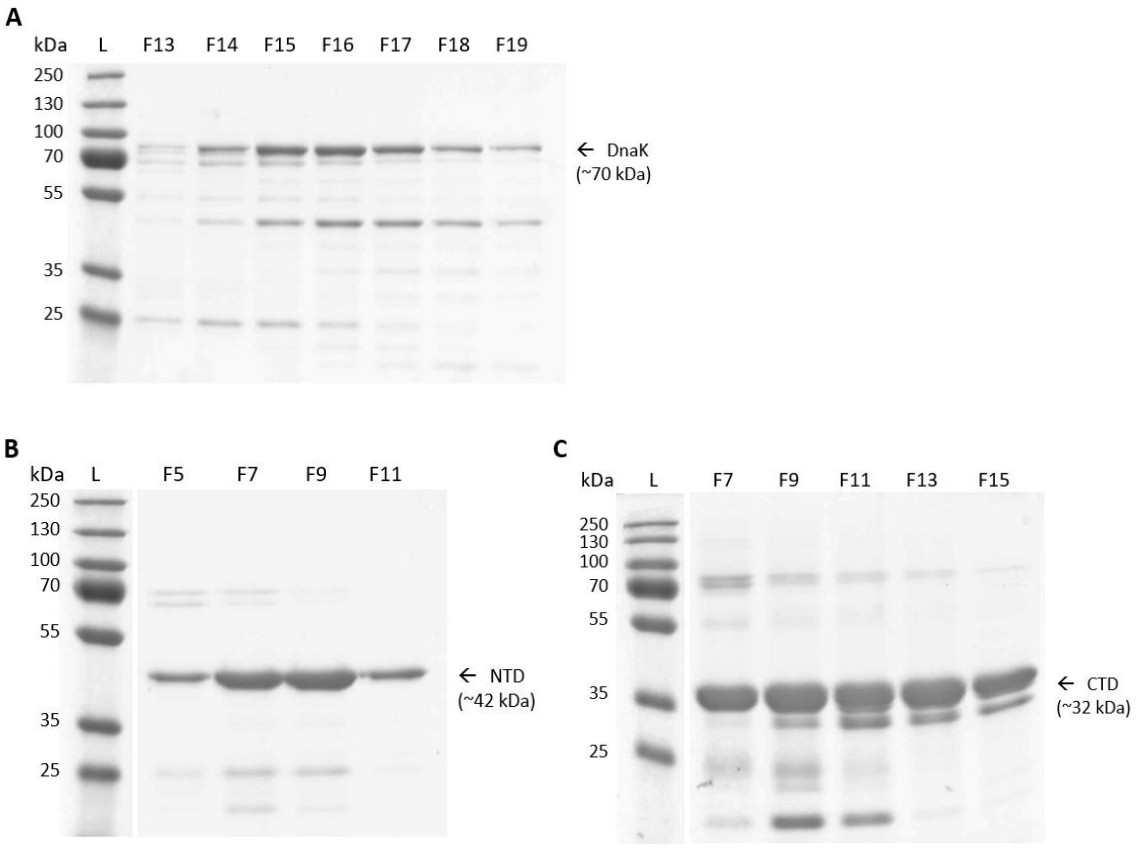

Figure 3.8. Purification of DnaK, NTD and CTD with an elution gradient. (A) Purification of DnaK, (B) purification of NTD, and (C) purification of CTD visualized on SDS-PAGE gels. Expression was induced in 1 L of culture when the OD_{600} reached 0.5-0.7. After 4 hours, samples (DnaK, NTD, and CTD) were spun down and mechanically lysed. Purification of samples was performed using a HisTrap HP (GE Life Sciences) Ni-NTA column placed in an AKTA FPLC system with an imidazole gradient (30-300 mM). Select fractions were then run on 12% SDS-PAGE gels. L represents the PageRuler Plus Prestained Protein Ladder (Thermo Scientific); F represents individual fractions. Gels (B) and (C) were cropped to remove unnecessary lanes.

In order to purify DnaK and CTD further, a secondary purification using ion-exchange chromatography was conducted. For this, a Q-sepharose column was used to bind each of the proteins. Because DnaK and CTD have a theoretical pI of 4.82 and 4.68, respectively, they will be negatively charged at pH 8.0, the pH of the buffer used, and will bind the positively charged Q-sepharose column. To elute the proteins from the column, a salt gradient is used (100-500 mM NaCl) to disrupt the ionic interactions. Select fractions were then run on SDS-PAGE gels (Figure 3.9).

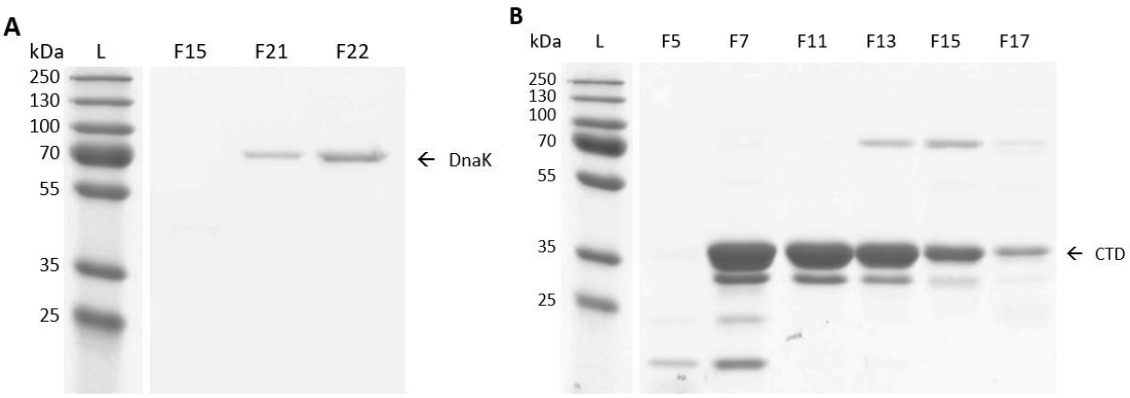

Figure 3.9. Secondary purification of DnaK and CTD using anion-exchange. (A) Purification of DnaK and (B) purification of CTD visualized on SDS-PAGE gels. Select fractions from DnaK and CTD nickel column purifications were combined and diluted with Q A buffer (20 mM TRIS-Cl, pH 8.0, 5% glycerol, 5 mM 2-mercaptoethanol) in order to reduce their NaCl concentration to 100 mM. Diluted samples were then loaded onto a HiTrap Q HP (GE Life Sciences) column and placed into an AKTA FPLC system. The column was washed and an elution gradient was carried out from 100-500 mM NaCl. Select fractions were then run on 12% SDS-PAGE gels. L represents the PageRuler Plus Prestained Protein Ladder (ThermoFisher Scientific); F represents individual fractions. Gels (A) and (B) were cropped to remove unnecessary lanes.

3.3 Selection of DNA aptamers that bind DnaK, NTD, and CTD

The *in vitro* selection strategy for obtaining aptamers, SELEX (Systematic Evolution of Ligands by Exponential Enrichment), is a relatively simple, but powerful technique (Figure 1.2). One begins with a large, randomized pool of DNA sequences, in our case 10^{14} oligonucleotides. These sequences share common primer binding sites at the 5'- and 3'- ends, flanking an internal 40 nucleotide random domain. The extremely diverse and large initial pool is first subjected to the target of interest, and washed to remove non-specific sequences. Those that bound the target of interest are then eluted and PCR amplified. The enriched pool is then subjected to several more rounds of selection, until target-binding is identified, at which point it is sent for sequencing.

Since our recombinant protein constructs contain a polyhistidine tag, we chose to use HisPur Ni-NTA magnetic beads in order to capture the targets. Purified DnaK, NTD, and CTD were bound to the beads and exposed to the initial library. After 30 minutes of equilibration the beads were separated with a magnet, washed several times, and the supernatant was disposed. Any remaining DNA still bound to the bead-bound proteins were eluted. Samples were then PCR amplified, however, the R primer was constructed to be longer than the F primer with an 18-atom hexaethyleneglycol spacer, which cannot be replicated. This leads to a regular sized sense strand and a larger anti-sense strand. The PCR amplified DNA was then run on a denaturing PAGE gel where the sense strands were then excised and purified.

The enriched library was then re-exposed to bead-bound protein and the selection processes was repeated nine times. A counter selection step was also introduced starting with the 4th round of selection. Before the enriched library was exposed to the target it was first exposed to the magnetic beads alone. Any non-specific sequences which bound the beads are then discarded. An additional counter selection step was also added during the positive selections starting round 6, whereby BSA was co-incubated with the target. By introducing BSA, which does not contain a polyhistidine tag and cannot bind the Ni-NTA magnetic beads, non-specific sequences recognizing a general protein target were eliminated during the wash step.

In order to determine whether a selection is complete or has been sufficiently enriched, the library can be tested for binding to the target protein; on the other hand, it can be sent directly for deep sequencing to determine the most abundant sequences for testing. For the NTD selection, which was completed first, the former was chosen. For this, the electrophoretic mobility shift assay (EMSA) was used. First, the library was radioactively labeled with ^{32}P and then incubated with the NTD protein in solution. After equilibration, samples were loaded onto a native PAGE gel and run at a moderate voltage, so as to keep the temperature low and any complexes stable. To visualize the radioactively labelled DNA, the gel was exposed on a phosphor plate, which was then scanned to determine the migration pattern of the DNA (Figure 3.10).

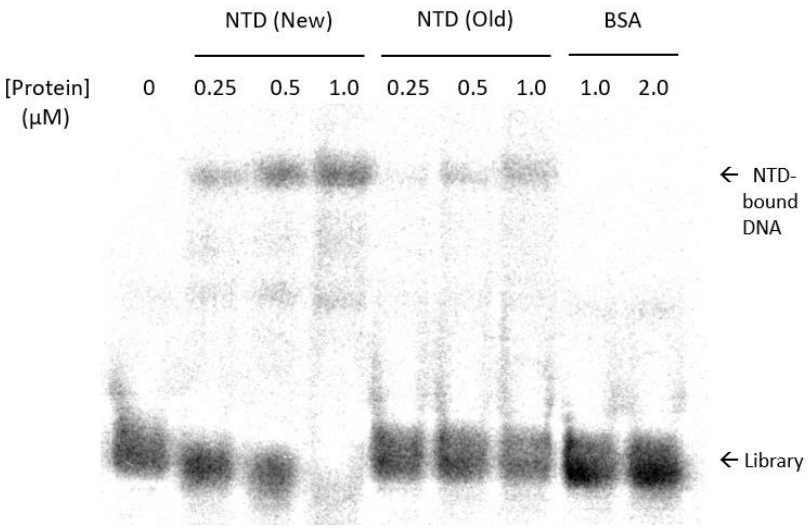

Figure 3.10. EMSA of NTD library after 8 rounds of selection. Various aliquots of radioactively labeled NTD library DNA (100 nM) were incubated with increasing concentrations of NTD protein, as well as BSA, in selection buffer (50 mM HEPES, pH 8.0, 100 mM NaCl, 10 mM KCl, 2 mM $MgCl_2$, 0.05% Tween 20) for 30 minutes. The samples were run on a 7% native PAGE gel. NTD (New) represent fresh samples of purified NTD protein, whereas NTD (Old) represent samples which were over three weeks old, stored at 4°C.

A successful selection should result in a library which experiences an upward shift on the gel, corresponding to the larger DNA-protein complex formed, compared to free DNA alone. This was indeed the case for the NTD selection. As can be seen in Figure 3.10, increasing amounts of DNA library are shifted upwards with increasing amounts of NTD, suggesting the presence of potential aptamer sequences within the library. A clear distinction can also be seen between old and fresh NTD protein, indicating degradation over time which can affect binding. Lastly, there was no indication of non-specific binding to BSA, up to 2 μM; thus, binding from within the library is

specific to the NTD protein. After confirmation of binding, the library was sent for deep sequencing to determine the most abundant sequences. In the case of the DnaK and CTD selections, after the 10^{th} round, the enriched libraries were sent immediately for deep sequencing. Once the data was analyzed, the top sequences were ranked in order of abundance (Figure 3.11). Because of the selective conditions and enrichment of the DNA library after each round of selection, these top sequences, then, represent the likeliest candidate aptamers.

3.4 Binding Characteristics of the Top Aptamer Candidates

In order to test which sequences bind to their target and which do not, the top two candidates from each selection were chemically synthesized by Integrated DNA Technologies and purified by denaturing PAGE. We then prepared an assay based on the selection protocol: the protein-bound bead binding assay. The procedure begins by capturing the target protein using the Ni-NTA magnetic beads that were used during the selection. The beads are then washed several times and resuspended in selection buffer. The ^{32}P labeled candidate aptamer is then introduced and incubated with the protein-bound beads for 30 min, with shaking, to keep the beads suspended in solution. After the equilibration period has concluded, the protein-bound beads are separated from the supernatant using a magnet and washed several times. The supernatant and beads are then run on a denaturing PAGE gel. To visualize the radioactively labelled DNA, the gel is exposed on a phosphor plate and scanned with a gel imaging system.

DnaK Top Aptamer Candidates, Round 10

Rank	Sequence	Copy No.	% of Total
1	GACCGGGTAGGATTATCCCGGCAGCTGTATGAAGCCAGTA	94780	14.6
2	CGCGGTCAACTACTGATTGAGAGAAGCATCCTTATAGCCC	92110	14.2
3	GTATGACGGCTGCAGTATTGTCGGCCGCACACGTCTGAGT	45564	7.0
4	GCGACGCTCTAGATATATGGGCGCGAGTGTCTGAATGTTC	44125	6.8
5	CATCAACATCGACAGAATGTCCCCTGAGGCGCTTCCGTGT	35062	5.4

Total Reads: 648954

NTD Top Aptamer Candidates, Round 8

Rank	Sequence	Copy No.	% of Total
1	CGCATGGTCTTGTACTCGTGCGGGCCGTCTGAGGGAAGTT	141561	34.5
2	GTGTGCGAGAAGACACACGGCATGTCTGAATGGTCTGTTG	27093	6.6
3	CGGAGCAGGTCCAGGCCATTTAATTGTGCTTACCGTAGAG	10980	2.7
4	GGACTGTGGTTATCTAACGTCTGAGTTCGGCACTAGATCC	4309	1.1
5	CGCATGGTCCTGTACTCGTGCGGGCCGTCTGAGGGAAGTT	4133	1.0

Total Reads: 409814

CTD Top Aptamer Candidates, Round 10

Rank	Sequence	Copy No.	% of Total
1	GTATTTGGTGTTGTGAGGGCGGGGGGTTTGGCGGAGATGG	35555	9.7
2	GGATAGGTGGAGTGTGAGGGTAGGGGGGTTTGGCTCCTGA	24490	6.7
3	CGGGTGGTGTGCGAGGGAATGGGGGATTGGATGGTAGTAG	23868	6.5
4	GGCCAGGGGGGATTGGTGTGGTGTGCGAGGGTGGGGCCTC	17344	4.7
5	GCGGGAAAGGGCTGAGGTGGCGTGTGAGGGAAAGGGGGTT	14765	4.0

Total Reads: 367024

Figure 3.11. List of the most abundant sequences from each selection. After the final round of selection, each of the enriched libraries from the DnaK, NTD, and CTD selections were sent for deep sequencing and analyzed. The most abundant sequences are listed along with the number of copies within the pool and the percentage of the pool the sequence represents. These sequences represent the top aptamer candidates.

If the candidate aptamer recognizes and binds the protein of interest it should be present in both the supernatant and bead fraction bands, if not, it will only be present in the supernatant fraction band.

The initial testing began with NTD aptamers 1 and 2, referring to their rank based on abundance in the final round of selection. These aptamers were initially synthesized without the 5' and 3' primer binding sites, however, the protein-bound bead binding assay as well as EMSA testing concluded little to no binding (not shown). We hypothesized that this could be due to the lack of the primer binding sites, as the sequencing data clearly suggested enrichment, and the initial EMSA using the NTD round 8 library showed binding to the target. The primer binding sites are usually removed because they are typically not involved in the binding domain of the selected aptamers. However, in this case, they seemed to be structurally important for the proper folding and binding of the aptamers. Because of the importance of the primer binding sites for the functional binding of NTD aptamers 1 and 2, they, along with the DnaK and CTD aptamers were synthesized with their primer binding sites intact. In each case we chose to test the top two aptamers from each selection (KAp1, KAp2; NAp1, NAp2; CAp1, CAp2).

Aptamer characterization began by testing the binding of each aptamer to its target as well as the other two targets (Figure 3.12). In this way, we were able to determine whether the aptamers bound their intended target, whether the domain-specific

aptamers also bound the full DnaK protein (as they should), and which of the domains, if any, would the DnaK aptamer recognize and bind. As can be seen, none of the aptamers tested bound non-specifically to the beads alone (NAp1 and NAp2 had already been tested separately for non-specific binding to the beads, which were not included in this assay). Furthermore, all aptamers also bound their intended target, indicating a successful selection. The NTD aptamers bound the NTD protein as well as the full DnaK protein, but not the CTD protein, as expected. The CTD aptamers also bound the CTD protein and full DnaK protein, but not the NTD protein, which was also expected. Interestingly, both DnaK aptamers bound the DnaK protein, but also only the NTD and not CTD protein. This suggests the NTD is the more aptagenic domain, for which sequences are more likely to bind when performing *in vitro* selection on the full DnaK protein.

3.5 Constructing the Heterobivalent Aptamer

Because all tested aptamers bound their respective targets, we chose to focus on the top aptamer for each protein moving forward (KAp1, NAp1, and CAp1). To construct the proposed heterobivalent aptamer we used a doubled stranded ligation method. First, CAp1, serving as the 3' component of the aptamer, was radiolabeled with ^{32}P in the form of a phosphate group from $\gamma^{32}P$-ATP. This serves two purposes: 1) it allows us to use radioactivity to track protein binding and 2) it phosphorylates the 5' terminal of CAp1, which is required by DNA ligase to form a phosphodiester bond with a neighbouring oligonucleotide.

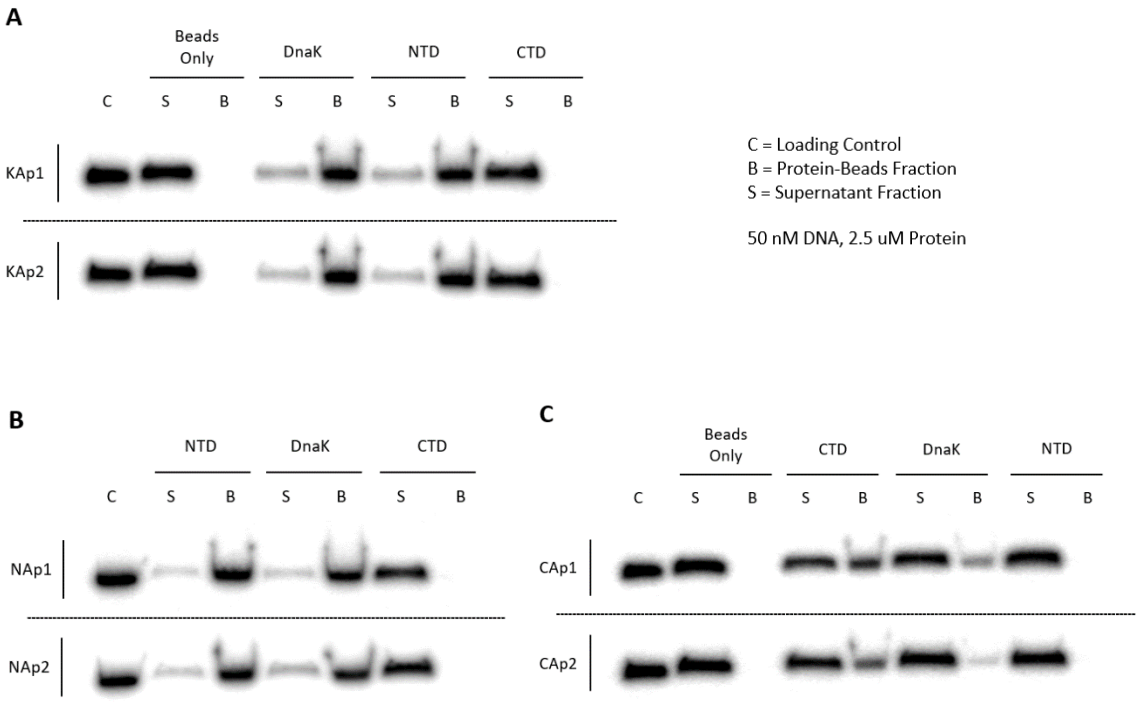

Figure 3.12. Protein-bound bead binding assay for the top DnaK, NTD, and CTD aptamers against target proteins. (A) KAp1 and KAp2 exposed to beads only (negative control), DnaK, NTD, and CTD protein. (B) NAp1 and NAp2 exposed to NTD, DnaK, and CTD protein. (C) CAp1 and CAp2 exposed to CTD, DnaK, and NTD protein. Proteins were captured with Ni-NTA magnetic beads and exposed to radiolabeled aptamers for 30 min with shaking. Beads were then separated from the supernatant using a magnet and washed several times with selection buffer. Supernatant and beads were then run on 10% denaturing PAGE gels.

Because only a fraction of the CAp1 can be labeled with hot (radioactive) ATP, cold ATP was subsequently used to complete phosphorylation of all CAp1 5’ termini. Next, equimolar concentrations of NAp1 and splint were added. The splint is an oligonucleotide sequence complimentary to the 3’ terminus of NAp1 and 5’ terminus of CAp1, which allows the two to be joined side by side in a double strand for DNA ligase

to perform the ligation. The sequences were then heated and cooled gradually to form the correct structure, at which point T4 DNA ligase was added to ligate NAp1 to CAp1.

Initial testing of this heterobivalent construct composed of unmodified NAp1 and CAp1 resulted in no binding to DnaK protein (not shown). This was not unexpected due to a glaring problem when combining the NTD and CTD aptamers: their 5' and 3' primer binding sites are complimentary to one another, resulting in the two sites binding each other and poisoning the resulting structure (Figure 3.13). As can be seen, the predicted secondary structure of the combined aptamers does not resemble that of the individual aptamers, thus the heterobivalent aptamer loses the structure and function of its individual members. Our initial reasoning for having the 5' primer binding site of the NTD library and 3' primer binding site of the CTD library complimentary to one another was so that we could bind the resulting aptamers together without ligation, but simply through hydrogen bonding. However, as was discovered, at least one of the primer binding sites of NAp1 is necessary for the functioning of the aptamer.

In order to construct a functional heterobivalent aptamer that is structurally similar to its component parts, we first had to investigate the structure and function of the individual aptamers. To begin, we determined the importance of the primer binding sites of NAp1 and CAp1 to their function using the protein-bound bead binding assay (Figure 3.14). The assay was carried out as before, but either in the presence or absence of a complimentary sequence to the 5' or 3' primer binding site.

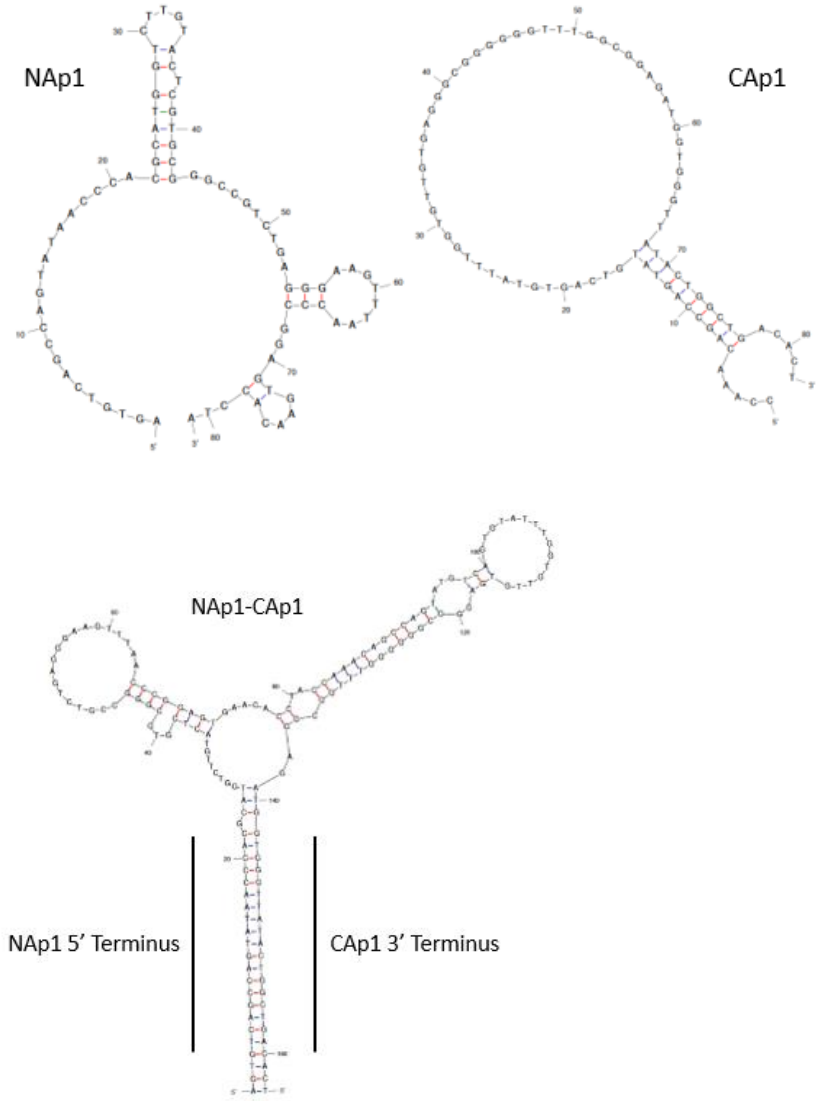

Figure 3.13. Predicted secondary structures of unmodified NAp1, CAp1, and the NAp1-CAp1 heterobivalent aptamer. Structures were predicted by the Mfold web server for nucleic acid folding and hybridization prediction. Folding conditions: 25°C, 100 mM Na^{+}, 2 mM Mg^{2+}.

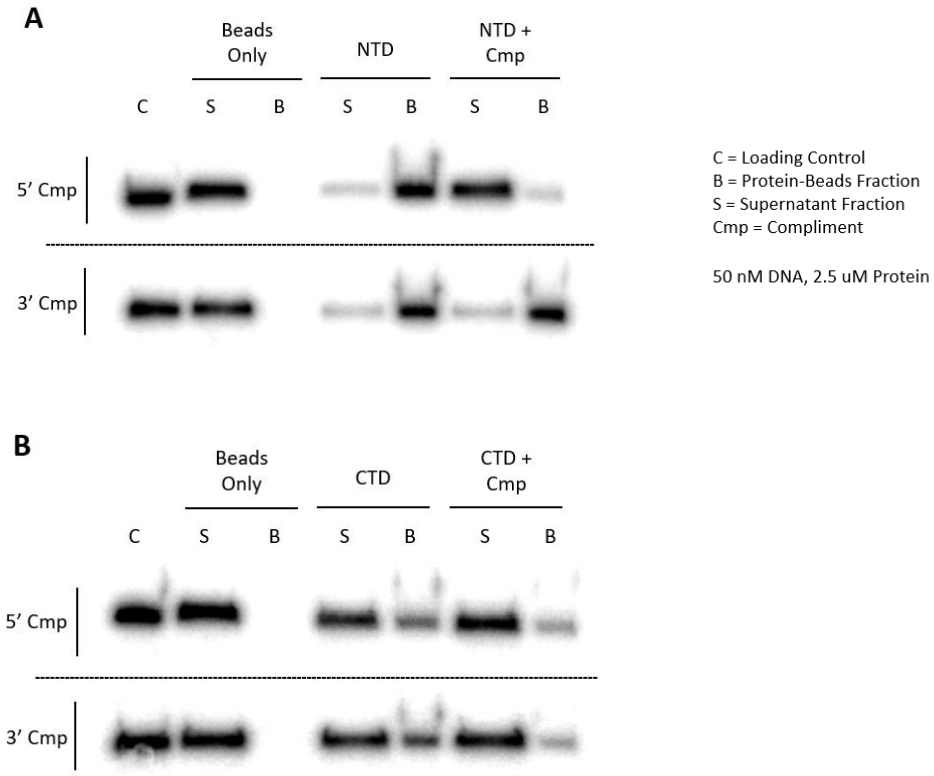

Figure 3.14. Protein-bound bead binding assay for NAp1 and CAp1 in the presence of 5' and 3' complimentary sequences. (A) NAp1 exposed to beads only (negative control), NTD, and NTD + 5' or 3' compliment. (B) CAp1 exposed to beads only, CTD, and CTD + 5' or 3' compliment. Proteins were captured with Ni-NTA magnetic beads and exposed to radiolabeled aptamers for 30 min with shaking. Beads were then separated from the supernatant using a magnet and washed several times with selection buffer. Supernatant and beads were then run on 10% denaturing PAGE gels.

As can be seen, NAp1 loses most of its binding capacity when the 5' site is blocked, but retains its binding capacity when the 3' site is blocked. This suggests that the 3' primer binding site is expendable, and at least a portion of the 5' primer binding site is required. On the other hand, CAp1 seems to lose only modest binding capacity when either the 5' or 3' sites are blocked. In this case, we hypothesize that neither site is

inherently required but rather necessary only to hold the large loop above the stem in place, stabilizing its folding. This is because the large, seemingly unstructured loop, is G rich and believed to form a G-quadruplex structure, which cannot be predicted by the Mfold algorithm. QGRS (Quadruplex forming G-Rich Sequence) Mapper, a software program based on algorithms for recognition and mapping of putative QGRS, identified CAp1 as containing a G-quadruplex, confirming our suspicion.

Because of the information obtained about the importance of the 5' and 3' termini of NAp1 and CAp1, we probed further modifications which would allow us to combine the two aptamers into a coherent heterobivalent construct retaining the individual functions of its component parts. First, we modified CAp1 by removing five nucleotides each from the 5' and 3' ends, which do not contribute to the structure. We then scrambled the sequences involved in the stem so that CAp1 would retain the stem but the 3' terminus would not be complimentary to the 5' terminus of NAp1. We then tested this modification, CAp1-M1, for target binding (Figure 3.15). Figure 3.15 A shows the predicted structure of CAp1-M1, which retains the structural features of CAp1 but has a scrambled stem sequence. Figure 3.15 B shows that CAp1-M1 binds CTD protein just as well as the unmodified CAp1, thus the modifications have retained structure and function.

Given the more transparent nature of the structure of NAp1, we proposed several modifications of the aptamer (Figure 3.16 A). The first modification, NAp1-M1,

truncated the 3’ end by 13 nucleotides so as to leave the predicted small loop structure intact. NAp1-M2 has the same truncation as M1, but also truncates the 5’ end by 10 nucleotides, leaving only half of the 5’ primer binding site intact. NAp1-M3 goes further than M1 by truncating 20 nucleotides from the 3’ end, effectively removing the whole 3’ primer binding site. NAp1-M4 also removes the 3’ primer binding site and truncates the 5’ end by 10 nucleotides, leaving half of the 5’ primer binding site intact. As can be seen in Figure 3.16 B, all proposed modifications resulted in target binding to the same degree as the unmodified NAp1. Thus it is clear, as with our previous tests, that the 3’ end is indeed unnecessary, and that the second half of the 5’ primer site is sufficient to maintain aptamer structure and function. Because all NAp1 modifications performed similarly well, we chose to continue work with the smallest of the lot, NAp1-M4.

Now that we had modified and truncated aptamers, the question remained, would they form a coherent heterobivalent structure that resembled its individual parts. The logical first step, then, would be to input the combined sequences into the Mfold algorithm and analyze the given output (Figure 3.17). According to Mfold, the NAp1-M4-CAp1-M1 construct (N4-C1), folds so as to maintain the structure of the individual aptamers contained within it. Next, we had to construct and test the heterobivalent aptamer for functionality. To do this, we used the double stranded ligation method, as before, except this time we included variable length spacers composed of thymidine repeats (10, 15, and 20), which did not affect the construct’s structure according to Mfold.

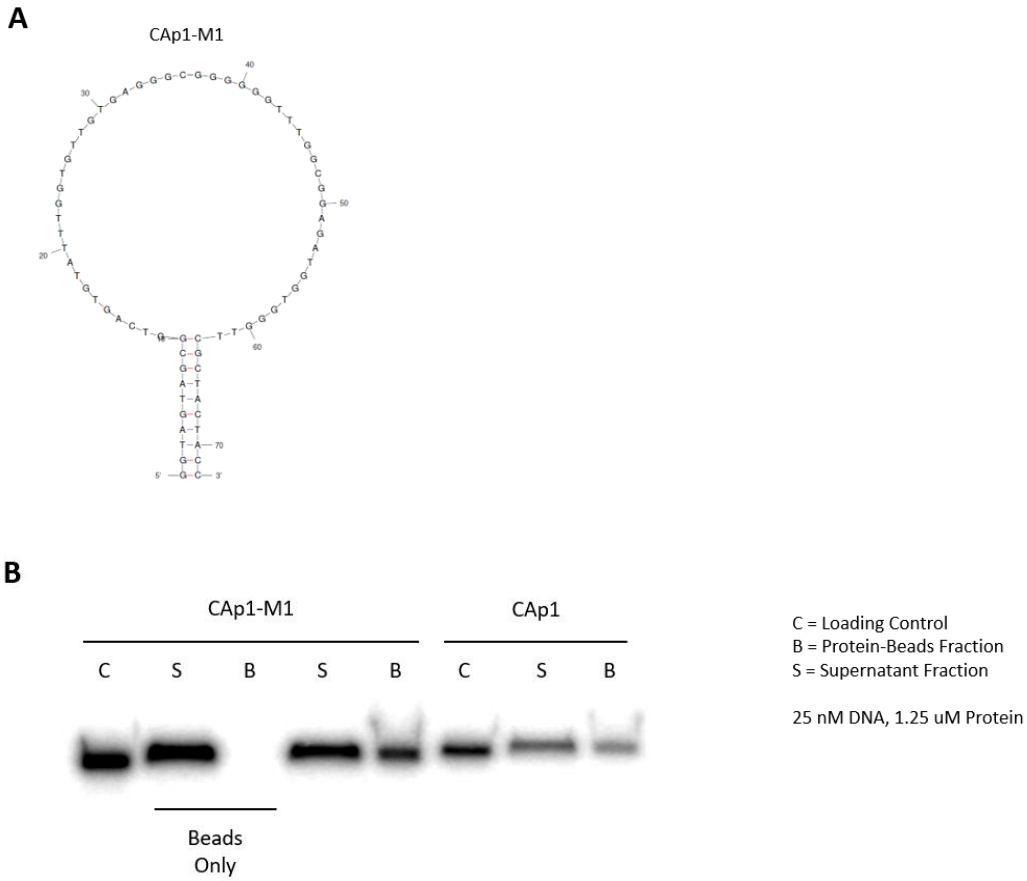

Figure 3.15. Predicted secondary structure of CAp1-M1 and its binding to CTD protein. (A) CAp1-M1 structure prediction by the Mfold web server. Folding conditions: 25°C, 100 mM Na^{+}, 2 mM Mg^{2+}. (B) Protein-bound bead binding assay to asses CAp1-M1 binding to CTD relative to unmodified CAp1. Proteins were captured with Ni-NTA magnetic beads and exposed to radiolabeled aptamers for 30 min with shaking. Beads were then separated from the supernatant using a magnet and washed several times with selection buffer. Supernatant and beads were then run on 10% denaturing PAGE gels.

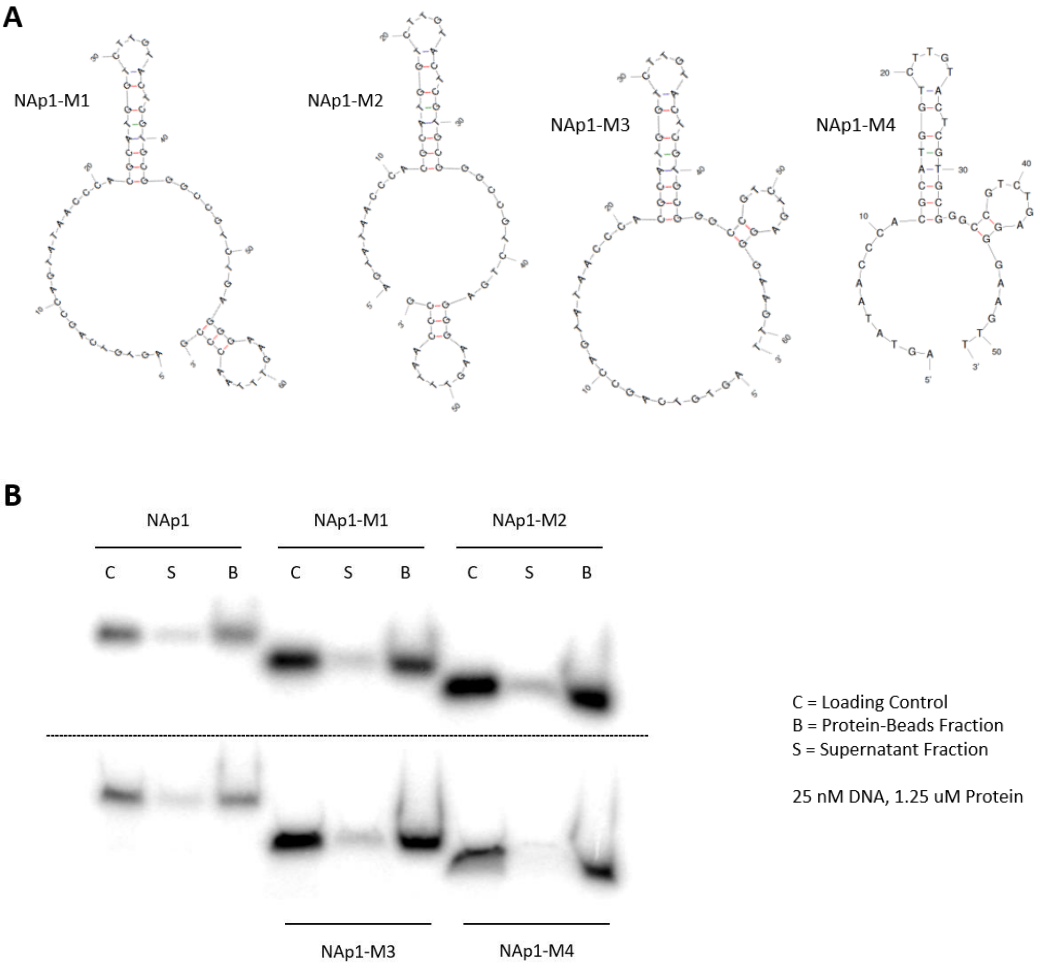

Figure 3.16. Predicted secondary structure of NAp1-M1, -M2, -M3, and -M4 and their binding to NTD protein. (A) NAp1-M1, -M2, -M3, and -M4 structure predictions by the Mfold web server. Folding conditions: 25°C, 100 mM Na^{+}, 2 mM Mg^{2+}. (B) Protein-bound bead binding assay to asses NAp1 modifications binding to NTD relative to unmodified NAp1. Proteins were captured with Ni-NTA magnetic beads and exposed to radiolabeled aptamers for 30 min with shaking. Beads were then separated from the supernatant using a magnet and washed several times with selection buffer. Supernatant and beads were then run on 10% denaturing PAGE gels.

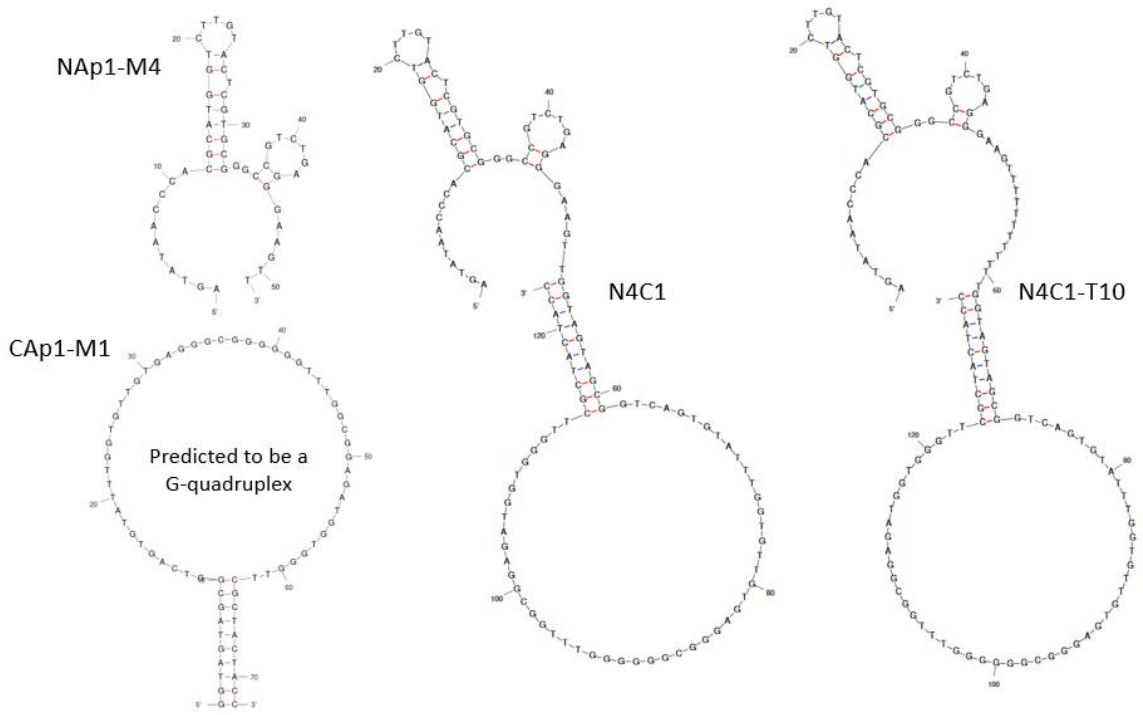

Figure 3.17. Predicted secondary structure of NAp1-M4 and CAp1-M1 combined into a single heterobivalent structure. Structure predictions by the Mfold web server. ApN4-C1 composed of NAp1-M4 and CAp1-M1, ApN4-T10-C1 includes a 10 T spacer between the two.

We included the spacers in order to give the individual aptamers within the heterobivalent structure room to maneuver and find their binding sites on the large globular protein target. If the construct were to too constrained or short, one of the two aptamer components may not be able to reach or bind its target site, essentially defeating the purpose of the dimerization. The variable length constructs were then tested for DnaK binding, relative to the single DnaK aptamer, KAp1, using the electrophoretic mobility shift assay (Figure 3.17).

Although the protein-bound bead binding assay was adequate for determining target binding, we found that the amount of protein captured by the beads was variable and less than expected, hence it was not a very good quantitative test. EMSA on the other hand is a well-recognized, quantitatively reliable, method. As can be seen in Figure 3.18, all of the heterobivalent constructs bound DnaK, and to such a degree that none were detectably left unbound. This was the first indication that the heterobivalent aptamer indeed possessed a higher affinity for the target than the single aptamer. Moving forward, we decided to focus on the smallest spacer variant, N4C1-T10.

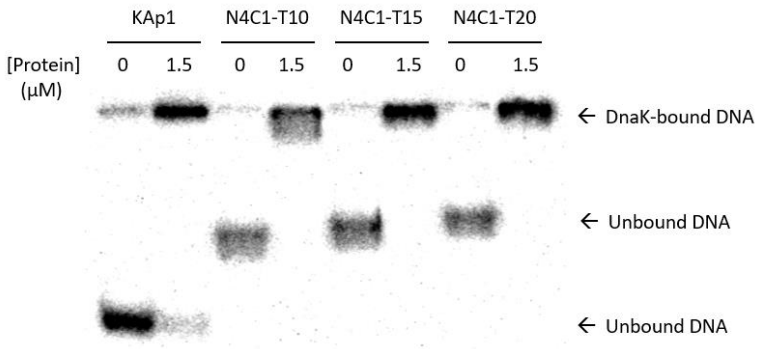

Figure 3.18. EMSA of various N4C1 heterobivalent constructs. Radioactively labeled aptamer DNA (25 nM) were incubated with or without DnaK protein, in selection buffer (50 mM HEPES, pH 8.0, 100 mM NaCl, 10 mM KCl, 2 mM $MgCl_2$, 0.05% Tween 20) for 30 minutes. The samples were run on a 10% native PAGE gel.

3.6 Relative Affinity and Specificity of the Heterobivalent DnaK Aptamer

In order to get a quantitatively reliable measurement of the affinity of the heterobivalent aptamer, N4C1-T10, as well as the individual components, NAp1-M4 and CAp1-M1, and the single DnaK aptamer KAp1, we used EMSA to visualize and quantify a titration curve for each (Figure 3.19), done in triplicate. From these gels, it is immediately clear that N4C1-T10 is in a league of its own, with the range of the curve in the single digits, as opposed to the rest of the single aptamers, which have a range in the triple digits.

Using these gels, bands were quantified with ImageJ, binding fractions were quantified and the data were input into GraphPad Prism. Binding curves were generated using non-linear regression (one site, specific binding) to determine the affinity of each aptamer (Figure 3.20). As can be seen, again, the heterobivalent aptamer is in a completely different class. Based on these curves, the K_d of N4C1-T10 is 1.66 nM, almost 100 fold better than that of the single DnaK aptamer KAp1, which has a K_d of 160 nM. The K_d of the individual components of the construct, NAp1-M4 and CAp1-M1, are 206 nM and 69.5 nM, respectively. It is clear from this example, then, that combining the two aptamers not only increases affinity, but the increase seen is more than just additive, but exponential in nature.

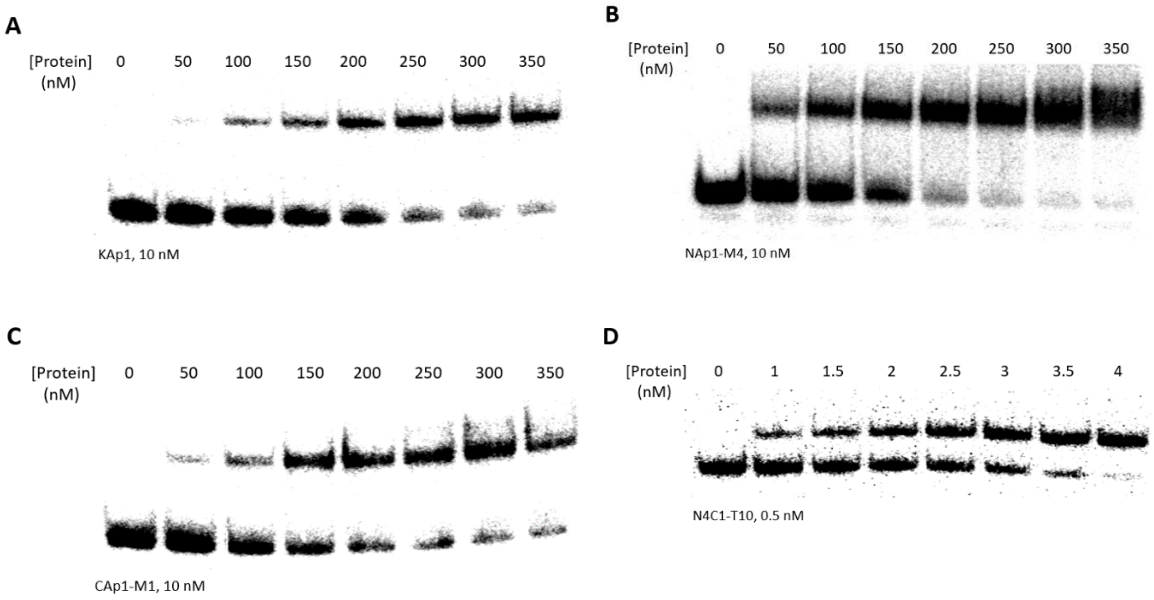

Figure 3.19. EMSA of KAp1, NAp1-M4, CAp1-M1, and N4C1-T10 for DnaK binding, titrated. (A) KAp1 binding DnaK titration, (B) NAp1-M4 binding DnaK titration, (C) CAp1-M1 binding DnaK titration, (D) N4C1-T10 binding DnaK titration. Radioactively labeled aptamer DNA were incubated with increasing concentrations of DnaK protein, in selection buffer (50 mM HEPES, pH 8.0, 100 mM NaCl, 10 mM KCl, 2 mM $MgCl_2$, 0.05% Tween 20) for 30 minutes. The samples were run on a 7% native PAGE gel.

Clearly, the effect of combining two single aptamers targeting different sites on a single target into one heterobivalent structure results in greatly increased affinity for the target; but does this combination also affect target specificity. To answer this question, we began by comparing N4C1-T10 to KAp1 binding of purified *C. difficile* DnaK vs *E. coli* DnaK (Figure 3.21). Because N4C1-T10 has a much greater affinity than KAp1, we used a lower concentration of DnaK for the former (Figure 3.21 A).

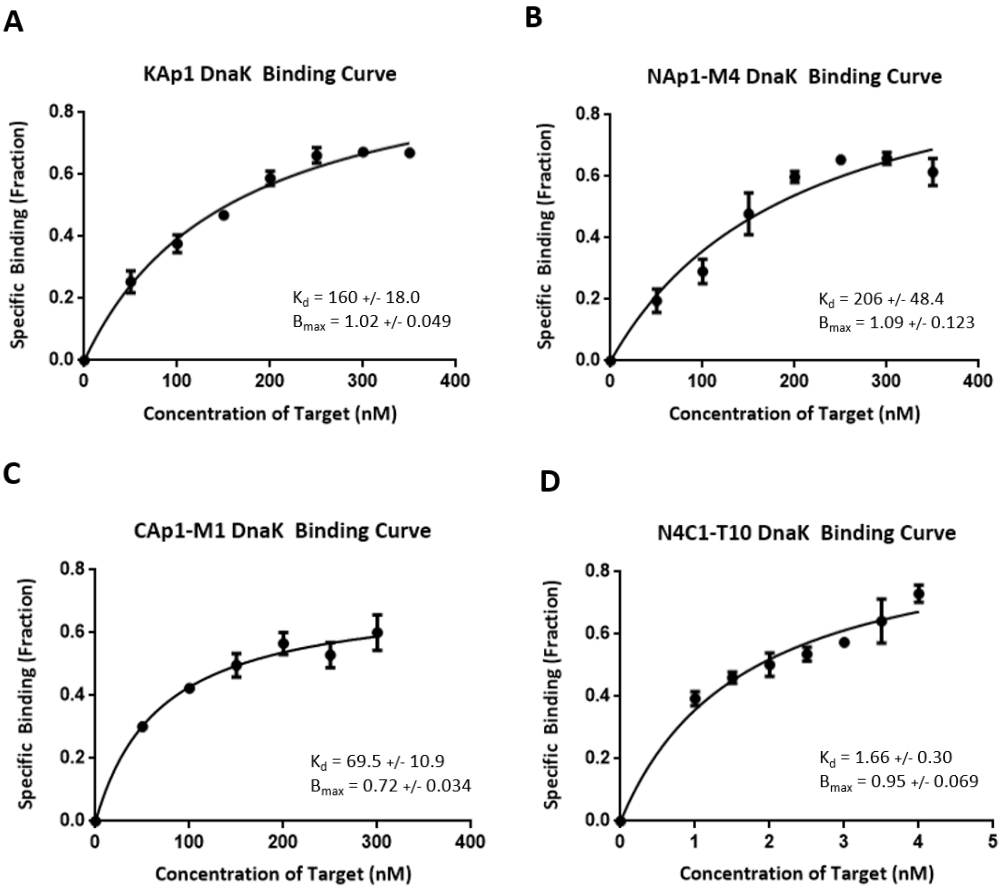

Figure 3.20. Binding curves for KAp1, NAp1-M4, CAp1-M1, and N4C1-T10. (A) Binding curve for KAp1, (B) binding curve for NAp1-M4, (C) binding curve for CAp1-M1, and (D) binding curve for N4C1-T10. Triplicate titrations (of DnaK) were carried out for each aptamer. Samples were run on 7% native PAGE gels and radioactively labeled DNA bands were visualized with a Typhoon Imager. Bands were quantified using ImageJ software and binding fractions were calculated. Data was used to generate binding curves using GraphPad Prism software using non-linear regression for one site specific binding.

As can be seen, neither aptamer had any detectable binding to purified *E. coli* DnaK. However, to be sure of any low level binding, we also carried out a titration for both aptamers against *E. coli* DnaK (Figure 3.21 B, C). Unexpectedly, it turned out that

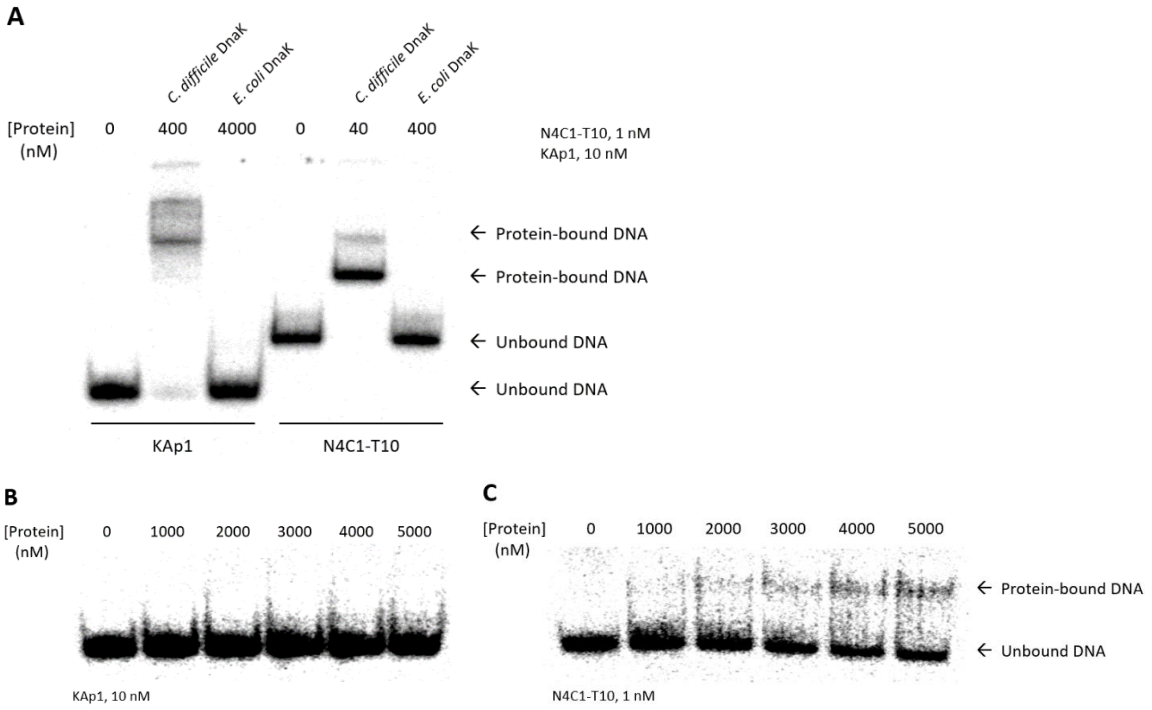

Figure 3.21. KAp1 vs N4C1-T10 binding to purified *C. difficile* and *E. coli* DnaK. (A) Comparative binding of KAp1 vs N4C1-T10 to purified *C. difficile* or *E. coli* DnaK, (B) titration of KAp1 binding to purified *E. coli* DnaK, (C) titration of N4C1-T10 binding to purified *E. coli* DnaK. Radioactively labeled aptamer DNA were incubated with various or increasing concentrations of purified DnaK protein (from *C. difficile* or *E. coli*), in selection buffer (50 mM HEPES, pH 8.0, 100 mM NaCl, 10 mM KCl, 2 mM $MgCl_2$, 0.05% Tween 20) for 30 minutes. The samples were run on a 7% native PAGE gel.

the heterobivalent aptamer had some low level binding to high concentrations of *E. coli* DnaK. When this titration was analyzed and fitted, an approximate K_d of 2637 nM was determined. This is almost a 1600-fold lower affinity than that of N4C1-T10 to *C. difficile* DnaK.

Although N4C1-T10 had minor non-specific binding to high concentrations of purified *E. coli* DnaK, a true measure of an aptamer's effectiveness can only be obtained

using a real-life or biological sample. To approximate such a sample, we grew several strains of bacteria with different levels of DnaK similarity, relative to *C. difficile.* These were *E. coli* (57%), *F. nucleatum* (66%), *B. subtilis* (69%), and *C. difficile* (100%). We then processed the cultures by chemical lysis and spun them down to yield only the soluble fraction, the crude intracellular matrix (CIM). The CIM was then run through a 100 kDa spin column, the filtrate was then collected and run through a 50 kDa spin column and the concentrate was collected. Hence, this fractionated intracellular matrix (FIM), contained all soluble intracellular proteins in the approximate range of 50 to 100 kDa (DnaK is ~70 kDa).

Each culture that was grown was also quantified by cell colony count, in triplicate, in order to approximate the colony forming unit (CFU) equivalent of each FIM used for aptamer binding (Figure 3.22). As can be seen, neither KAp1 or N4C1-T10 had any non-specific binding to bacterial FIM, even with 10 fold more *E. coli* and *B. subtilis* FIM, and both bound to target in *C. difficile* FIM. As expected, N4C1-T10, bound much more (in this case all of the aptamer was bound) to target. This shows that N4C1-T10 not only binds its target in a complex biological sample, but does not show any non-specific binding to homologous proteins.

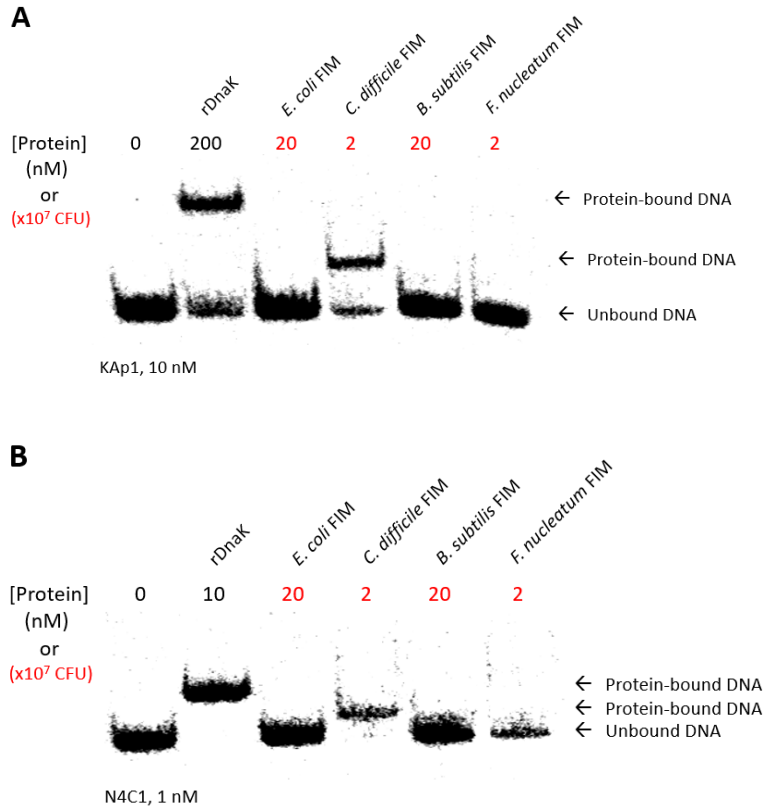

Figure 3.22. KAp1 vs N4C1-T10 biological sample specificity test. (A) KAp1 and (B) N4C1-T10 binding and specificity to purified (recombinant) *C. difficile* DnaK, *E. coli* FIM, *C. difficile* FIM, *B. subtilis* FIM, and *F. nucleatum* FIM. Radioactively labeled aptamer DNA were incubated with purified DnaK protein or various bacterial FIM, given in CFU equivalents, in selection buffer (50 mM HEPES, pH 8.0, 100 mM NaCl, 10 mM KCl, 2 mM $MgCl_2$, 0.05% Tween 20) for 30 minutes. The samples were run on 7% native PAGE gels.

Chapter 4 – Discussion

C. difficile infection (CDI) is a major cause of hospital-acquired infections, morbidity, and mortality in developed nations, which have experienced dramatic increases in infections and mortality. Current detection methods for CDI can lack the desired sensitivity, be time consuming, or expensive. Aptamers can serve as alternative tools, having numerous advantages over conventional methods that use antibodies or gene amplification. However, aptamers isolated by current methods may be susceptible to non-specific binding in complex or biological samples. To get around this problem, we propose a new approach in isolating aptamers: multiple domain selection (MDS). This approach utilizes the power of multiple, independent, selections on separate components of one target in order to obtain uniquely specific aptamers. We hypothesized that combining the aptamers into a single heterobivalent structure would result in increased affinity and specificity for the target, resulting in less non-specific binding, and overall greater utility in any assay.

In the current study, we identified *C. difficile* DnaK as a potential candidate for MDS. DnaK was chosen because it contains two distinct domains that can be targeted for selection. It also appears to be found both within the cell and extracellularly, and is also upregulated during a clinically relevant heat stress, which can be useful for diagnostic testing. In order to determine where DnaK could be split into its respective domains, the linker region had to be determined. The structure of *C. difficile* DnaK had not yet been

determined, but research indicated that the linker region is highly conserved between species and is known in *E. coli*. Using this information, the *E. coli* amino acid sequence was aligned with the *C. difficile* sequence to determine the likely linker region. This information was used to construct clones of full length DnaK as well as its constituent domains (NTD and CTD).

One potential problem that could arise in such a proposed method is that the individual domains of the target, when separated, may not form correctly. Therefore, target selection is a critical decision in the process. Proteins, in general, are great targets for MDS because many contain individual domains, which fold independently as part of the overall structure. However, many proteins are also single domain proteins, which may not be able to retain their native structure when split. Another consideration is very large proteins that have many domains. These complex targets may have inter-domain interactions essential for the proper folding of individual domains, which may not fold properly once otherwise isolated. With these considerations in mind, DnaK is certainly a great target. It has more than one domain but is not large and overly complex. Furthermore, its two domains are separated by a flexible linker, making separation easier.

Although MDS inherently requires more work than a standard single-target selection, there is an advantage in performing multiple selections. With the advent of next generation sequencing, a lot of information can be obtained from *in vitro* selection. For one, a near complete list of the most abundant sequences is obtained once the selection is

complete. This information can be analyzed to determine the characteristics of high and low abundance sequences. Furthermore, sequences can be compared between the individual selections. For example, it is widely believed that many selection targets contain one "aptagenic" site, which outcompetes other possible sites for aptamer binding. This leads to a pool of aptamers that seemingly only bind one site on the target of interest. This is also why it is difficult to obtain single aptamers to distinct sites on one target, without performing MDS.

When comparing the individual selections, it was clear that sequences from the DnaK selection had similarities with those from the NTD selection. However, neither had overt similarity with the CTD selection. It would follow, then, that the NTD contained the "aptagenic" site. Indeed, this was observed when testing the top aptamer sequences from the DnaK selection, which all bound to the NTD protein but not the CTD protein. However, this does not mean that potential CTD binding sequences did not exist within the DnaK selection pool. Although there were no identical sequences between the two, 14 low ranked species in the DnaK selection had characteristics of the CTD core motif. Therefore, it is plausible that sequences specific to distinct sites on one target could be pulled from an individual selection, and warrants further investigation.

It was interesting to observe the relative affinity of the individual aptamers versus the heterobivalent aptamer, once tested. KAp1 (160 nM) and NAp1-M4 (206 nM) were the closest in affinity, which is not unexpected as they both likely target the same site.

CAp1-M1 (70 nM), however, had a better affinity than the NTD aptamer, which seems unexpected because it targets the non-aptagenic site. This could indicate that although non-aptagenic sites on a target are harder to select for, it is not because of relative affinity differences with the aptagenic site. What was exciting, however, was the amplification in affinity once the two aptamers were combined. N4C1-T10 (1.66 nM) has an affinity almost 100 fold better than KAp1, showcasing the power of multi-aptamer unification. This technique, then, has tremendous value for improving aptamer affinity, and therefore any assay utilizing aptamers.

What about specificity? The increase in affinity is certainly a positive outcome, but can a heterobivalent aptamer targeting two sites on one target increase its specificity for that target. This was difficult to establish compared to KAp1, which did not experience non-specific binding to purified *E. coli* DnaK, nor any of the bacterial FIM samples. However, at very high concentrations of purified *E. coli* DnaK, N4C1-T10 did exhibit low level binding. This was an interesting finding, because it meant the relatively very high affinity of the heterobivalent aptamer would allow non-specific binding to a similar target at very high concentration. Realistically speaking, this has little to do with real world specificity, but rather highlights the power of the relative affinity increase. N4C1-T10 showed much stronger binding to *C. difficile* FIM than KAp1, and no non-specific binding was detected with any other bacterial FIM. Therefore, in a clinically relevant biological sample, the affinity increase outweighs any potential non-specific

binding, which would require a similar target to be in very high concentration for even low level binding to be detected.

4.1 Conclusion

The overall goal of this project was to test whether a heterobivalent construct composed of two specific aptamers, targeting distinct sites on one target, would result in greater affinity and specificity for the target of interest. The affinity for the target certainly increased dramatically, allowing the heterobivalent aptamer to bind very low levels of target. Specificity, on the other hand, was not as clear cut. It appears the affinity gained by the heterobivalent construct allowed low-level binding in high concentrations of a similar target. However, this did not translate when using biological samples, which did not show any non-specific binding. MDS is certainly worth consideration for those researchers and clinicians who would like to greatly improve aptamer affinity and overall utility. Further work is warranted to investigate the use of such aptamers in real world applications.

References

Aravind, A., Jeyamohan, P., Nair, R., Veeranarayanan, S., Nagaoka, Y., Yoshida, Y., … Kumar, D. S. (2012). AS1411 aptamer tagged PLGA-lecithin-PEG nanoparticles for tumor cell targeting and drug delivery. *Biotechnology and Bioengineering*, *109*, 2920–31.

Balamurugan, S., Obubuafo, A., Soper, S. A., & Spivak, D. A. (2008). Surface immobilization methods for aptamer diagnostic applications. *Analytical and Bioanalytical Chemistry*, *390*, 1009–21.

Ban, N., Nissen, P., Hansen, J., Moore, P. B., Steitz, T. A., Agrawal, R. K., … Carson, M. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science*, *289*, 905–20.

Barbut, F., Delmée, M., Brazier, J. S., Petit, J. C., Poxton, I. R., Rupnik, M., … Tvede, M. (2003). A European survey of diagnostic methods and testing protocols for Clostridium difficile. *Clinical Microbiology and Infection : The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases*, *9*, 989–96.

Bartlett, J. G., & Gerding, D. N. (2008). Clinical recognition and diagnosis of Clostridium difficile infection. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *46 Suppl 1*, S12-8.

Bertelsen, E. B., Chang, L., Gestwicki, J. E., & Zuiderweg, E. R. P. (2009). Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 8471–6.

Bevilacqua, P. C., & Yajima, R. (2006, October). Nucleobase catalysis in ribozyme mechanism. *Current Opinion in Chemical Biology*.

Boetzkes, A., Felkel, K. W., Zeiser, J., Jochim, N., Just, I., & Pich, A. (2012). Secretome analysis of Clostridium difficile strains. *Archives of Microbiology*, *194*, 675–87.

Bolton, D. C., McKinley, M. P., & Prusiner, S. B. (1982). Identification of a protein that purifies with the scrapie prion. *Science*, *218*, 1309–11.

Brauner-Osborne, H., Wellendorph, P., & Jensen, A. (2007). Structure, Pharmacology and Therapeutic Prospects of Family C G-Protein Coupled Receptors. *Current Drug Targets*, *8*, 169–184.

Breaker, R. R., & Joyce, G. F. (1995). A DNA enzyme with Mg2+-dependent RNA phosphoesterase activity. *Chemistry and Biology*, *2*, 655–660.

Bruno, J. G., Carrillo, M. P., & Phillips, T. (2008). In vitro antibacterial effects of antilipopolysaccharide DNA aptamer-C1qrs complexes. *Folia Microbiologica*, *53*, 295–302.

Bunka, D. H. J., Platonova, O., & Stockley, P. G. (2010). Development of aptamer therapeutics. *Current Opinion in Pharmacology*, *10*, 557–62.

Burckhardt, F., Friedrich, A., Beier, D., & Eckmanns, T. (2008). Clostridium difficile surveillance trends, Saxony, Germany. *Emerging Infectious Diseases*, *14*, 691–2.

Burnham, C.-A. D., & Carroll, K. C. (2013). Diagnosis of Clostridium difficile infection: an ongoing conundrum for clinicians and for clinical laboratories. *Clinical Microbiology Reviews*, *26*, 604–30.

Carroll, K. C. (2011). Tests for the diagnosis of Clostridium difficile infection: the next generation. *Anaerobe*, *17*, 170–4.

Chan, M. Y., Rusconi, C. P., Alexander, J. H., Tonkens, R. M., Harrington, R. A., & Becker, R. C. (2008). A randomized, repeat-dose, pharmacodynamic and safety study of an antidote-controlled factor IXa inhibitor. *Journal of Thrombosis and Haemostasis : JTH*, *6*, 789–96.

Chapple, K. E., Bartel, D. P., & Unrau, P. J. (2003). Combinatorial minimization and secondary structure determination of a nucleotide synthase ribozyme. *RNA*, *9*, 1208–1220.

Cheng, C., Chen, Y. H., Lennox, K. A., Behlke, M. A., & Davidson, B. L. (2013). In vivo SELEX for Identification of Brain-penetrating Aptamers. *Molecular Therapy. Nucleic Acids*, *2*, e67.

Cho, M., Xiao, Y., Nie, J., Stewart, R., Csordas, A. T., Oh, S. S., … Soh, H. T. (2010). Quantitative selection of DNA aptamers through microfluidic selection and high-throughput sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 15373–8.

Cohen, S. H., Gerding, D. N., Johnson, S., Kelly, C. P., Loo, V. G., McDonald, L. C., … Wilcox, M. H. (2010). Clinical practice guidelines for Clostridium difficile infection in adults: 2010 update by the society for healthcare epidemiology of America (SHEA) and the infectious diseases society of America (IDSA). *Infection Control and Hospital Epidemiology : The Official Journal of the Society of Hospital*

Epidemiologists of America, *31*, 431–55.

Crick, F. H. C. (1970). Central Dogma of Molecular Biology. *Nature*.

de Jong, E., de Jong, A. S., Bartels, C. J. M., van der Rijt-van den Biggelaar, C., Melchers, W. J. G., & Sturm, P. D. J. (2012). Clinical and laboratory evaluation of a real-time PCR for Clostridium difficile toxin A and B genes. *European Journal of Clinical Microbiology & Infectious Diseases : Official Publication of the European Society of Clinical Microbiology*, *31*, 2219–25.

Delmée, M., Van Broeck, J., Simon, A., Janssens, M., & Avesani, V. (2005). Laboratory diagnosis of Clostridium difficile-associated diarrhoea: a plea for culture. *Journal of Medical Microbiology*, *54*, 187–91.

Dey, A. K., Griffiths, C., Lea, S. M., & James, W. (2005). Structural characterization of an anti-gp120 RNA aptamer that neutralizes R5 strains of HIV-1. *RNA (New York, N.Y.)*, *11*, 873–84.

Dey, A. K., Khati, M., Tang, M., Wyatt, R., Lea, S. M., & James, W. (2005). An aptamer that neutralizes R5 strains of human immunodeficiency virus type 1 blocks gp120-CCR5 interaction. *Journal of Virology*, *79*, 13806–10.

Di Bella, S., Musso, M., Cataldo, M. A., Meledandri, M., Bordi, E., Capozzi, D., … Petrosillo, N. (2013). Clostridium difficile infection in Italian urban hospitals: data from 2006 through 2011. *BMC Infectious Diseases*, *13*, 146.

Dorsam, R. T., & Gutkind, J. S. (2007). G-protein-coupled receptors and cancer. *Nature Reviews. Cancer*, *7*, 79–94.

Duclair, S., Gautam, A., Ellington, A., & Prasad, V. R. (2015). High-affinity RNA Aptamers Against the HIV-1 Protease Inhibit Both In Vitro Protease Activity and Late Events of Viral Replication. *Molecular Therapy. Nucleic Acids*, *4*, e228.

Eastwood, K., Else, P., Charlett, A., & Wilcox, M. (2009). Comparison of nine commercially available Clostridium difficile toxin detection assays, a real-time PCR assay for C. difficile tcdB, and a glutamate dehydrogenase detection assay to cytotoxin testing and cytotoxigenic culture methods. *Journal of Clinical Microbiology*, *47*, 3211–7.

Eggertson, L., & Sibbald, B. (2004). Hospitals battling outbreaks of C. difficile. *CMAJ : Canadian Medical Association Journal = Journal de l'Association Medicale Canadienne*, *171*, 19–21.

Ekland, E. H., Szostak, J. W., & Bartel, D. P. (1995). Structurally complex and highly active RNA ligases derived from random RNA sequences. *Science*, *269*, 364–370.

Ellington, A. D., & Szostak, J. W. (1990). In vitro selection of RNA molecules that bind specific ligands. *Nature*, *346*, 818–822.

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

Elliott, B., Squire, M. M., Thean, S., Chang, B. J., Brazier, J. S., Rupnik, M., & Riley, T. V. (2011). New types of toxin A-negative, toxin B-positive strains among clinical isolates of Clostridium difficile in Australia. *Journal of Medical Microbiology*, *60*, 1108–11.

Ferré-D'Amaré, A. R., Zhou, K., & Doudna, J. A. (1998). Crystal structure of a hepatitis delta virus ribozyme. *Nature*, *395*, 567–574.

Furuya-Kanamori, L., Robson, J., Soares Magalhães, R. J., Yakob, L., McKenzie, S. J., Paterson, D. L., … Clements, A. C. A. (2014). A population-based spatio-temporal analysis of Clostridium difficile infection in Queensland, Australia over a 10-year period. *The Journal of Infection*, *69*, 447–55.

Gragoudas, E. S., Adamis, A. P., Cunningham, E. T., Feinsod, M., & Guyer, D. R. (2004). Pegaptanib for neovascular age-related macular degeneration. *The New England Journal of Medicine*, *351*, 2805–16.

Gravel, D., Miller, M., Simor, A., Taylor, G., Gardam, M., McGeer, A., … Mulvey, M. (2009). Health care-associated Clostridium difficile infection in adults admitted to acute care hospitals in Canada: a Canadian Nosocomial Infection Surveillance Program Study. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *48*, 568–76.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., & Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, *35*, 849–857.

Gumerlock, P. H., Tang, Y. J., Meyers, F. J., & Silva, J. (1991). Use of the polymerase chain reaction for the specific and direct detection of Clostridium difficile in human feces. *Reviews of Infectious Diseases*, *13*, 1053–60.

Gupta, A., & Khanna, S. (2014). Community-acquired Clostridium difficile infection: an increasing public health threat. *Infection and Drug Resistance*, *7*, 63–72.

Guthrie, J. W., Hamula, C. L. A., Zhang, H., & Le, X. C. (2006). Assays for cytokines

using aptamers. *Methods (San Diego, Calif.)*, *38*, 324–30.

Hall, I., & O'Toole, E. (1935). Intestinal flora in new-born infants: With a description of a new pathogenic anaerobe, Bacillus difficilis. *American Journal of Diseases of Children*, *49*, 390–402.

Hall, K. B., & Kranz, J. K. (1999). Nitrocellulose filter binding for determination of dissociation constants. *Methods in Molecular Biology (Clifton, N.J.)*, *118*, 105–14.

Hasegawa, H., Taira, K., Sode, K., & Ikebukuro, K. (2008). Improvement of Aptamer Affinity by Dimerization. *Sensors*, *8*, 1090–1098.

He, M., Miyajima, F., Roberts, P., Ellison, L., Pickard, D. J., Martin, M. J., … Lawley, T. D. (2013). Emergence and global spread of epidemic healthcare-associated Clostridium difficile. *Nature Genetics*, *45*, 109–13.

Held, D. M., Kissel, J. D., Patterson, J. T., Nickens, D. G., & Burke, D. H. (2006). HIV-1 inactivation by nucleic acid aptamers. *Frontiers in Bioscience : A Journal and Virtual Library*, *11*, 89–112.

Hellman, L. M., & Fried, M. G. (2007). Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nature Protocols*, *2*, 1849–61.

Hermann, T., & Patel, D. J. (2000). Adaptive recognition by nucleic acid aptamers. *Science*, *287*, 820–825.

Hink, T., Burnham, C.-A. D., & Dubberke, E. R. (2013). A systematic evaluation of methods to optimize culture-based recovery of Clostridium difficile from stool specimens. *Anaerobe*, *19*, 39–43.

Humphries, R. M., Uslan, D. Z., & Rubin, Z. (2013). Performance of Clostridium difficile toxin enzyme immunoassay and nucleic acid amplification tests stratified by patient disease severity. *Journal of Clinical Microbiology*, *51*, 869–73.

Jaenisch, R., & Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature Genetics*, *33 Suppl*, 245–254.

Jaffe, G. J., Eliott, D., Wells, J. A., Prenner, J. L., Papp, A., & Patel, S. (2016). A Phase 1 Study of Intravitreous E10030 in Combination with Ranibizumab in Neovascular Age-Related Macular Degeneration. *Ophthalmology*, *123*, 78–85.

Jayasena, S. D. (1999). Aptamers: an emerging class of molecules that rival antibodies in

diagnostics. *Clinical Chemistry*, *45*, 1628–50.

Jilma, B., Paulinska, P., Jilma-Stohlawetz, P., Gilbert, J. C., Hutabarat, R., & Knöbl, P. (2010). A randomised pilot trial of the anti-von Willebrand factor aptamer ARC1779 in patients with type 2b von Willebrand disease. *Thrombosis and Haemostasis*, *104*, 563–70.

Johnson, S. W., Kanatani, M., Humphries, R. M., & Uslan, D. Z. (2013). Clinical impact of switching conventional enzyme immunoassay with nucleic acid amplification test for suspected Clostridium difficile-associated diarrhea. *American Journal of Infection Control*, *41*, 373–5.

Kane, P. M., Yamashiro, C. T., Wolczyk, D. F., Neff, N., Goebl, M., & Stevens, T. H. (1990). Protein splicing converts the yeast TFP1 gene product to the 69-kD subunit of the vacuolar H(+)-adenosine triphosphatase. *Science*, *250*, 651–7.

Kato, N., Ou, C. Y., Kato, H., Bartley, S. L., Brown, V. K., Dowell, V. R., & Ueno, K. (1991). Identification of toxigenic Clostridium difficile by the polymerase chain reaction. *Journal of Clinical Microbiology*, *29*, 33–7.

Kavruk, M., Celikbicak, O., Ozalp, V. C., Borsa, B. A., Hernandez, F. J., Bayramoglu, G., … Arica, M. Y. (2015). Antibiotic loaded nanocapsules functionalized with aptamer gates for targeted destruction of pathogens. *Chemical Communications (Cambridge, England)*, *51*, 8492–5.

Keefe, A. D., & Cload, S. T. (2008). SELEX with modified nucleotides. *Current Opinion in Chemical Biology*, *12*, 448–56.

Keefe, A. D., Pai, S., & Ellington, A. (2010). Aptamers as therapeutics. *Nature Reviews. Drug Discovery*, *9*, 537–50.

Kelly, C. P., & LaMont, J. T. (1998). Clostridium difficile infection. *Annual Review of Medicine*, *49*, 375–390.

Khanna, S., Baddour, L. M., Huskins, W. C., Kammer, P. P., Faubion, W. A., Zinsmeister, A. R., … Pardi, D. S. (2013). The epidemiology of Clostridium difficile infection in children: a population-based study. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *56*, 1401–6.

Khanna, S., Pardi, D. S., Aronson, S. L., Kammer, P. P., Orenstein, R., St Sauver, J. L., … Zinsmeister, A. R. (2012). The epidemiology of community-acquired Clostridium difficile infection: a population-based study. *The American Journal of Gastroenterology*, *107*, 89–95.

Kim, H., Riley, T. V, Kim, M., Kim, C. K., Yong, D., Lee, K., … Park, J.-W. (2008). Increasing prevalence of toxin A-negative, toxin B-positive isolates of Clostridium difficile in Korea: impact on laboratory diagnosis. *Journal of Clinical Microbiology*, *46*, 1116–7.

Kim, J., Pai, H., Seo, M.-R., & Kang, J. O. (2012). Clinical and microbiologic characteristics of tcdA-negative variant Clostridium difficile infections. *BMC Infectious Diseases*, *12*, 109.

Kim, J., Smathers, S. A., Prasad, P., Leckerman, K. H., Coffin, S., & Zaoutis, T. (2008). Epidemiological Features of Clostridium difficile-Associated Disease Among Inpatients at Children's Hospitals in the United States, 2001–2006. *Pediatrics*, *122*, 1266–1270.

Kim, Y. S., Chung, J., Song, M. Y., Jurng, J., & Kim, B. C. (2014). Aptamer cocktails: enhancement of sensing signals compared to single use of aptamers for detection of bacteria. *Biosensors & Bioelectronics*, *54*, 195–8.

Klein, D. J., & Ferré-D'Amaré, A. R. (2006). Structural basis of glmS ribozyme activation by glucosamine-6-phosphate. *Science*, *313*, 1752–1756.

Koo, H. L., Van, J. N., Zhao, M., Ye, X., Revell, P. A., Jiang, Z.-D., … DuPont, H. L. (2014). Real-time polymerase chain reaction detection of asymptomatic Clostridium difficile colonization and rising C. difficile-associated disease rates. *Infection Control and Hospital Epidemiology*, *35*, 667–73.

Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., & Cech, T. R. (1982). Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. *Cell*, *31*, 147–157.

Lafontaine, D. A., Norman, D. G., & Lilley, D. M. J. (2001). Structure, folding and activity of the VS ribozyme: Importance of the 2-3-6 helical junction. *EMBO Journal*, *20*, 1415–1424.

Lambert, P. J., Dyck, M., Thompson, L. H., & Hammond, G. W. (2009). Population-based surveillance of Clostridium difficile infection in Manitoba, Canada, by using interim surveillance definitions. *Infection Control and Hospital Epidemiology*, *30*, 945–51.

Lance George, W., Goldstein, E. C., Sutter, V., Ludwig, S., & Finegold, S. (1978). AETIOLOGY OF ANTIMICROBIAL-AGENT-ASSOCIATED COLITIS. *The Lancet*, *311*, 802–803.

Langer, T., Lu, C., Echols, H., Flanagan, J., Hayer, M. K., & Hartl, F. U. (1992). Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. *Nature*, *356*, 683–9.

Lanis, J. M., Heinlen, L. D., James, J. A., & Ballard, J. D. (2013). Clostridium difficile 027/BI/NAP1 encodes a hypertoxic and antigenically variable form of TcdB. *PLoS Pathogens*, *9*, e1003523.

Larson, H. E., Price, A. B., Honour, P., & Borriello, S. P. (1978). CLOSTRIDIUM DIFFICILE AND THE AETIOLOGY OF PSEUDOMEMBRANOUS COLITIS. *The Lancet*, *311*, 1063–1066.

Lau, M. W. L., Cadieux, K. E. C., & Unrau, P. J. (2004). Isolation of fast purine nucleotide synthase ribozymes. *Journal of the American Chemical Society*, *126*, 15686–93.

Lee, C. H., Lee, Y. J., Kim, J. H., Lim, J. H., Kim, J.-H., Han, W., … Lee, S.-W. (2013). Inhibition of hepatitis C virus (HCV) replication by specific RNA aptamers against HCV NS5B RNA replicase. *Journal of Virology*, *87*, 7064–74.

Lessa, F. C., Mu, Y., Bamberg, W. M., Beldavs, Z. G., Dumyati, G. K., Dunn, J. R., … McDonald, L. C. (2015). Burden of Clostridium difficile Infection in the United States. *The New England Journal of Medicine*, *372*, 825–834.

Levy, M., Cater, S. F., & Ellington, A. D. (2005). Quantum-dot aptamer beacons for the detection of proteins. *Chembiochem : A European Journal of Chemical Biology*, *6*, 2163–6.

Li, Y., Liu, Y., & Breaker, R. R. (2000). Capping DNA with DNA. *Biochemistry*, *39*, 3106–3114.

Liu, J., Mazumdar, D., & Lu, Y. (2006). A simple and sensitive “dipstick” test in serum based on lateral flow separation of aptamer-linked nanostructures. *Angewandte Chemie (International Ed. in English)*, *45*, 7955–9.

Lodish, H. F. (1981, July). Post-translational modification of proteins. *Enzyme and Microbial Technology*.

Loo, V. G., Poirier, L., Miller, M. A., Oughton, M., Libman, M. D., Michaud, S., … Dascal, A. (2005). A predominantly clonal multi-institutional outbreak of Clostridium difficile-associated diarrhea with high morbidity and mortality. *The New England Journal of Medicine*, *353*, 2442–9.

Lorsch, J. R., & Szostak, J. W. (1994). In vitro evolution of new ribozymes with polynucleotide kinase activity. *Nature*, *371*, 31–6.

Luna, R. A., Boyanton, B. L., Mehta, S., Courtney, E. M., Webb, C. R., Revell, P. A., & Versalovic, J. (2011). Rapid stool-based diagnosis of Clostridium difficile infection by real-time PCR in a children's hospital. *Journal of Clinical Microbiology*, *49*, 851–7.

Lyytikäinen, O., Turunen, H., Sund, R., Rasinperä, M., Könönen, E., Ruutu, P., & Keskimäki, I. (2009). Hospitalizations and deaths associated with Clostridium difficile infection, Finland, 1996-2004. *Emerging Infectious Diseases*, *15*, 761–5.

Maasch, C., Buchner, K., Eulberg, D., Vonhoff, S., & Klussmann, S. (2008). Physicochemical stability of NOX-E36, a 40mer L-RNA (Spiegelmer) for therapeutic applications. *Nucleic Acids Symposium Series (2004)*, *52*, 61–2.

Mallikaratchy, P. R., Ruggiero, A., Gardner, J. R., Kuryavyi, V., Maguire, W. F., Heaney, M. L., … Scheinberg, D. A. (2011). A multivalent DNA aptamer specific for the B-cell receptor on human lymphoma and leukemia. *Nucleic Acids Research*, *39*, 2458–69.

Mayer, M. P. (2013). Hsp70 chaperone dynamics and molecular mechanism. *Trends in Biochemical Sciences*, *38*, 507–14.

McDonald, L. C., Killgore, G. E., Thompson, A., Owens, R. C., Kazakova, S. V, Sambol, S. P., … Gerding, D. N. (2005). An epidemic, toxin gene-variant strain of Clostridium difficile. *The New England Journal of Medicine*, *353*, 2433–41.

McDonald, L. C., Owings, M., & Jernigan, D. B. (2006). Clostridium difficile infection in patients discharged from US short-stay hospitals, 1996-2003. *Emerging Infectious Diseases*, *12*, 409–15.

Michael Liss, Birgit Petersen, Hans Wolf, * and, & Prohaska, E. (2002). An Aptamer-Based Quartz Crystal Protein Biosensor.

Murphy, M. B. (2003). An improved method for the in vitro evolution of aptamers and applications in protein detection and purification. *Nucleic Acids Research*, *31*, 110e–110.

Ng, E. W. M., Shima, D. T., Calias, P., Cunningham, E. T., Guyer, D. R., & Adamis, A. P. (2006). Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nature Reviews. Drug Discovery*, *5*, 123–32.

Nguyen, D. H., DeFina, S. C., Fink, W. H., & Dieckmann, T. (2002). Binding to an RNA aptamer changes the charge distribution and conformation of malachite green. *Journal of the American Chemical Society*, *124*, 15081–15084.

NICKENS, D. G. (2003). Inhibition of HIV-1 reverse transcriptase by RNA aptamers in Escherichia coli. *RNA*, *9*, 1029–1033.

Nieuwlandt, D., West, M., Cheng, X., Kirshenheuter, G., & Eaton, B. E. (2003). The first example of an RNA urea synthase: Selection through the enzyme active site of human neutrophile elastase. *ChemBioChem*, *4*, 651–654.

Ninichuk, V., Clauss, S., Kulkarni, O., Schmid, H., Segerer, S., Radomska, E., … Anders, H.-J. (2008). Late onset of Ccl2 blockade with the Spiegelmer mNOX-E36-3'PEG prevents glomerulosclerosis and improves glomerular filtration rate in db/db mice. *The American Journal of Pathology*, *172*, 628–37.

Nissen, P., Hansen, J., Ban, N., Moore, P. B., & Steitz, T. A. (2000). The structural basis of ribosome activity in peptide bond synthesis. *Science*, *289*, 920–30.

Nitsche, A., Kurth, A., Dunkhorst, A., Pänke, O., Sielaff, H., Junge, W., … Kage, A. (2007). One-step selection of Vaccinia virus-binding DNA aptamers by MonoLEX. *BMC Biotechnology*, *7*, 48.

Nonaka, Y., Sode, K., & Ikebukuro, K. (2010). Screening and improvement of an anti-VEGF DNA aptamer. *Molecules (Basel, Switzerland)*, *15*, 215–25.

O'Brien, J. A., Lahue, B. J., Caro, J. J., & Davidson, D. M. (2007). The emerging infectious challenge of clostridium difficile-associated disease in Massachusetts hospitals: clinical and economic consequences. *Infection Control and Hospital Epidemiology*, *28*, 1219–27.

Oh, S. S., Qian, J., Lou, X., Zhang, Y., Xiao, Y., & Soh, H. T. (2009). Generation of highly specific aptamers via micromagnetic selection. *Analytical Chemistry*, *81*, 5490–5.

Ozer, A., Pagano, J. M., & Lis, J. T. (2014). New Technologies Provide Quantum Changes in the Scale, Speed, and Success of SELEX Methods and Aptamer Characterization. *Molecular Therapy. Nucleic Acids*, *3*, e183.

Paige, J. S., Wu, K. Y., & Jaffrey, S. R. (2011). RNA mimics of green fluorescent protein. *Science (New York, N.Y.)*, *333*, 642–6.

Pancholi, P., Kelly, C., Raczkowski, M., & Balada-Llasat, J. M. (2012). Detection of

toxigenic Clostridium difficile: comparison of the cell culture neutralization, Xpert C. difficile, Xpert C. difficile/Epi, and Illumigene C. difficile assays. *Journal of Clinical Microbiology*, *50*, 1331–5.

Peery, A. F., Dellon, E. S., Lund, J., Crockett, S. D., McGowan, C. E., Bulsiewicz, W. J., … Shaheen, N. J. (2012). Burden of gastrointestinal disease in the United States: 2012 update. *Gastroenterology*, *143*, 1179-87–3.

Pépin, J., Valiquette, L., Alary, M.-E., Villemure, P., Pelletier, A., Forget, K., … Chouinard, D. (2004). Clostridium difficile-associated diarrhea in a region of Quebec from 1991 to 2003: a changing pattern of disease severity. *CMAJ : Canadian Medical Association Journal = Journal de l'Association Medicale Canadienne*, *171*, 466–72.

Pépin, J., Valiquette, L., & Cossette, B. (2005). Mortality attributable to nosocomial Clostridium difficile-associated disease during an epidemic caused by a hypervirulent strain in Quebec. *CMAJ : Canadian Medical Association Journal = Journal de l'Association Medicale Canadienne*, *173*, 1037–42.

Perler, F. B. (2002). InBase: the Intein Database. *Nucleic Acids Research*, *30*, 383–384.

Peterson, L. R., Manson, R. U., Paule, S. M., Hacek, D. M., Robicsek, A., Thomson, R. B., & Kaul, K. L. (2007). Detection of toxigenic Clostridium difficile in stool samples by real-time polymerase chain reaction for the diagnosis of C. difficile-associated diarrhea. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *45*, 1152–60.

Peterson, L. R., Mehta, M. S., Patel, P. A., Hacek, D. M., Harazin, M., Nagwekar, P. P., … Robicsek, A. (2011). Laboratory testing for Clostridium difficile infection: light at the end of the tunnel. *American Journal of Clinical Pathology*, *136*, 372–80.

Petrella, L. A., Sambol, S. P., Cheknis, A., Nagaro, K., Kean, Y., Sears, P. S., … Gerding, D. N. (2012). Decreased cure and increased recurrence rates for Clostridium difficile infection caused by the epidemic C. difficile BI strain. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *55*, 351–7.

Planche, T. D., Davies, K. A., Coen, P. G., Finney, J. M., Monahan, I. M., Morris, K. A., … Wilcox, M. H. (2013). Differences in outcome according to Clostridium difficile testing method: a prospective multicentre diagnostic validation study of C difficile infection. *The Lancet. Infectious Diseases*, *13*, 936–45.

Planche, T., & Wilcox, M. (2011). Reference assays for Clostridium difficile infection: one or two gold standards? *Journal of Clinical Pathology*, *64*, 1–5.

Pley, H. W., Flaherty, K. M., & McKay, D. B. (1994, November 3). Three-dimensional structure of a hammerhead ribozyme. *Nature*. Nature Publishing Group.

Poniková, S., Tlučková, K., Antalík, M., Víglaský, V., & Hianik, T. (2011). The circular dichroism and differential scanning calorimetry study of the properties of DNA aptamer dimers. *Biophysical Chemistry*, *155*, 29–35.

Prusiner, S. B. (1982). Novel proteinaceous infectious particles cause scrapie. *Science*, *216*, 136–44.

Purtha, W. E., Coppins, R. L., Smalley, M. K., & Silverman, S. K. (2005). General Deoxyribozyme-Catalyzed Synthesis of Native 3‘−5‘ RNA Linkages. *Journal of the American Chemical Society*, *127*, 13124–13125.

Rideout Ill, W. M., Eggan, K., & Jaenisch, R. (2001). the Genome Nuclear Cloning and Epigenetic Reprogramming of the Genome. *Science*, *293*, 1093–1098.

Robertson, D. L., & Joyce, G. F. (1990, March 29). Selection in vitro of an RNA enzyme that specifically cleaves single-stranded DNA. *Nature*.

Romig, T. S., Bell, C., & Drolet, D. W. (1999). Aptamer affinity chromatography: *Journal of Chromatography B: Biomedical Sciences and Applications*, *731*, 275–284.

Rosenbaum, D. M., Rasmussen, S. G. F., & Kobilka, B. K. (2009). The structure and function of G-protein-coupled receptors. *Nature*, *459*, 356–63.

Rouphael, N. G., O’Donnell, J. A., Bhatnagar, J., Lewis, F., Polgreen, P. M., Beekmann, S., … McDonald, L. C. (2008). Clostridium difficile-associated diarrhea: an emerging threat to pregnant women. *American Journal of Obstetrics and Gynecology*, *198*, 635.e1-6.

Rupert, P. B., & Ferré-D’Amaré, A. R. (2001). Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. *Nature*, *410*, 780–786.

Rupnik, M., Wilcox, M. H., & Gerding, D. N. (2009). Clostridium difficile infection: new developments in epidemiology and pathogenesis. *Nat Rev Micro*, *7*, 526–536.

Santoro, S. W., & Joyce, G. F. (1997). A general purpose RNA-cleaving DNA enzyme. *Proceedings of the National Academy of Sciences of the United States of America*,

94, 4262–4266.

Sassanfar, M., & Szostak, J. W. (1993). An RNA motif that binds ATP. *Nature*, *364*, 550–3.

Schütze, T., Wilhelm, B., Greiner, N., Braun, H., Peter, F., Mörl, M., … Glökler, J. (2011). Probing the SELEX process with next-generation sequencing. *PloS One*, *6*, e29604.

Shangguan, D., Li, Y., Tang, Z., Cao, Z. C., Chen, H. W., Mallikaratchy, P., … Tan, W. (2006). Aptamers evolved from live cells as effective molecular probes for cancer study. *Proceedings of the National Academy of Sciences*, *103*, 11838–11843.

Shangguan, D., Li, Y., Tang, Z., Cao, Z. C., Chen, H. W., Mallikaratchy, P., … Tan, W. (2006). Aptamers evolved from live cells as effective molecular probes for cancer study. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 11838–43.

She, R. C., Durrant, R. J., & Petti, C. A. (2009). Evaluation of enzyme immunoassays to detect Clostridium difficile toxin from anaerobic stool culture. *American Journal of Clinical Pathology*, *131*, 81–4.

Shi, H., Fan, X. C., Ni, Z. Y., & Lis, J. T. (2002). Evolutionary dynamics and population control during in vitro selection and amplification with multiple targets. *Rna-a Publication of the Rna Society*, *8*, 1461–1470.

Shoji, A., Kuwahara, M., Ozaki, H., & Sawai, H. (2007). Modified DNA aptamer that binds the (R)-isomer of a thalidomide derivative with high enantioselectivity. *Journal of the American Chemical Society*, *129*, 1456–1464.

Soler, P., Nogareda, F., & Cano, R. (2008). Rates of Clostridium difficile infection in patients discharged from Spanish hospitals, 1997-2005. *Infection Control and Hospital Epidemiology*, *29*, 887–9.

Sreedhara, A., Li, Y., & Breaker, R. R. (2004). Ligating DNA with DNA. *Journal of the American Chemical Society*, *126*, 3454–3460.

Stamper, P. D., Alcabasa, R., Aird, D., Babiker, W., Wehrlin, J., Ikpeama, I., & Carroll, K. C. (2009). Comparison of a commercial real-time PCR assay for tcdB detection to a cell culture cytotoxicity assay and toxigenic culture for direct detection of toxin-producing Clostridium difficile in clinical samples. *Journal of Clinical Microbiology*, *47*, 373–8.

Stellrecht, K. A., Espino, A. A., Maceira, V. P., Nattanmai, S. M., Butt, S. A., Wroblewski, D., … Musser, K. A. (2014). Premarket evaluations of the IMDx C. difficile for Abbott m2000 Assay and the BD Max Cdiff Assay. *Journal of Clinical Microbiology*, *52*, 1423–8.

Steurer, M., Montillo, M., Scarfò, L., Mauro, F. R., Andel, J., Wildner, S., … Gobbi, M. (2014). Results from a Phase IIa Study of the Anti-CXCL12 Spiegelmer Olaptesed Pegol (NOX-A12) in Combination with Bendamustine/Rituximab in Patients with Chronic Lymphocytic Leukemia. *Blood*, *124*, 1996.

Stojanovic, M. N., & Kolpashchikov, D. M. (2004). Modular aptameric sensors. *Journal of the American Chemical Society*, *126*, 9266–70.

Stoltenburg, R., Reinemann, C., & Strehlitz, B. (2005). FluMag-SELEX as an advantageous method for DNA aptamer selection. *Analytical and Bioanalytical Chemistry*, *383*, 83–91.

Suga, H., Lohse, P. A., & Szostak, J. W. (1998). Structural and kinetic characterization of an acyl transferase ribozyme. *Journal of the American Chemical Society*, *120*, 1151–1156.

Ternan, N. G., Jain, S., Srivastava, M., & McMullan, G. (2012). Comparative transcriptional analysis of clinically relevant heat stress response in Clostridium difficile strain 630. *PloS One*, *7*, e42410.

Thiel, W. H., Bair, T., Peek, A. S., Liu, X., Dassie, J., Stockdale, K. R., … Giangrande, P. H. (2012). Rapid identification of cell-specific, internalizing RNA aptamers with bioinformatics analyses of a cell-based aptamer selection. *PloS One*, *7*, e43836.

Tok, J. B.-H., & Fischer, N. O. (2008). Single microbead SELEX for efficient ssDNA aptamer generation against botulinum neurotoxin. *Chemical Communications (Cambridge, England)*, 1883–5.

Tuerk, C., & Gold, L. (1990a). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, *249*, 505–510.

Tuerk, C., & Gold, L. (1990b). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, *249*, 505–510.

Tuerk, C., MacDougal, S., & Gold, L. (1992). RNA pseudoknots that inhibit human immunodeficiency virus type 1 reverse transcriptase. *Proceedings of the National*

Academy of Sciences of the United States of America, *89*, 6988–92.

Turgeon, D. K., Novicki, T. J., Quick, J., Carlson, L., Miller, P., Ulness, B., … Fritsche, T. R. (2003). Six rapid tests for direct detection of Clostridium difficile and its toxins in fecal samples compared with the fibroblast cytotoxicity assay. *Journal of Clinical Microbiology*, *41*, 667–70.

Uy, R., & Wold, F. (1977). Posttranslational covalent modification of proteins. *Science*, *198*, 890–6.

Vaish, N. K., Larralde, R., Fraley, A. W., Szostak, J. W., & McLaughlin, L. W. (2003). A novel, modification-dependent ATP-binding aptamer selected from an RNA library incorporating a cationic functionality. *Biochemistry*, *42*, 8842–8851.

Valiquette, L., Low, D. E., Pépin, J., & McGeer, A. (2004). Clostridium difficile infection in hospitals: a brewing storm. *CMAJ : Canadian Medical Association Journal = Journal de l'Association Medicale Canadienne*, *171*, 27–9.

van den Berg, R. J., Bruijnesteijn van Coppenraet, L. S., Gerritsen, H.-J., Endtz, H. P., van der Vorm, E. R., & Kuijper, E. J. (2005). Prospective multicenter evaluation of a new immunoassay and real-time PCR for rapid diagnosis of Clostridium difficile-associated diarrhea in hospitalized patients. *Journal of Clinical Microbiology*, *43*, 5338–40.

Vater, A., & Klussmann, S. (2015, January). Turning mirror-image oligonucleotides into drugs: The evolution of Spiegelmer® therapeutics. *Drug Discovery Today*.

Walsh, R., & DeRosa, M. C. (2009). Retention of function in the DNA homolog of the RNA dopamine aptamer. *Biochemical and Biophysical Research Communications*, *388*, 732–5.

Wang, K. Y., McCurdy, S., Shea, R. G., Swaminathan, S., & Bolton, P. H. (1993). A DNA aptamer which binds to and inhibits thrombin exhibits a new structural motif for DNA. *Biochemistry*, *32*, 1899–1904.

Wang, W., Billen, L. P., & Li, Y. (2002). Sequence diversity, metal specificity, and catalytic proficiency of metal-dependent phosphorylating DNA enzymes. *Chemistry and Biology*, *9*, 507–517.

Whatley, A. S., Ditzler, M. A., Lange, M. J., Biondi, E., Sawyer, A. W., Chang, J. L., … Burke, D. H. (2013). Potent Inhibition of HIV-1 Reverse Transcriptase and Replication by Nonpseudoknot, "UCAA-motif" RNA Aptamers. *Molecular Therapy. Nucleic Acids*, *2*, e71.

Wiegand, T. W., Janssen, R. C., & Eaton, B. E. (1997). Selection of RNA amide synthases. *Chemistry & Biology*, *4*, 675–683.

Williams, K. P., Ciafre, S., & Tocchini-Valentini2, G. P. (1995). Selection of novel Mg2+-dependent self-cleaving ribozymes. *The EMBO Journal*, *14*, 4551–4557.

Wise, A., Gearing, K., & Rees, S. (2002). Target validation of G-protein coupled receptors. *Drug Discovery Today*, *7*, 235–246.

Wren, B., Clayton, C., & Tabaqchali, S. (1990). Rapid identification of toxigenic Clostridium difficile by polymerase chain reaction. *Lancet*, *335*, 423.

Yang, D., Meng, X., Yu, Q., Xu, L., Long, Y., Liu, B., … Zhu, H. (2013). Inhibition of hepatitis C virus infection by DNA aptamer against envelope protein. *Antimicrobial Agents and Chemotherapy*, *57*, 4937–44.

Yunusov, D., So, M., Shayan, S., Okhonin, V., Musheev, M. U., Berezovski, M. V, & Krylov, S. N. (2009). Kinetic capillary electrophoresis-based affinity screening of aptamer clones. *Analytica Chimica Acta*, *631*, 102–7.

Zhang, B., & Cech, T. R. (1997). Peptide bond formation by in vitro selected ribozymes. *Nature*, *390*, 96–100.

Zhao, Z., Xu, L., Shi, X., Tan, W., Fang, X., & Shangguan, D. (2009). Recognition of subtype non-small cell lung cancer by DNA aptamers selected from living cells. *The Analyst*, *134*, 1808–14.

Zhou, W., Huang, P.-J. J., Ding, J., & Liu, J. (2014). Aptamer-based biosensors for biomedical diagnostics. *The Analyst*, *139*, 2627–40.

Zhu, Q., Shibata, T., Kabashima, T., & Kai, M. (2012). Inhibition of HIV-1 protease expression in T cells owing to DNA aptamer-mediated specific delivery of siRNA. *European Journal of Medicinal Chemistry*, *56*, 396–9.

Zilberberg, M. D., Shorr, A. F., & Kollef, M. H. (2008). Increase in adult Clostridium difficile-related hospitalizations and case-fatality rate, United States, 2000-2005. *Emerging Infectious Diseases*, *14*, 929–31.

Zimmermann, B., Gesell, T., Chen, D., Lorenz, C., & Schroeder, R. (2010). Monitoring genomic sequences during SELEX using high-throughput sequencing: neutral SELEX. *PloS One*, *5*, e9169.

Appendix I: Oligonucleotides

Oligonucleotide	Sequence (5'-3')
NAp1 (w/primers)	AGTGTCAGCCAGTATAACCCACGCATGGTCTTGTACTC GTGCGGGCCGTCTGAGGGAAGTTTAACCCGGAGTGAAC ACCTA
CAp1 (w/primers)	CCAAACAGCCAGTATGTCAGTGTATTTGGTGTTGTGAGG GCGGGGGGTTTGGCGGAGATGGTGGGTTATACTGGCTGA CACT
KAp1 (w/primers)	AACAGCCCAAGTCAGGTCTATGACCGGGTAGGATTATCC CGGCAGCTGTATGAAGCCAGTAGTAGTTCTGACTGACGC TTGG
NAp1 5' Cmp	TGGGTTATACTGGCTGACACT
NAp1 3' Cmp	TAGGTGTTCACTCCGGGTTA
CAp1 5' Cmp	ACTGACATACTGGCTGTTTGG
CAp1 3' Cmp	AGTGTCAGCCAGTATAACCCA
CAp1-M1	GGTAGTAGCGGTCAGTGTATTTGGTGTTGTGAGGGCGGG GGGTTTGGCGGAGATGGTGGGTTCGCTACTACC
NAp1-M4	AGTATAACCCACGCATGGTCTTGTACTCGTGCGGGCCGT CTGAGGGAAGTT
NAp1-M4-T10	AGTATAACCCACGCATGGTCTTGTACTCGTGCGGGCCGT CTGAGGGAAGTTTTTTTTTTT
N4C1-T10 Splint	ATACACTGACCGCTACTACCAAAAAAAAAAAACTTCCCTC

Table A1. List of oligonucleotides used during the project.