Theophylline IMAGEtags (Intracellular Multi Aptamer Genetic tags): the development and evaluation of an RNA reporter system based on the theophylline aptamer

Hans Eirik Haarberg
Iowa State University

Follow this and additional works at: http://lib.dr.iastate.edu/etd
Part of the Neurosciences Commons

Recommended Citation
http://lib.dr.iastate.edu/etd/10578

This Thesis is brought to you for free and open access by the Graduate College at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Theophylline IMAGEtags (intracellular multi aptamer genetic tags): the development and evaluation of an RNA reporter system based on the theophylline aptamer

by

Hans Eirik Haarberg

A thesis submitted to the graduate faculty in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Neuroscience

Program of Study Committee:
Marit Nilsen-Hamilton, Major Professor
Donald S. Sakaguchi
David Hannapel

Iowa State University
Ames, Iowa
2009

Copyright © Hans Eirik Haarberg, 2009. All rights reserved.
# TABLE OF CONTENTS

## ABSTRACT

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. General Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Current reporter genes</td>
<td>1</td>
</tr>
<tr>
<td>Fluorescent proteins</td>
<td>3</td>
</tr>
<tr>
<td>In Vivo Imaging</td>
<td>6</td>
</tr>
<tr>
<td>RNA Imaging</td>
<td>8</td>
</tr>
<tr>
<td>Aptamers</td>
<td>9</td>
</tr>
<tr>
<td>Aptamer structure and stabilization</td>
<td>10</td>
</tr>
<tr>
<td>Applications of aptamers</td>
<td>11</td>
</tr>
<tr>
<td>Riboswitches - natural RNAs with aptamer-like properties</td>
<td>12</td>
</tr>
<tr>
<td>Application of aptamers for imaging</td>
<td>13</td>
</tr>
<tr>
<td>Sulforhodamine-B and fluorescein aptamers</td>
<td>13</td>
</tr>
<tr>
<td>Fluorescence switches and “Light-up” aptamers</td>
<td>14</td>
</tr>
<tr>
<td>Theophylline aptamer for imaging</td>
<td>17</td>
</tr>
<tr>
<td>Theophylline</td>
<td>17</td>
</tr>
<tr>
<td>The theophylline aptamer</td>
<td>19</td>
</tr>
<tr>
<td>The IMAGEtag reporter system</td>
<td>22</td>
</tr>
<tr>
<td>Collaborator contributions</td>
<td>22</td>
</tr>
</tbody>
</table>

## CHAPTER 2. CLONING OF MULTIPLE TANDEM REPEATS OF PROTEIN DOMAINS AND APTAMERS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>24</td>
</tr>
<tr>
<td>Introduction</td>
<td>25</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>27</td>
</tr>
<tr>
<td>Materials</td>
<td>27</td>
</tr>
<tr>
<td>Cloning of initial repetitive sequences from synthetic oligonucleotides</td>
<td>27</td>
</tr>
<tr>
<td>Construction of longer repetitive sequences</td>
<td>28</td>
</tr>
<tr>
<td>Results and Discussion</td>
<td>28</td>
</tr>
<tr>
<td>Repetitive aptamer sequences constructed using oligonucleotides</td>
<td>28</td>
</tr>
<tr>
<td>Multiplication of the original repetitive sequences</td>
<td>29</td>
</tr>
<tr>
<td>Other repetitive constructs</td>
<td>31</td>
</tr>
<tr>
<td>Stability of repetitive constructs</td>
<td>31</td>
</tr>
<tr>
<td>Summary</td>
<td>32</td>
</tr>
</tbody>
</table>
CHAPTER 3. STUDY OF THE THEOPHYLLINE IMAGETAG AND LIGANDS

Abstract
Introduction
Materials and Methods
  Plasmids and Reagents
  Cloning of IMAGETags into expression vectors
  Assay of theophylline toxicity in CCL64 cells
  Expression of IMAGETags in CCL64 cells and analysis by real time RT-PCR
  Uptake and efflux of rhodamine-B and GK38
  Determining the intracellular localization of GK38 by confocal microscopy
  Efflux of theophylline-[8-3H]
  Quantification of theophylline-[8-3H] in cells with and without IMAGETags
  Determination of the stability of IMAGETag RNA

Results and Discussion
  Cloning of IMAGETags into expression vector
  Toxicity of theophylline in CCL64 cells
  Expression of IMAGETags in CCL64 cells
  Uptake and efflux of rhodamine-B and GK38
  Cellular localization of GK 38
  Efflux of theophylline-[8-3H]
  Effect of IMAGETags on the intracellular concentration of theophylline-[8-3H]
  Stability and half-life of the IMAGETag RNA in CLL64 cells

CHAPTER 4. CONCLUSIONS AND FUTURE DIRECTIONS

LITERATURE CITED
ABSTRACT

Reporter genes, such as luciferase and green fluorescent protein (GFP), have simplified efforts to study transcriptional regulatory elements, tag endogenous proteins, and label and track individual cells in vivo. While these reporters represent a major advance in cellular imaging they have significant limitations. Luciferase, GFP and other reporter genes are restricted by a high cellular energy requirement for their synthesis and a long lag time between the initiation of transcription and reporter protein expression. Also, the method for study of endogenous RNAs utilizing reporter genes is quite complex and limited.

RNA reporters could allow determination of rapid changes in gene expression and may also be used as a tag to make RNA fusion constructs to track other RNAs in vivo. As there is no known naturally produced RNA with inherent fluorescence sufficient for imaging, our approach has been to use RNA aptamers to bind a target molecule with a fluorescent or radioactive label. Aptamers, single stranded RNA or DNA molecules, have high binding affinity and specificity, low molecular mass, and have been selected to recognize enzymes, receptors, growth factors and small molecules.

This work describes part of an effort to develop an RNA tag and reporter gene using the RNA aptamer for theophylline. For the tag to be successful in vivo, the ligand cannot be a natural component of the cell. Theophylline and its aptamer were chosen because theophylline is a small molecule drug and its aptamer has proven functional in bacteria, yeast and cultured human cells.
The proposed IMAGEtag (intracellular multi aptamer genetic tag) will consist of a string of multiple tandem aptamers that can be expressed in cells as RNAs. Multiaptamers were constructed in order to increase the binding capacity of the IMAGEtags. Initially, short multiaptamers were cloned using synthetic oligonucleotides. Elongated multiaptamers were constructed utilizing a modified recursive directional ligation (RDL) technique that employs the non palindromic restriction enzymes BsaI and BsmAI. The multiaptamers were cloned into a mammalian expression vector containing a β-actin promoter and CMV enhancer and expressed in CCL64 cells. Expression and stability of the IMAGEtags were analyzed by real time reverse transcription PCR. The ligands for this IMAGEtag were radioactive theophylline-[³H] and GK38, a theophylline analog labeled with the fluorescent molecule rhodamine-B. GK38 was examined utilizing assays for cellular efflux and localization. Efflux assays were also performed with rhodamine-B, and theophylline-[³H]. Theophylline-[³H] was used to investigate whether IMAGEtag expression had an effect on the cellular concentration of theophylline.

Cloning of a theophylline multiaptamer using synthetic oligonucleotides resulted in a construct with six aptamer repeats, which was used to create constructs with 18, 90 and 270 repeats by the RDL technique. Expression of the 18 repeat multiaptamer in CCL64 cells resulted in a maximum RNA level 24 hours after transfection, and the RNA was determined to have high stability with an approximate half-life of 43 hours. Using a rhodamine-theophylline conjugate to track theophylline in the cells revealed that it binds to endogenous cellular
components localized mainly in the Golgi apparatus. Efflux assays showed that 55-60% of the conjugate remains bound to the cells after 1 hour, compared to 30-40% for rhodamine-B and 60-80% for theophylline. It was determined that expression of the theophylline IMAGEtag did not have any effect on the cellular concentration of theophylline. However, it was not determined if the tandemly linked theophylline aptamers in this RNA were capable of binding theophylline.

The observed binding of theophylline and rhodamine-theophylline to cellular components creates a high background signal that obscures a potential signal from the IMAGEtag. This could limit the effectiveness of a reporter system based on theophylline and the theophylline aptamer, and other aptamers and ligands may be more suitable for use as reporters. Alternatively, utilizing Förster resonance energy transfer (FRET), a pair of fluorescently labeled aptamer ligands could generate a unique signal when binding to adjacent aptamers in an IMAGEtag.
CHAPTER 1. GENERAL INTRODUCTION

This work was part of an effort to develop an RNA tag and reporter gene using the RNA aptamer for theophylline. An RNA reporter could provide a more direct measure of transcriptional activity than enzyme based reporter genes, and potentially allow determination of rapid changes in gene expression. Additionally, the RNA reporter may also be used as a tag to make RNA fusion constructs, to track other RNAs in the cell and organism, in much the same way as fluorescent proteins have been used to study other proteins. Since there is no known naturally produced RNA with inherent fluorescence sufficient for imaging, our approach has been to use RNA aptamers that can bind to a target molecule with a fluorescent label. My colleagues, PhD candidates, Ilchung Shin and Muslum Ilgu, also participated in this project with their focus being on aptamers against aminoglycoside antibiotics (kanamycin, tobramycin and neomycin).

Current reporter genes

Reporter genes were developed as tools to study transcriptional regulatory elements, namely promoters and enhancers. To determine the activity of a promoter, a vector is created in which the promoter is placed upstream of the reporter gene to regulate its expression. The reporter gene product could be easily detected and quantified. The transcriptional activity of the promoter could then be estimated based on the amount of reporter gene product. Vectors that express reporter genes from a constitutive promoter could also be used to
determine the efficiency of transfection or other gene delivery methods (Alam and Cook, 1990).

Early reporter genes include chloramphenicol acetyltransferase (CAT), luciferase, human growth hormone (hGH), alkaline phosphatase and β-galactosidase. hGH is secreted from the cells and can be detected by radioimmunologic assay of the culture medium. The rest are enzymes and would be assayed based on their activity, usually by lysing and extracting the components of the cells. β-galactosidase and luciferase can also be analyzed in situ, to identify specific cells that express these enzymes. CAT was initially the most used reporter gene, but has for the most part been replaced by luciferase (Alam and Cook, 1990).

Luciferase from the firefly Photinus pyralis catalyzes the ATP driven oxidative decarboxylation of firefly luciferin, resulting in the emission of a photon of light. Firefly luciferase can be quantified by measuring the emission of light with a luminometer. It gained popularity because the luciferase assay is simpler and yields higher sensitivity (30- to 1000-fold) compared to CAT (de Wet et al., 1987; Alam and Cook, 1990).

The availability of reporter genes has simplified efforts to quantify the activity of transcriptional regulatory elements. Reporter genes have also been valuable in studies to determine the efficiency of gene delivery methods. The application of reporter genes has expanded to imaging of cells, as well as intracellular components, with the discovery of fluorescent proteins.
Fluorescent proteins

Fluorescent proteins (FP) do not require a substrate, and their fluorescence can be detected in living cells. Fluorescent microscopy enables researchers to study changes in expression in specific cells, and determine the intracellular location of the FP. The locations, movements and interactions of proteins have been elucidated by generating FP fusion proteins. FPs with different colors can be used to study several processes simultaneously, and certain pairs of FPs can be used in Förster resonance energy transfer (FRET) experiments to study interactions between different proteins (Chudakov et al., 2005; Shaner et al., 2005).

Green fluorescent protein (GFP) was discovered in 1962 during isolation of the bioluminescent protein aequorin from the jellyfish *Aequoria victoria* (Shimomura et al., 1962). Aequorin emits blue light (470 nm), while the intact organism emits green light (509 nm). It was demonstrated that the luminescence from aequorin is transferred to GFP, which emits green light. GFP has a maximal excitation at 400 nm with a second small peak at 480 nm. Its emission maximum is 508-509 nm, and the quantum yield is 0.72 when excited at 470 nm (Morise et al., 1974).

Development of GFP as a reporter gene began with the cloning and sequencing of the GFP gene and cDNA in 1992. The cDNA encodes a 238 amino acid protein with a calculated molecular weight of 26.9 kDa (Prasher et al., 1992). The first report of expression of functional GFP in cells other than *A. victoria* was published two years later in 1994. GFP expressed in *Escherichia*
coli and Caenorhabditis elegans displayed identical fluorescence properties to natively expressed GFP (Chalfie et al., 1994). The chromophore (p-hydroxybenzylidene-imidazolidinone) is formed spontaneously by cyclization and dehydrogenation from the three amino acids Ser65-Tyr66-Gly67 in the GFP polypeptide (Cody et al., 1993; Heim et al., 1994).

Much work has been done to improve and modify certain properties of GFP, such as the efficiency of folding and chromophore formation, exogenous expression levels, fluorescence intensity and the wavelengths of excitation and emission. Variants of GFP were created by inducing random mutations of the 20 amino acids in and around the chromophore. These mutants were expressed in E. coli. Mutants with improved excitation at 488 nm, for use with the 488 nm argon laser and fluorescein-isothiocyanate (FITC) filters, were selected by fluorescence-activated cell sorting (FACS). One mutant (GFPmutl; F64L, S65T) exhibited a 35-fold increase in fluorescence, when excited at 488 nm. This mutant also displayed improved folding, resulting in an approximately 100-fold increase in fluorescence in E. coli 1 hour after induction of FP expression (Cormack et al., 1996). Based on this mutant, enhanced green fluorescent protein (eGFP) was developed by a group at Clontech introducing silent mutations to optimize for preferred human codons. This resulted in a 4-fold increase in expression of eGFP in human cells compared to GFPmutl (Yang et al., 1996; Zhang et al., 1996).

Similar mutation and screening experiments have been used to develop color variants of GFP such as cyan (eCFP) and yellow (YFP) fluorescent proteins.
(Heim and Tsien, 1996; Miyawaki et al., 1997). The brightness of eCFP was relatively low (33% of GFP), but the improved variant Cerulean has a 2.5-fold increase in brightness compared to eCFP, as well as improved resistance to photobleaching (Rizzo et al., 2004). Improved variants of YFP, such as Citrine and Venus, were developed to decrease sensitivity to pH and chloride ions, which were the main problem with the original YFP (Nagai et al., 1989; Griesbeck et al., 2001). Blue fluorescent proteins (BFP) have also been developed from GFP, but they have not been commonly used due to low brightness and very low resistance to photobleaching. These properties have been improved in newer forms of BFP such as Azurite (Mena et al., 2006).

Six additional fluorescent proteins were discovered in a variety of nonbioluminescent anthozoa species, including DsRed, a tetrameric red fluorescent protein from *Discosoma striata* (Matz et al., 1999). Mutation and selection of DsRed were used to decrease the maturation half-time from 24 hours to less than 1 hour, and eliminate the partial green fluorescence exhibited by the original protein (Bevis and Glick, 2002). Additional modifications were made to disrupt oligomerization, and generate the monomer mRFP1 (Campbell et al., 2002). Mutations of the mRFP1 sequence were made to create a number of fluorescent proteins in the orange to far red spectrum, notably mOrange, tdTomato, mCherry (Shaner et al., 2004) and mPlum (Wang et al., 2004).
**In Vivo Imaging**

Several gene reporter systems have been developed for noninvasive imaging *in vivo* in order to determine the location and level of gene expression. Potential reporter systems have been explored using various imaging modalities including bioluminescence, fluorescence and radioactivity (Tsien, 2003).

Firefly luciferase has been a popular reporter gene because its bioluminescence is sufficient to detect fewer than a hundred cells. The long wavelength light (>600 nm) can penetrate through several mm of tissue, although increasing depth leads to both a weaker signal and lower resolution due to scattered photons (Tsien, 2003). For fluorescent proteins, the limited penetration of light through tissues is a challenge for both excitation as well as detection of emission. However, recent advances in optical imaging technology have greatly enhanced the utility of bioluminescent and fluorescent reporters *in vivo*. Detection of bioluminescence *in vivo* is significantly improved by the use of charge coupled device (CCD) cameras that are cooled to -120°C. The increased sensitivity and signal-to-noise ratio of CCD cameras makes it possible to detect luciferase activity in deeper tissues (Contag and Bachmann, 2002). Although bioluminescent imaging does not provide high resolution, advances in fluorescent microscopy make it possible to image fluorescence at single cell resolution within tissues (Mempel et al., 2004). Recently, bimodal reporter systems utilizing both bioluminescence and fluorescence were used to study human neural stem cells (hNSCs) and glioma cells implanted into the brains of mice. This system provides the possibility for both real time imaging of global cell survival and migration
using bioluminescence, as well as determining the fates of implanted cells at single cell resolution using fluorescent proteins (Shah et al., 2008).

Two imaging methods that utilize radioactive isotope labeled tracers have been widely used in humans for medical purposes (e.g. detecting tumors). Single photon emission computed tomography (SPECT) and positron emission tomography (PET) are used to detect isotopes that emit gamma rays and positrons respectively. Utilizing SPECT, gamma rays can penetrate through tissue and be detected directly by a gamma camera which is rotated around the subject to create three dimensional images. Positrons do not penetrate through the tissue, but instead collide with electrons and undergo annihilation, which causes two gamma rays to be emitted at close to 180 degrees. The simultaneous detection of two gamma rays is used to localize the tracer and generate the tomographic image in a PET scan. Differences in the underlying principles, equipment and isotopes used in PET and SPECT, give each method certain limitations and advantages. The major advantage of PET is that it has over ten times the sensitivity of SPECT. Additionally, the positron emitting isotopes used in PET, such as $^{11}$C, $^{13}$N, $^{15}$O and $^{18}$F, can be used to substitute atoms or groups in organic molecules. However, different isotopes cannot be distinguished with PET because the gamma rays created from positron annihilations have the same energy (511 keV). Gamma rays emitted directly from atomic nuclei can have a range of energies, and dual-isotope SPECT can be used to image several isotopes simultaneously provided that the energies of their gamma rays are sufficiently different (Gambhir, 2002).
Intracellular enzymes, cell surface receptors or membrane transporters can serve as reporter genes for PET or SPECT. A popular PET reporter is the herpes simplex virus 1 thymidine kinase (HSV1-tk), which can trap thymidine analogs such as $^{18}$F-fluoroganciclovir (FGCV) inside cells by phosphorylation (Gambhir et al., 1999). The dopamine type 2 receptor (D2R) has been used as a PET reporter with the ligand $^{18}$F-fluoroethylspiperone (FESP) (MacLaren et al., 1999). Membrane transporters can accumulate a tracer by pumping it into the cell. The sodium-iodide symporter has recently been used to image cardiac-derived stem cells implanted into the rat heart. The tracers $^{124}$I and $^{99m}$Tc were used for PET and SPECT imaging, respectively (Terrovitis et al., 2008).

RNA Imaging

There have been important developments in protein imaging, but currently available methods for detecting and imaging RNA have significant limitations. Nucleic acids can be detected and imaged using fluorescent in situ hybridization (FISH). FISH utilizes fluorescently labeled nucleic acid probes that are complementary to the target RNA or DNA sequence. FISH has been used mainly in applications where the target is DNA, such as localizing a specific sequence on chromosomes (Volpi and Bridger, 2008). However, FISH can also be utilized to study mRNA expression (Femino et al., 1998) and cellular distribution (Guzowski et al., 2001). Although FISH is very useful for identification and localization of specific mRNA species, this technique requires that the cells or tissues to be studied are fixed.
A method of tracking RNA in living cells was developed by labeling an mRNA with six repeats of a 19-nucleotide RNA stem loop, which serve as targets for the RNA binding protein MS2. The labeled RNA could then be detected by binding to MS2-GFP fusion proteins (Bertrand et al., 1998). The disadvantage of this method is that the MS2-GFP protein has to be expressed simultaneously with the labeled RNA. The fluorescent signal of GFP is always present and is redistributed upon binding to the RNA. Additionally, the extra RNA sequence and multiple MS2-GFP fusion proteins give the tag a mass of approximately 270kDa, which could have significant effects on the movement and function of the RNA (Tsien, 2003).

Aptamers

In 1990, two groups independently developed an in vitro selection method for generating synthetic oligonucleotides that could bind with high affinity and specificity to chosen target molecules. Ellington and Szostak selected RNA sequences that could bind to 7 different dyes, and named them aptamers after aptus, a Latin word meaning “to fit” (Ellington and Szostak, 1990). Concurrently, Tuerk and Gold developed an RNA aptamer for T4 DNA polymerase and coined the term SELEX (Systematic Evolution of Ligands by EXponential enrichment) for the selection process (Tuerk and Gold, 1990).

The SELEX procedure starts with a pool of synthetic oligonucleotides consisting of a sequence of random nucleotides flanked by a constant primer binding sequence at both ends. Oligonucleotides in the pool are selected for
binding to a molecule of interest by incubation with an immobilized form of the target molecule followed by washing to remove non binding sequences and retrieval of binding sequences by elution with the target molecule. The bound fraction is amplified by PCR under low fidelity conditions so that some additional variability is introduced in the amplified product. Repeating the selection and amplification steps exponentially enriches the pool for binding sequences. Individual sequences are cloned and sequenced, and a number of sequences are analyzed to determine possible consensus sequences (Ellington and Szostak, 1990; Tuerk and Gold, 1990).

**Aptamer structure and stabilization**

Aptamers, which are composed of nucleic acids, form a variety of intricate structures. Stem loops are common feature of many aptamer structures. Mismatched bases within aptamer stem loops, often composed of pairs of purines, form binding pockets for small molecule ligands. Base stacking stabilizes the interaction between RNA aptamers and ligands by hydrophobic interactions and hydrogen bonding (Patel and Suri, 2000).

RNA is sensitive to both enzymatic and non-enzymatic degradation. The original hairpin loop of the theophylline aptamer was found to be susceptible to cleavage catalyzed by Mg$^{2+}$ in solution (Zimmermann et al., 2000). This degradation was abolished by replacing the original loop with a GAAA tetraloop, which is a common feature of many natural RNAs (Heus and Pardi, 1991).
Additional stabilization is required for \textit{in vivo} use, particularly if the application is extracellular. Several modifications are used to stabilize nucleic acids and make them more resistant to degradation by nucleases. Modified bases such as 2'-fluoro- and 2'-amino- pyrimidines were used to stabilize the hammerhead ribozyme (Pieken et al., 1991). An aptamer for HIV Type-1 nucleocapsid protein was protected from degradation by linking the ends together to make it circular (Kim et al., 2002). Mirror image aptamers selected for D-adenosine (Klussmann et al., 1996) and L-arginine (Nolte et al., 1996) display much higher stability \textit{in vivo} than their natural counterparts. Locked nucleic acids (LNA) exhibit higher melting temperatures than RNA or DNA, and an RNA aptamer modified as an LNA was also resistant to nucleases (Schmidt et al., 2004).

\textbf{Applications of aptamers}

Aptamers have gained interest as potential therapeutics due to their high binding affinities and specificities (comparable to antibodies) yet low molecular mass (a 40 nucleotide RNA would have a mass of approximately 13kDa). Aptamers have been selected against enzymes, receptors, growth factors, prions, toxins and viral capsule proteins (Gopinath, 2007).

A 2'-F-pyrimidine aptamer was selected against vascular endothelial growth factor (VEGF). This aptamer binds with very high affinity to the isoform VEGF\textsubscript{165} ($K_D$ 49 and 130 pM), but does not bind to isoform VEGF\textsubscript{121} (Ruckman et al., 1998). Pegaptanib, a version of the VEGF aptamer conjugated to
polyethylene glycol (PEG), was approved for the treatment of age related macular degeneration by the FDA in 2004 (Kourlas and Schiller, 2006).

Many aptamers have also been selected against small organic molecules. Molecules that bind to the ribosome and other cellular RNAs have been especially popular targets for aptamer selection. There are aptamers for a large number of antibiotics including tetracycline, chloramphenicol, viomycin, lividomycin, and several aminoglycosides (kanamycin, tobramycin, neomycin, streptomycin). Other targets include theophylline, dopamine, several nucleotides, amino acids, and cofactors. The structures of many aptamer small molecule complexes have been determined by NMR spectroscopy (Famulok, 1999; Patel and Suri, 2000). Aptamers have been used to develop biosensors for several drugs and other small molecules, including theophylline (Rankin et al., 2006; Ferapontova et al., 2008), adenosine (Lu et al., 2008) and cocaine (Baker et al., 2006). Others aptamers have been used to control gene expression in vivo (Weigand and Suess, 2007; Tuleuova et al., 2008).

**Riboswitches - natural RNAs with aptamer-like properties**

More recently, it has been discovered that many RNAs with properties similar to aptamers occur naturally in bacteria and some eukaryotes. The majority of these are riboswitches that are often present in mRNAs and are involved in the regulation of translation. A thiamine (vitamin B₁) riboswitch was found in an mRNA in *E. coli*. The mRNA encodes an enzyme involved in the biosynthesis of thiamine, and binding of thiamine to the riboswitch inhibits ribosome binding and
translation. Thus the riboswitch is responsible for feedback inhibition and limits the synthesis of the enzyme when the concentration of thiamine is sufficient (Winkler et al., 2002a). A similar feedback mechanism involving the biosynthesis of flavin mononucleotide (FMN, vitamin B$_2$) was found in *Bacillus subtilis*. Although the FMN riboswitch was present in an mRNA, the mechanism of inhibition seems to involve premature transcriptional termination rather than translational inhibition (Winkler et al., 2002b).

**Application of aptamers for imaging**

There have been several attempts to develop RNA tags based on aptamers. The first aptamers were selected for the fluorophores sulforhodamine B and fluorescein with the goal of sequestering the respective fluorophores (Holeman et al., 1998). It was then found that an aptamer for malachite green can induce structural changes in its ligand. Malachite green is not fluorescent on its own, but becomes fluorescent upon binding to its aptamer (Babendure et al., 2003). Several aptamers have recently been developed against different dyes in attempts to make aptamer dye pairs that become fluorescent when they associate (Sparano and Koide, 2007; Constantin et al., 2008; Sando et al., 2008). There has not yet been any demonstration of successful application of any RNA tag in living cells.

**Sulforhodamine-B and fluorescein aptamers**

Aptamers were selected for two commonly used fluorescent molecules sulforhodamine B and fluorescein. The selection process was performed using a
buffer with intracellular ion concentrations (100 mM K\(^+\), 10 mM Na\(^+\), 5 mM Mg\(^{2+}\)).

The sulforhodamine binding aptamer was selected from a pool of oligonucleotides with a random sequence of 72 nucleotides. An initial sequence was determined after seven rounds of selection for binding to sulforhodamine agarose. The last step included an elution with fluorescein to select against any RNA that was not specific for sulforhodamine. A 54 nucleotide truncated form of the aptamer was specific for sulforhodamine, and a K\(_D\) of 310 ± 60 nM was determined by fluorescence anisotropy. An aptamer for fluorescein was selected from a doped pool of RNA based on the sulforhodamine aptamer (Holeman et al., 1998).

The fluorescein and sulforhodamine aptamers were biotinylated and bound to separate streptavidin beads. The labeled beads were able to concentrate their respective ligands in a solution containing 100 nM sulforhodamine and 125 nM fluorescein. The authors of this work discuss the possibility of using the aptamers to label RNA transcripts, and to measure the proximity between different aptamers by FRET (Holeman et al., 1998). There has not been any report of further progress using these aptamers.

**Fluorescence switches and “Light-up” aptamers**

An aptamer for the dye malachite green (MG) was found to influence the fluorescent properties of its ligand. MG has a very low fluorescence when it is free in solution, but binding to the MG aptamer increased the quantum yield by 2360-fold (binding also shifts the maximum absorbance from 618 nm to 630 nm).
The quantum yield of bound MG is 0.187, which is still low compared to most commonly used fluorophores. To improve fluorescence, an indolinyl derivative (IMG) with higher unbound quantum yield than MG was tested. IMG exhibited a similar enhancement of fluorescence (2090-fold), and the bound form of IMG has a quantum yield of 0.324 (Babendure et al., 2003).

Sparano and Koide developed an RNA sensor by conjugating dichlorofluorescein (DCF) with two aniline moieties. Aniline is able to quench the fluorescence of DCF by photoinduced electron transfer (PET). When a fluorescent molecule is excited by absorption of light, an electron jumps from the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular orbital (LUMO). When the excited electron returns to the HOMO, the extra energy is released as a photon. PET occurs when an adjacent molecule (quencher) has a HOMO energy level between the HOMO and LUMO levels of the fluorophore. An electron can then jump from the HOMO of the quencher to the HOMO of the excited fluorophore, which results in quenching by hindering the excited electron in the LUMO to return to the HOMO. The efficiency of PET depends on the energy level of the HOMO of the quencher and the distance between quencher and fluorophore. In an attempt to control PET efficiency, an aptamer was developed for the aniline quencher. One aptamer was found to increase the fluorescence of the DCF-aniline conjugate by 13-fold at 100 µM. This demonstrates that an aptamer could be developed to control fluorescence by means of PET, but it has some limitations. The selected aptamer was over 100 nucleotides long, and its binding affinity was not reported, although high
concentrations of aptamer (> 10 µM) are required to affect fluorescence (Sparano and Koide, 2007).

Other RNA sensing probes have been developed based on fluorophores that are known to have fluorescent properties that change depending on their molecular interactions. The DNA stain Hoechst 33258 binds to AT rich regions of double stranded DNA (dsDNA), whereby its fluorescence is increased. A number of derivatives were tested for binding to dsDNA, and addition of two t-butyl groups was found to abolish its binding to dsDNA. A DNA aptamer was then selected for this derivative and binding to the aptamer resulted in a 191-fold increase in fluorescence (Sando et al., 2007). RNA aptamers were then selected for the same compound in an attempt to make a ‘blue fluorescent RNA’. A 29 nucleotide aptamer was able to enhance the fluorescence of the modified Hoechst by 56-fold (to a quantum yield of 0.26) at a concentration of 200 nM. To test its function, the aptamer was incorporated into a construct to produce a tagged luciferase mRNA. The fluorescence of the dye increased during in vitro transcription of the tagged mRNA but not with untagged mRNA (Sando et al., 2008).

A similar approach was used to develop probes based on a modified cyanine dye. Dimethyl indole red (DIR) is a modified cyanine dye, with reduced affinity for DNA. A bulky dimethyl indole group and a sulphonate group were intended to interfere with intercalation and provide repulsion from anionic phosphate groups of nucleic acids respectively. An aptamer selected for DIR was shown to increase its fluorescence by 60-fold while DNA and non specific RNA
resulted in a 2-fold increase. The actual quantum yields for either free or bound dye were not reported (Constantin et al., 2008).

**Theophylline aptamer for imaging**

Although there are numerous small molecule aptamers available that could be used to make an RNA tag, there are certain considerations that have to be taken into account regarding both the aptamer and its ligand. For the tag to eventually be successful *in vivo*, the ligand cannot be a natural component of the cell. Otherwise, the endogenous ligand would compete for aptamer binding, and reduce its capacity to bind to the labeled (fluorescent or radioactive) ligand. This excludes the amino acids and vitamin/cofactors. Additionally, the ligand should not be toxic at concentrations used for imaging. The aptamer should have high affinity and selectivity, and function in the intracellular environment.

Theophylline and its aptamer were chosen because theophylline is used as a drug, and the concentrations expected to be suitable for imaging are well below typically therapeutic doses. Theophylline also diffuses through the blood brain barrier, so the system could potentially be used in the brain. The theophylline aptamer has a reasonable affinity, and very high specificity for theophylline compared to similar compounds such as caffeine. The theophylline aptamer has also been demonstrated to function inside of cells.

**Theophylline**

Theophylline is a methylxanthine related to caffeine that has been used as a bronchodilator to treat asthma and COPD (chronic obstructive pulmonary
disease). The mechanism underlying the therapeutic effect of theophylline has not been completely determined. Theophylline is a known adenosine receptor antagonist and a phosphodiesterase (PDE) inhibitor, but it has been suggested that it may have other targets. It can increase the intracellular concentration of cyclic nucleotides (cAMP and cGMP) by inhibition of PDE. This was thought to be the main cause of its bronchodilatory effect, but it has been shown that PDE function is reduced by merely 5-20% at the therapeutic plasma concentration of 10-20 mg/l (55-110 µM) (Barnes and Pauwels, 1994). Theophylline has also been found to have anti-inflammatory effects, which are at least partly mediated by histone deacetylase (HDAC). The stimulation of HDAC activity was not observed with specific PDE inhibitors or adenosine receptor antagonists (Ito et al., 2002). This also indicates that there are additional targets for theophylline, but these have not been identified.

In humans, theophylline is metabolized in the liver by cytochrome P450 (CYP) enzymes 1A1, 1A2, 2D6 and 2E1 (Ha et al., 1995). The major metabolites are 1,3-dimethyluric acid (DMU), 3-methylxanthine (3-MX) and 1-methyl-uric acid (1-MU). It is excreted in the urine as 16.6 ± 6.5% unchanged theophylline, while DMU, 1-MU and 3-MX account for 44.3 ± 7.0%, 24.3 ± 4.8% and 12.9 ± 3.4% respectively (Gundert-Remy et al., 1983).

Cellular uptake of theophylline is facilitated by the organic human anion transporter 2 (hOat2). The intracellular concentration of theophylline in *Xenopus laevis* oocytes was increased by 5.9-fold by the expression of hOat2 after a 48 hour incubation. The uptake follows Michaelis-Menten kinetics, and the $K_m$ for
theophylline is 12.6 µM (Kobayashi et al., 2005). Oat2 is expressed mainly in the liver and kidney (Sun et al., 2001).

Theophylline can cross the blood brain barrier, and reaches a maximum concentration 20-30 min after intravenous injection of 10 or 30 mg/kg. The distribution was not affected when p-glycoprotein was inhibited by administration of 10 mg/kg cyclosporine (Tsai and Liu, 2004). The ratio of total concentration in brain to plasma is 0.31 at steady state. The unbound fraction of theophylline in plasma and brain are 0.29 and 0.31 respectively (Becker and Liu, 2006).

**The theophylline aptamer**

An RNA aptamer that binds theophylline was selected by a modified SELEX procedure. Five rounds of regular selection were performed by capturing the RNA on a sepharose column to which was linked 1-carboxypropyl theophylline, followed by elution with 0.1 M theophylline. Another three rounds were performed where an additional step (counter-
SELEX) was included to select against any RNA that could bind the structurally similar caffeine (Fig. 1.1A). In these steps the column with bound RNA was washed with 0.1 M caffeine before elution with theophylline. The resulting pool was cloned and 16 sequences were found to contain 15 conserved bases. A 42 nucleotide aptamer (TCT8-4) containing the conserved sequence was found to have high affinity for theophylline $K_d$ (dissociation constant) of 0.32 µM, which was 10000-fold higher than for caffeine (Jenison et al., 1994).

A smaller 33 nucleotide form of the theophylline aptamer (Fig. 1.1B) was made by truncating the terminal stem and replacing the original loop with a more stable GAAA tetraloop. NMR spectroscopy of a complex of this aptamer with theophylline reveals that the theophylline is stacked in between bases in the aptamer and forms hydrogen bonds with a cytosine and a uracil. The cytosine interacts with a hydrogen atom at nitrogen 7 in theophylline, which is replaced by a methyl group in caffeine (Zimmermann et al., 1997). High affinity binding requires 5 mM Mg$^{2+}$, Mn$^{2+}$ or Co$^{2+}$, and a metal binding site was found in the theophylline binding pocket of the aptamer. The binding pocket that forms around theophylline is not created in the presence of caffeine, which only shows stacking interactions with the terminal base pair of the aptamer (Zimmermann et al., 2000).

A fluorescent sensor that can detect theophylline at a concentration of 1-4 µM, was assembled by combining the theophylline aptamer with hybridizing DNA probes labeled with fluorophore and quenchers (Rankin et al., 2006). The theophylline aptamer was also used to make the first electrochemical sensor
based on an RNA aptamer. This sensor has a dynamic range of 200 nM to 10 µM of theophylline, and shows no interference from 1 mM caffeine (Ferapontova et al., 2008).

The theophylline aptamer has been used to make several artificial RNA constructs that change conformation in response to theophylline, and can be used to regulate various intracellular processes in bacteria, yeast and cultured human cells. An artificial riboswitch was inserted into an mRNA to regulate translation in *Bacillus subtilis*. Theophylline (6 mM) was used to induce switching between different stem loop structures adjacent to the ribosomal binding site of the mRNA, which in turn reduces the level of translation of the mRNA construct by 8-fold (Suess et al., 2004).

Another group applied the theophylline aptamer to make various trans-acting regulatory modules (riboregulators). The riboregulators were tested for their ability to reversibly bind to a target mRNA and inhibit translation in yeast (*Saccharomyces cerevisiae*). These riboregulators contain an antisense sequence that is complementary to a target mRNA, another sequence that can bind and sequester the antisense sequence, and an RNA aptamer that responds to a target molecule with a conformational change. The antisense sequence is sequestered in the free RNA and released by binding of theophylline to its aptamer. The most sensitive construct inhibits translation at a theophylline concentration around 100 µM (Bayer and Smolke, 2005).

Aptamer controlled regulation of RNAi (RNA interference) in HEK 293 (human embryonic kidney) cells was also demonstrated using the theophylline
aptamer. Short hairpin RNAs (shRNAs) were constructed with the loop region replaced by the theophylline aptamer. Theophylline was used to regulate shRNA processing by DICER and consequently RNAi activity. Inhibition of RNAi was observed first with 100 µM and increased with theophylline concentration up to 10 mM (An et al., 2006).

The IMAGEtag reporter system

The proposed IMAGEtag (intracellular multi aptamer genetic tag) will consist of a string of multiple tandem aptamers that can be expressed in the cells as RNA. The IMAGEtag will not encode any protein, but could be used to make fusion construct with endogenous mRNAs so that these can be localized and tracked in the cell. Detection of the IMAGEtags relies on the aptamers’ binding to a ligand molecule that is labeled with a fluorescent moiety or radioactive element, which would be added to the cells before analysis. To increase the binding capacity and the potential signal of such tags, constructs with several aptamers have been created by employing a cloning technique based on methods previously used for creating repetitive polypeptides (Meyer and Chilkoti, 2002; Won and Barron, 2002).

Collaborator contributions

The cloning method presented in Chapter 2 was developed by me based on previously published methods (Meyer and Chilkoti, 2002; Won and Barron, 2002) and utilized by me and my colleagues to clone other multiaptamer constructs as part of the Nilsen-Hamilton laboratory IMAGEtag project. Relevant
data generated from these collaborations is included in Chapter 2 of this thesis. Tobramycin multiaptamers were prepared by Ilchung Shin, and neomycin, kanamycin and alternating neomycin-tobramycin multiaptamers were created by Muslum Ilgu.

GK38, a conjugate of theophylline and Rhodamine-B utilized for several experiments presented in chapter 3, was synthesized by the graduate research assistant Vinayak Gupta in the lab of Dr. George Kraus (Department of Chemistry, Iowa State University).
CHAPTER 2. CLONING OF MULTIPLE TANDEM REPEATS OF PROTEIN DOMAINS AND APTAMERS

A manuscript prepared for submission to the journal *BioTechniques*

Hans Eirik Haarberg, Muslum Ilgu, Ilchung Shin, Marit Nilsen-Hamilton

**Abstract**

Constructing and cloning repetitive DNA sequences can be a technical challenge. Yet, tandemly repeated sequences have been employed to increase the binding capacities of DNA segments and protein domains and have applications in studies of the structure and function of nucleic acids and protein polymers. The recursive directional ligation and PCR-based methods of assembling DNA repetitive sequences that have been developed to date have limited application because they are technically complex, time consuming, cost inefficient, and/or specifically designed for each DNA sequence to be multiplied. Here we describe a simplified method for efficiently and cost effectively generating tandem repeats of a variety of sequences and lengths of DNA that can be used without modification for each application. Whereas other reported methods for generating repeated DNA sequences require significant modification to generate tandem repeats of different sequences of interest, our method is broadly applicable, without modification, to most DNA sequences. In this method, the restriction enzymes *BsmAI* and *Bsal* are utilized to generate tandem repeats
of a variety of DNA sequences. Here we give examples of the generation of repeated aptamer and polypeptide sequences.

**Introduction**

There has been an interest in generation of tandem repetitive DNA sequences for several applications. Repetitive sequences have been used to study the structure and function of DNA (Tchernaenko et al., 2003). Tandem repeats of transcriptional regulatory elements have been used in affinity chromatography for purification of DNA binding proteins (Kadonaga and Tjian, 1986) and to study the activity of transcription factors (Veldman et al., 1985; Schafer and Fournier, 1992). Other DNA constructs have been prepared for expressing repetitive protein domains. Tandem repeats have been used to study functional protein domains such as the Ras-binding domain of C-Raf (Augsten et al., 2006). Many synthetic repetitive polypeptides have properties similar to natural proteins such as extracellular matrix proteins (Welsh and Tirrell, 2000) and spider web proteins (Zhou et al., 2001). An increase in a need for repetitive polypeptides, to be used as components of the extracellular matrix, is expected with further advances in tissue replacement therapy (Kim and Mooney, 1998).

To make constructs for expressing repetitive polypeptide sequences, Meyer and Chilkoti developed the recursive directional ligation (RDL) technique (Meyer and Chilkoti, 2002). RDL allows a fragment contained in a plasmid within two selected restriction enzyme sites to be multiplied to create longer repetitive sequences, and can be repeated to generate constructs of increasing length. The
method utilizes restriction enzymes with split palindromic recognition sequences. This makes it possible to perform seamless multiplication of a fragment by excising the fragment with both enzymes, and ligating it back into the same vector that has been digested with only one of these enzymes. However, with this method, part of the recognition sequence of each enzyme is within the sequence that is being multiplied. Consequently, different restriction enzymes must be used for multiplying different sequences.

Won and Barron developed a method for making repetitive sequences that encoded polypeptides, without DNA sequence restrictions, using a pair of type IIS restriction enzymes that cleave the DNA at one side of a recognition sequence that is asymmetrical (Won and Barron, 2002). However, the strategy required PCR and numerous enzymatic reactions.

Our modified RDL technique utilizes two non palindromic restriction enzymes, Bsal and BsmAI, to make constructs with repetitive sequences. The method is time-efficient, reliable and can be used for most DNA sequences without adjustments. Following this modified RDL method, we created DNA sequences encoding a protein with repeats of the 25 amino acid C-terminal domain of the bacterial magnetosome forming protein Mms6 (Arakaki et al., 2003), and RNAs consisting of multiple copies of the aptamers for theophylline (Jenison et al., 1994), tobramycin (Jiang et al., 1997), kanamycin (Werstuck and Green, 1998) and neomycin (Jiang et al., 1999).
Materials and Methods

Materials

The cloning vector pZErO-2 and Vector NTI Software were from Invitrogen (Carlsbad, CA, USA). E. coli strain NEB 5-alpha and restriction endonucleases AflIII, Apol, Bsal and BsmAI were obtained from New England Biolabs (Ipswich, MA, USA). Calf intestinal alkaline phosphatase (CIAP), T4 polynucleotide kinase and T4 DNA ligase were purchased from Promega (Madison, WI, USA). QIAquick Gel Extraction Kit and QIAprep Spin Miniprep Kit were obtained from QIAGen (Valencia, CA, USA). Sequencing primers 2819 (5’-GGCCT TTTTA CGGTT CCTGG-3’) and 2820 (5’-GCCGC TCCCG ATTCG CAG-3’), and all oligonucleotides used to make inserts were synthesized by the Iowa State University DNA Facility (DSSF). All sequencing was done by the DSSF.

Cloning of initial repetitive sequences from synthetic oligonucleotides

The vector for cloning the repetitive sequences was created by digesting the pZErO-2 vector with restriction enzymes AflIII and Apol. The fragment of 1854 base pairs was isolated using gel electrophoresis, and purified with QIAquick gel extraction kit. Two pairs of complementary oligonucleotides were used to create inserts containing between two and seven repeats (150-250bp) of the monomer sequences (Table 1), flanked by restriction sites for Bsal and BsmAI for subsequent multiplication. Complementary oligonucleotides were annealed by slow cooling from 95°C to 25°C and phosphorylated using the T4 polynucleotide kinase. The double stranded inserts were cloned into the vector with T4 DNA
ligase and transformed into *E. coli* strain NEB 5-alpha by heat shock (1 min, 42°C).

**Construction of longer repetitive sequences**

The initial repetitive sequences were cut out of the vector using *Bsm*AI, and purified using agarose gel electrophoresis. A separate fraction of the vector was cut with *Bsa*I, which cuts at a single site at the 5’ end of the multiaptamer, leaving the aptamers in the vector. The *Bsa*I digested vector was then dephosphorylated, and the multiaptamer fragment that had been excised with *Bsm*AI was cloned into the dephosphorylated vector (Fig. 2.1). Clones were analyzed for insertions by digestion with *Bsm*AI, and sequenced in each direction with primers 2819 and 2820.

**Results and Discussion**

**Repetitive aptamer sequences constructed using oligonucleotides**

Digestion of the pZErO-2 vector with *Afl*III and *Apo*I removed parts of the original pZErO-2 vector that was considered unnecessary for the purpose of this cloning method. Oligonucleotides were used to create five repeats of the theophylline aptamer. Analysis of eight clones was done by digestion with *Bsm*AI followed by agarose gel electrophoresis. Four constructs appeared to have an insert, but the sizes were different and only one appeared to have five repeats as expected. Sequencing showed that the clones contained 3, 4, 5 and 6 aptamer repeats.
Plasmid pHEH14 (Fig. 2.2) containing 6 repeats of the theophylline aptamer was used for further multiplication to make longer repeating sequences.

**Multiplication of the original repetitive sequences**

The multiplication of aptamer sequence repeats was achieved by utilizing the restriction enzymes *BsaI* and *BsmAI*. These restriction enzymes are unusual because their recognition sequences are not palindromic, and the restriction enzymes cut the DNA outside of their recognition sequence. Consequently, the restriction enzyme sites remain unaltered after digestion, and four-base overhangs are created from the adjacent base pairs. The sequence between the restriction sites is designed so that four of the bases at either end of the fragment are repeated at the other end in order to create compatible overhangs after digestion (Fig. 2.2). Additionally, the recognition sequence for *BsmAI* is also present within the sequence for *BsaI*, so that any sequence cut by *BsaI* is also cut by *BsmAI* but not *vice versa*. With this construct, the fragment between the two restriction enzyme sites can be cut out from the vector with *BsmAI*, while digestion with *BsaI* only cuts the vector open at one end of the fragment. The fragment that was cut out with *BsmAI* can be ligated back into the same vector but that has been digested with *BsaI*, resulting in multiplication of the number of repeats.

This procedure was first used to multiply the 6 repeats of the theophylline aptamers in vector pHEH14. Twenty colonies were prepared and analyzed by digestion with *BsmAI* and agarose gel electrophoresis (Fig. 2.3). From the sizes
of the fragments, 14 of the 20 clones appeared to have 12 repeats (Size 444bp) of the aptamer. One of these clones was sequenced and the presence of 12 aptamers was confirmed. One clone had a larger fragment, and was found by sequencing to contain 18 repeats. Two clones had smaller fragments, one of which was sequenced and determined to contain 9 repeats. The constructs with 12 and 18 repeats were expected, and result from incorporation of one or two 6-aptamer fragments respectively. The presence of 9 repeats can only have resulted if several aptamer sequences were removed from the construct during cloning. pHEH15 containing 18 repeats was the longest repetitive sequence obtained and this clone was used in the next round of multiplication.

The same method was used in the second round of cloning to multiply the 18 repeats contained in vector pHEH15. The result was three clones with fragments of sizes corresponding to approximately 270, 108 and 90 repeats (Fig. 2.4). These sizes correspond to 15, 6 and 5 fragments of the original 18 repeats respectively. The fourth clone appears to have two different fragments and possibly contains two distinct constructs.

The exact number of aptamers in these constructs is hard to determine from the estimated size of the fragments on the gel. Sequences of the clones from each side of the construct confirmed the correct sequence of the adjacent restriction sites and approximately 30 repeats of the multiaptamers at each end. Sequencing was however limited to about 1000-1200 bases, which corresponds to approximately 30 aptamers. Because the same sequence is repeated
throughout the construct it is impossible to design any unique primers for sequencing within the fragment.

**Other repetitive constructs**

Including the cloning of theophylline multiaptamers, this method was used to make similar constructs of other sequences. The technique has proven successful for cloning multiple repeats of the sequence of aptamers for kanamycin, neomycin and tobramycin, as well as a sequence encoding the 25 amino acid C-terminus of the Mms6 protein (Table 2).

**Stability of repetitive constructs**

Secondary structure of DNA, particularly in the form of direct repeats and palindromic sequences, can destabilize the DNA during replication and result in deletions, insertions and rearrangements (Bi and Liu, 1994; Rosche et al., 1995; Bi and Liu, 1996). This feature of bacterial DNA synthesis and repair is probably the basis of the recurrent observations with several aptamer constructs that the number of repeats obtained did not match a multiple of the number of repeats in the fragments used for the cloning. In the case of the theophylline, tobramycin, neomycin and kanamycin aptamers, the clones obtained always included complete aptamer sequences with no partial aptamer sequences. This result might be due to the strong secondary stem loop structures of these aptamers that probably result in the formation of cruciform DNA and recombination or DNA repair involving breakage at one or the other side of the cruciform. Except for the longest repeated sequence (270 tandem theophylline aptamers) all other stem-
loop aptamer structures cloned by this method were stable as demonstrated by successful cloning of the repetitive sequences into other vectors. The correct number of repeats were confirmed by restriction digestion and sequencing.

By contrast with the theophylline and aminoglycoside aptamers, the tetracycline aptamer forms a more complex structure with several stems and loops. Only fragments of two aptamers were obtained when cloning of oligonucleotide repeats of the tetracycline aptamer was attempted. Repeats of the length represented by the tetracycline aptamer are deleted by a recA-independent mechanism, which in pBR322 can be influenced by another D7 region within the plasmid acting in cis (Bi and Liu, 1996). Thus, the stability of clones containing repeated sequences is restricted by the bacterial replication and recombination systems and the cis-acting features in the plasmid. For this reason it is desirable, as we have done, to use a small (stripped down) plasmid for cloning the repetitive sequences so as to minimize the presence of cis-acting features that promote recA-independent recombination.

Summary

In summary, we demonstrate a method that can be used to clone repetitive DNA sequences of a variety of forms, some with strong internal structures (aptamers) and others without stable internal structures (mature Mms6 coding sequence). The method is time and cost-effective and results in plasmids containing multiple repeated sequences can be stably maintained.
Acknowledgements

This work was supported by a grant to MNH from the National Institutes of Health (R01HL078659). We thank Lee Bendickson for technical support and advice.

Competing interests

MNH is owner and President of Molecular Express Inc., Iowa

Tables

Table 1. Monomer sequences

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theophylline aptamer*</td>
<td>GGTGATACCAGCGAAGGGCCTTGGCCAGCACC</td>
</tr>
<tr>
<td>Tetracycline Aptamer*</td>
<td>GGCCTAAAACATACCAGAGAAATCTGGAGAGTGGAAGAT ACACCACCTAGGCC</td>
</tr>
<tr>
<td>Mms6-C25</td>
<td>TACGCGTATATGAAGAGCCGTCGATATCGAATCGGCAGA</td>
</tr>
<tr>
<td></td>
<td>GCGACGAGGAAACTCGGACGCACGCCTGGCC</td>
</tr>
<tr>
<td>Kanamycin aptamer*</td>
<td>GCGCAGUGGCCAUAGAACCAUGCC</td>
</tr>
<tr>
<td>Tobramycin aptamer*</td>
<td>GGCACGAGGUUAGCUAACACUCGUGCC</td>
</tr>
<tr>
<td>Neomycin aptamer*</td>
<td>GGACUGGGCGAGAUUAGUAGUAGUAGUCC</td>
</tr>
</tbody>
</table>

* Each repeat of the multiple aptamers also contained a linker of four As.
Table 2. Constructs with repetitive sequences

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Source (monomer repeats)</th>
<th>Number of repeats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theophylline aptamer</td>
<td>Oligonucleotides (5)</td>
<td>3*, 4*, 5, 6*</td>
</tr>
<tr>
<td></td>
<td>Vector (6)</td>
<td>9*, 12, 18</td>
</tr>
<tr>
<td></td>
<td>Vector (18)</td>
<td>90, 108, 270</td>
</tr>
<tr>
<td>mms6-C25</td>
<td>Oligonucleotides (2)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Vector (2)</td>
<td>4, 6, 8</td>
</tr>
<tr>
<td>Tetracycline aptamer</td>
<td>Oligonucleotides (3)</td>
<td>2 partial aptamers</td>
</tr>
<tr>
<td>Tobramycin aptamer</td>
<td>Oligonucleotides</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Vector (5)</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Vector (10)</td>
<td>14*</td>
</tr>
<tr>
<td>Neomycin aptamer</td>
<td>Oligonucleotides (6)</td>
<td>6, 8*</td>
</tr>
<tr>
<td></td>
<td>Vector (8)</td>
<td>16, 23*, 24, 32</td>
</tr>
<tr>
<td>Kanamycin aptamer</td>
<td>Oligonucleotides (7)</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Vector (7)</td>
<td>9*, 12*, 14, 28</td>
</tr>
<tr>
<td>Alternating</td>
<td>Oligonucleotides (4)</td>
<td>4</td>
</tr>
<tr>
<td>Alternating Neomycin-Tobramycin aptamers</td>
<td>Oligonucleotides (4)</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Vector (4)</td>
<td>8</td>
</tr>
</tbody>
</table>

* The number of repeats obtained was not expected from the source used.
Legends to Figures

Figure 2.1. Multiplication of repetitive sequences. A: The double stranded recognition sequences (highlighted) and location of cut sites (arrows) for *Bsa*I and *Bsm*AI. N indicates any base (A, C, G or T) B: Outline of the strategy for multiaptamer cloning. The cloning vector contains two recognition sequences for *Bsm*AI, one on each side of the multiaptamer sequence. The multiaptamer sequence is cut out of the vector with *Bsm*AI (1), and purified by agarose gel electrophoresis (2). One of the *Bsm*AI sites contains an additional GC base pair and is recognized by *Bsa*I, while the other site contains a TA base pair and is therefore protected from cleavage by *Bsa*I. Digestion with *Bsa*I cuts open the vector, while leaving the original multiaptamers in the vector (3). The digested vector is then dephosphorylated (4) to prevent self ligation of the vector in the final step. The purified multiaptamer is ligated back into the single cut vector to create several copies of the multiaptamer sequence (5). Several multiaptamer fragments can potentially be ligated into the vector. Because the restriction enzyme sites are unaffected, this method can be repeated to create constructs with increasing numbers of repeats.

Figure 2.2. Plasmid *pHEH14* containing 6 repeats of the theophylline aptamer. The vector contains only the origin of replication and the kanamycin resistance gene of the original pZErO-2 vector. Tandem repeats of the theophylline aptamer are flanked by the restriction sites for multiplying the repetitive sequence (*Bsa*I and *Bsm*AI), and primers 2819 and 2820 for sequencing from each side of the construct.
Figure 2.3. Agarose gel electrophoresis of clones from the first multiaptamer cloning. M: DNA marker (from top: 1500, 850, 400, 200 and 100bp). 1-20: clones digested with BsmAI. Each clone has two bands from the vector (840 and 1049bp), in addition to a band from the multiaptamer. Clones 1, 4 and 13 were sequenced and determined to contain 12, 9 and 18 repeats respectively.

Figure 2.4. Agarose gel electrophoresis of clones from the second multiaptamer cloning. M: DNA marker (from top: 10, 8, 6, 5, 4, 3.5, 3, 2.5, 2 and 1.5kb). 1-4: Clones digested with BsaI and Apol. Each clone has a band from the vector (1872bp), and a band from the multiaptamer. Clones 1, 2 and 3 are estimated to contain 270 (9990bp), 108 (3990bp) and 90 (3330bp) aptamer repeats respectively.
Figures

A

BsaI: 5’-GGTCTCNNNNNN 3’-CCAGAGNNNNNN

B

BsmAI: 5’-GTCTCNNNNN 3’-CAGAGNNNNNN

Cloning vector with multiaptamer flanked by BsaI and BsmAI sites (bold)

1. Excise multiaptamer from vector with BsmAI

2. Gel purify multiaptamer

3. Cut vector adjacent to multiaptamer with BsaI

4. Dephosphorylate vector to prevent ligation of vector without insert

5. Ligate additional copy of multiaptamer into vector

(Figure 2.1)
(Figure 2.2)
CHAPTER 3. STUDY OF THE THEOPHYLLINE IMAGETAG AND LIGANDS

Abstract

An efficient RNA reporter is expected to have several significant advantages for studying gene expression and RNA, and several potential reporters have been developed utilizing RNA aptamers. However, no such reporter has been used successfully to study RNA expressed in living cells. The theophylline multiaptamer and fluorescent and radioactively labeled theophylline were studied in cultured CCL-64 cells in an attempt to develop a cellular RNA reporter system. The multiaptamer RNA was expressed from a plasmid containing the β-actin promoter and a cytomegalovirus (CMV) enhancer by transient transfection and RNA levels were determined using real time reverse transcription PCR. The interactions of the aptamer ligands with the cells were studied using both fluorescent and radioactive labeled theophylline. The theophylline multiaptamer RNA was expressed and shown to have high stability in the CCL-64 cells. The ligands successfully entered the cells, but bound to endogenous cellular components and over 50 percent of the ligands taken up by the cells remained inside the cells after one hour. The binding of ligand to the cells created a high background signal that was independent of aptamer expression.
Introduction

A successful IMAGEtag would offer several advantages over current protein reporters including significant reductions in both energy requirement and response time, and would also allow endogenous RNAs to be tagged to investigate RNA localization and transport.

Several groups have attempted to develop aptamer based RNA reporters. These attempts include selecting aptamers that can bind and sequester fluorophores from solution (Holeman et al., 1998) and aptamers that can increase the fluorescence of target molecules (Babendure et al., 2003); (Sparano and Koide, 2007; Constantin et al., 2008; Sando et al., 2008). Although several candidate reporters have exhibited promising results in in vitro assays, no aptamer tag has so far been used successfully in vivo or in cultured cells.

The reasons for the limitations of previous RNA reporters are not clear, but it is apparent that several conditions have to be satisfied for an aptamer based RNA tag to be successful in living cells. The strength of the signal depends on both the intracellular concentration of the RNA aptamer itself, and that a high proportion of the aptamers bind to the ligand. Thus, the aptamers should be efficiently transcribed, stable and have a high affinity for the ligand, while the ligand must efficiently traverse the cell membrane to interact with the aptamer. In addition, the ligand should not bind to cellular components because this would create a high background signal that would be present in all cells and compete with the specific signal from aptamer bound ligands.
The theophylline multiaptamer transcripts and fluorescent and radioactive ligands were studied in cell culture to determine if they satisfy the requirements for a functional IMAGEtag. The mink (*Mustela vison*) lung epithelial cell line CCL-64 from ATCC was used for all cell culture studies. This cell line has been used by other members of the Nilsen-Hamilton lab and was chosen for these experiments because of a particularly high transfection efficiency attained with these cells.

The IMAGEtag RNA was expressed in CCL-64 cells from a mammalian expression vector containing a β-actin promoter and a cytomegalovirus (CMV) enhancer. The expression vector was introduced into the cells by lipofection, and real time reverse transcription PCR was utilized to determine the level of IMAGEtag RNA present in the cells. Time course studies of RNA were performed to determine the time of maximal expression and the stability of the RNA.

The cellular uptake, efflux and retention were investigated for both fluorescent and radioactive labeled theophylline. Additionally, the location of the fluorescent labeled theophylline was determined using fluorescent confocal microscopy.

**Materials and Methods**

**Plasmids and Reagents**

The expression vector with CMV enhancer and β-actin promoter: pKS35-CMV-UTC was obtained from C. Tuggle, Department of Animal Science, Iowa State University). The expression vector with eGFP (enhanced green fluorescent
protein): peGFP was purchased from Clontech (Palo Alto, CA). SalI and DNase (RNase free) were from New England Biolabs (Ipswich, MA). HindIII, T4 polynucleotide kinase and T4 DNA ligase were from Promega (Madison, WI). Cordycepin, DMEM (high glucose), rhodamine-B base, theophylline, theophylline-[8-3H], Nonidet P-40 substitute and Hoechst 33250 were from Sigma-Aldrich (St. Louis, MO). Penicillin-streptomycin was from Gibco BRL (Grand Island, NY). YO-PRO-1, lipofectamine 2000 and TRizol reagent were from Invitrogen (Carlsbad, CA). Full Velocity and Brilliant II qRT-PCR kits were from Stratagene (La Jolla, CA). HEPES (enzyme grade) and Scintiverse BD Cocktail were from Fisher Scientific (Pittsburgh, PA). Agarose was from ISC BioExpress (Kaysville, UT). Bovine calf serum was from Hyclone (Logan, UT). NanoDrop-1000 spectrophotometer was from NanoDrop Technologies (Wilmington, DE). All sequencing was done by the Iowa State University DNA Facility (DSSF). The following oligonucleotides for use as primers or inserts were synthesized by the DSSF: 144 (5’-CCAAC CGACT GCTGT CA C-3’), 224 (5’-GCTTA ATACG ACTCA CTATA GGGCG GCTAT TCTCG CAGGA-3’), 1902 (5’-AATGC CAGCC CCAGC ATCGA AG-3’), 1905 (5’-CCCAC CCCCA ATGTG TCTGT-3’), 2823 (5’-TCGAG CGGCC GCTCG AGGAT CCTAA TACGA CTCAC TATAG GGAGA CAAGA ATACG CTCAA GTCGA CGCCG AAAAG CCG AA GCTTC GACAG GAGGC TCACA ACAGG CTAGC-3’), 2824 (5’-AGCTG CTAGC CTGTT GTGAG CCTCC TGTCG AAGCT TCGGC TTTTC GGCGT CGA CT TGAGC GTATT CTTGT CTCCC TATAG TGAGT CGTAT TAGGA TCCTC
GAGCG GCCGC -3'), 2837 (5'-GGGGT GGCTT TTAGG ATGGC AA-3'), 2839 (5'-CGACA GGAGG CTCAC AACAG GCTA-3').

GK38 is a conjugate of theophylline and Rhodamine-B (Fig. 3.1). The synthesis of GK38 was performed by the graduate research assistant Vinayak Gupta in the lab of Dr. George Kraus (Department of Chemistry, Iowa State University).

**Cloning of IMAGEtags into expression vectors**

An expression vector for the insertion of theophylline and other multiaptamers was constructed from vector pKS35-CMV-UTC. A fragment containing lipocalin 2 cDNA was removed from the vector by sequential digestion with *Sal*I and *Hind*III and the vector was purified by agarose gel electrophoresis. An insert containing a multiple cloning site for insertion of multiaptamer constructs, flanked by primer binding sites, was constructed from oligonucleotides 2823 and 2824. The oligonucleotides were annealed to create a double stranded insert with *Sal*I and *Hind*III overhangs. The annealed oligonucleotides were phosphorylated with T4 polynucleotide kinase and cloned into the digested vector with T4 DNA Ligase. Ligation mixtures were transformed into E. coli by heat shock (42°C for 1 min),
and clones were sequenced with primers 144 and 224. The resulting expression vector was named pHEH18.

Vector pHEH18 was digested with *Hind*III, followed by precipitation with 2-Propanol and digestion with *Sal*I. Multiaptamer constructs were excised from cloning vectors pHEH15, pHEH16.3 and pHEH16.1, by sequential digestion with *Sal*I and *Hind*III. Digested expression vector and multiaptamer fragments were purified by gel electrophoresis with 0.5% agarose. The vector was ligated with each of the multiaptamer fragments, transformed and sequenced as described previously.

**Assay of theophylline toxicity in CCL64 cells**

CCL64 (Mink epithelial cell line) cells were incubated in medium composed of 90% DMEM high glucose, 10% bovine calf serum and 10 u/mL penicillin/streptomycin for 24 h before treatment. The medium on the cells was replaced with new medium with or without theophylline (2.5 µM, 10 µM, 50 µM, 250 µM, 1 mM, 5 mM or 10 mM). After 24 h of incubation with theophylline, the cells were stained with YO-PRO-1 and Hoechst 33250. Staining was observed by fluorescent microscopy (Zeiss Axiovert 40 CLF), using DAPI and FITC filters, and pictures were taken with each filter.

**Expression of IMAGEtags in CCL64 cells and analysis by real time RT-PCR**

Cells were transfected with plasmids pHEH21.2 (multiaptamer) or peGFP (enhanced green fluorescent protein) using lipofectamine 2000. The plasmid and lipofectamine were diluted and combined in serum free DMEM for complex
formation. Transfection was performed with 2 ng/µL plasmid DNA and 4 ng/µL lipofectamine in DMEM, 8% bovine calf serum. No antibiotic was used during these experiments.

RNA samples were prepared by dissolving the cells using TRIzol reagent and pipeting up and down approximately 10 times, followed by chloroform extraction and 2-propanol precipitation. The RNA yield was determined with a NanoDrop 1000, and samples were treated with RNase free DNase. RNA was quantified by one-step reverse transcription real time PCR using the Full Velocity qRT-PCR kit. IMAGEtag transcripts were PCR amplified with primers 2837 and 2839. The mink GAPDH gene was used as a reference, and PCR was performed with primers 1902 and 1905.

**Uptake and efflux of rhodamine-B and GK38**

Confluent CCL64 cells were washed and incubated for 30-40 min with Hanks balanced salt solution (HBSS; pH 7.4, 137 mM NaCl, 5.4 mM KCl, 1.3 mM CaCl₂, 0.5mM MgCl₂, 0.6 mM MgSO₄, 4.2 mM NaHCO₃, 0.4 mM KH₂PO₄, 0.3 mM Na₂HPO₄). For measuring uptake, the cells were incubated with HBSS containing either rhodamine-B or GK38. After specified periods of time the dye solutions were removed, the cells were washed two times with HBSS, and the cells were harvested by scraping into lysis buffer (pH 7.2, 20 mM HEPES, 0.5% Nonidet P-40)

For efflux measurement, the cells were incubated in HBSS containing 1.8 µM of either dye (rhodamine-B or GK38) for 30 min. The cells were then rinsed
twice with HBSS, and incubated in HBSS for efflux of the dye. After the designated periods of efflux the HBSS was removed and cells were scraped into lysis buffer. For both uptake and efflux experiments, the amount of dye in the cells was determined by measuring fluorescence with a Cary Eclipse fluorometer (Varian). Samples were excited at 550 nm and emission was measured at 578.

**Determining the intracellular localization of GK38 by confocal microscopy**

CCL64 cells were incubated in HBSS with 10 µM GK38 at 37°C for 1 h. The cells were rinsed once and incubated in HBSS without GK38 for three hours before analysis. Confocal microscopy was performed at 60x magnification with a Nikon Eclipse TE200 microscope at the Roy J. Carver lab for ultrahigh resolution biological microscopy. GK38 was excited with a 568 nm Kr/Ar laser and fluorescence was detected at 580-620 nm through a TRITC filter. A stack of 29 slices with 0.5 µm thickness was collected.

**Efflux of theophylline-[8-³H]**

CCL64 cells were grown on cover slips and incubated at 37°C for 1 h in HBSS with 1.1 µM tritium labeled theophylline (activity: 2 µCi/ml). The cells were then washed twice with HBSS, and new HBSS solution was added for the efflux of theophylline. Following the specified durations of efflux, the HBSS was removed and the cells washed once with HBSS. Cover slips were collected and put in scintillation vials with Scintiverse BD Cocktail. Tritium activity was measured with a scintillation counter for 5 min per sample.
Quantification of theophylline-[8-^3^H] in cells with and without IMAGEtags

To determine the effect of aptamer expression on the cellular theophylline concentration, cells were transfected with either plasmids pHEH21.2 (multiaptamer) or peGFP using lipofectamine 2000. Transfection was performed with 2 ng/µL plasmid DNA and 4 ng/µL lipofectamine in DMEM, 8% bovine calf serum.

Cells were rinsed with HBSS 24 h after transfection, and incubated with either 100 nM, 0.5 µM, 2 µM or 10µM theophylline (2 µCi/ml activity each). After incubation with radiolabeled theophylline the cells were rinsed three times with HBSS, and collected using lysis buffer (pH 7.2, 20 mM HEPES, 0.5% Nonidet P-40). The radioactivity level was determined using Scintiverse BD Cocktail, scintillation counter for 4 min per sample.

Determination of the stability of IMAGEtag RNA

Cells were transfected with multiaptamer plasmid pHEH21.2 using lipofectamine 2000 (Invitrogen). Plasmid and lipofectamine were diluted and combined in serum-free DMEM for complex formation. Transfection was performed with 2 ng/µL plasmid DNA and 4 ng/µL lipofectamine in 92% DMEM supplemented with 8% bovine calf serum.

RNA synthesis was inhibited with 50 µg/µL cordycepin 36 h after transfection. The initial RNA sample (time 0) was collected 30 min after addition of cordycepin. Cells were collected with TRlzol reagent, followed by chloroform extraction and 2-propanol precipitation. RNA yield was determined with a
NanoDrop 2000 (Thermo Scientific), and samples were treated with RNase free DNase. RNA was quantified by one-step reverse transcription RT PCR using BrilliantII qRT-PCR kit. IMAGEtages transcripts were PCR amplified with primers 2837 and 2839. The mink GAPDH gene was amplified for comparison with primers 1902 and 1905.

Results and Discussion

Cloning of IMAGEtags into expression vector

Two successful multiaptamer expression vectors were created. Vector pHEH21.2 was produced from the multiaptamer fragment of pHEH15, and contained 18 repeats of the theophylline aptamer. Vector pHEH21.5 was cloned from the pHEH16.3 containing about 90 repeats. The third cloning vector used (pHEH16.1) was expected to contain 270 repeats, since the original cloning produced a fragment of ~10kb (Chapter 2). However, digestion with SalI and HindIII produced a fragment with a size of approximately 1kb. This fragment was also purified and ligated with the expression vector, but no colonies were produced after transformation.

Toxicity of theophylline in CCL64 cells

The toxicity of theophylline in CCL64 cells was investigated using the blue fluorescent dye YO-PRO-1. Apoptotic cells become permeable to YO-PRO-1, while normal cells do not. Due to this property, YO-PRO-1 selectively stains the
nucleus of apoptotic cells (Idziorek et al., 1995). The cells were also stained with Hoechst-33250, which stains the nuclei of all cells.

The cells were incubated in their regular medium with theophylline concentrations of 2.5 µM, 10 µM, 50 µM, 250 µM, 1 mM, 5 mM and 10 mM for 24 hours before assay with YO-PRO-1 and Hoechst-33250. No cells were stained by YO-PRO-1 even at the highest theophylline concentrations (Fig. 3.2).

It was not expected that theophylline would be toxic at the normal therapeutic plasma concentration range of 55-110 µM (Barnes and Pauwels, 1994), and the results from this experiment indicate that theophylline is not toxic to CCL64 cells at concentrations up to 10 mM.
Expression of IMAGEtags in CCL64 cells

The amount of transcribed IMAGEtag RNA in the cells after transient transfection was determined by reverse transcription and real time PCR. Transcription of the IMAGEtag RNA (18 tandem theophylline aptamers) begins within one hour of transfection with the expression vector pHEH21.2. The level of IMAGEtag transcripts is relatively low until 12 h, but increases exponentially until 24 h after transfection (Fig. 3.3).

![Graph showing the level of IMAGEtag RNA over time after transfection.](image)

**Figure 3.3.** Relative amount of IMAGE-tag RNA in CCL64 cells from 2 to 24 h after transfection with IMAGE-tag (18 repeat theophylline aptamer) expression vector pHEH21.2. Average of three samples with standard deviation error bars.

The RNA reaches a maximum level approximately 24 h after transfection. There is a rapid increase in IMAGEtag RNA level from 12 to 24 h, followed by a steady decrease from 24 to 72 h (Fig. 3.4).
To test the time course of the uptake of the dyes, GK38 and rhodamine-B, confluent CCL64 cells were incubated with the dye for 10 to 50 minutes. After incubation, the cells were washed three times to remove extracellular dye, prior to cell collection with lysis buffer. To determine the amount of dye present in the cells, the fluorescence of the lysis buffer was measured. Both dyes enter the cells rapidly and both GK38 and rhodamine-B reach a steady intracellular level within 10 to 15 minutes (Fig. 3.5).
The efflux of GK38 from CCL64 cells was investigated by incubating the cells with GK38 for 30 minutes, washing three times, and incubating in buffer without dye for 10 to 60 min. The amount of dye remaining in the cells was determined by measuring the fluorescence of the lysed cells at 578 nm emission; 550 nm excitation. GK38 decreased to 60-70% at 20 minutes and started to level off after this. At 60 minutes, the fluorescence was just below 60% and it is likely that a majority of the dye would remain in the cell for several hours (Fig. 3.6).

In a comparison of rhodamine-B and GK38, rhodamine-B had a slightly higher rate of efflux than GK38. After 1 h, the fluorescence of GK38 was decreased to 55-60%, while the fluorescence of Rhodamine-B was decreased to 30-40%. GK38 efflux also appeared to slow down after 20-30 min, while the efflux of rhodamine-B was nearly linear during the 60 minutes (Fig. 3.7).
Figure 3.6. Fluorescence of GK38 remaining in cells after incubation in buffer without dye for 10 to 60 minutes. Average of three samples with standard deviation error bars.

Figure 3.7. Comparison of rhodamine-B and GK38 remaining in cells after 0, 30 and 60 minutes of efflux. Averages of three samples with standard deviation error bars.
A significant difference between the two compounds was found when comparing the fluorescence of the cell lysates and the solutions of HBSS with either dye in which the cells were incubated. The HBSS solutions were diluted 1:10 in lysis buffer and the fluorescence of these were measured along with the cell lysates. While the fluorescence of the HBSS solution with Rhodamine-B was almost twice as high as HBSS with GK38, the fluorescence of the lysates with Rhodamine-B were only a third of those with GK38. Thus GK38 was taken up by the cells at a 6-fold higher ratio than Rhodamine-B, when compared to the extracellular concentrations.

**Cellular localization of GK 38**

The conclusion from the efflux experiments was that a large fraction of the GK38 taken up by the cells remained associated with the cells. Therefore, confocal microscopy was utilized to determine its cellular location. GK38 was found localized within the cell with the highest concentration being in one area outside of the nucleus; possibly the Golgi apparatus. Pictures were obtained 3 h after removal of GK38, at which time the fluorescence was still high (Fig. 3.8).

**Efflux of theophylline-[8-^3^H]**

The efflux rate for theophylline was investigated using the method employed for rhodamine-B and GK38, with the exception that quantification was done using scintillation counting instead of fluorescence measurements. The rate of theophylline efflux was found to be similar to that of GK38 and rhodamine-B, although a larger fraction of theophylline remained in the cells than for the other
two compounds. About 60-80% of the theophylline remained in the cells after 1 h (Fig. 3.9). This may also explain why more of the GK38 remains in the cell compared Rhodamine-B, being driven by theophylline.

Figure 3.8. Fluorescence microscopy of GK38 at four vertical layers of CCL64 cells (excitation with 568 nm Kr/Ar laser and TRITC emission filter)
Effect of IMAGEtags on the intracellular concentration of theophylline-[8-3H]

CCL64 cells transformed with either pHEH21.2 (IMAGEtag) or peGFP were incubated with different concentrations of theophylline to determine if expression of the theophylline IMAGEtag affected the concentration of theophylline inside of the cells. There was no significant difference between the level of theophylline in cells expressing the theophylline multiaptamer or eGFP (Fig. 3.10). This could be because the amount of theophylline binding to endogenous cell components is too high compared to the amount of aptamers expressed. It is also possible that the aptamers were not capable of binding theophylline.
Stability and half-life of the IMAGEtag RNA in CLL64 cells

The half life of the IMAGEtag RNA was estimated by measuring the amount of IMAGEtag in the cells at three time points (with 24 h intervals) after inhibition of mRNA synthesis by cordycepin. The rate of IMAGEtag decay was slow and slightly more than half of the RNA was degraded by 48 h (Fig. 3.11). An exponential trendline was fitted to the data points, and the equation of the trendline was used to calculate a half-life of 43.3 h.

Figure 3.10. Comparison of the amount of theophylline in cells expressing the IMAGE-tag and cells expressing eGFP. Average of three samples with standard deviation error bars.
Figure 3.11. Decay of IMAGEtag RNA in CCL64 cells after inhibition of transcription. Average of six samples with standard deviation error bars. An exponential trendline is fitted to the data points, and the equation and $R^2$ for the trendline is displayed in the upper right corner.

$$y = 2E+06e^{-0.016x}$$

$R^2 = 0.9652$
CHAPTER 4. CONCLUSIONS AND FUTURE DIRECTIONS

The proposed RNA reporter system may offer several advantages over current protein reporters, such as a significant reduction in energy consumption. This system also has the potential to detect rapid changes in gene expression because the time required to transcribe RNA is minimal compared to the time required for translation and maturation of fluorescent proteins, which are currently being used as promoter reporters. The RNA reporter could also be used as a tag to track RNA, analogous to the way fluorescent proteins have been used to tag proteins. However, because many protein molecules are produced from each RNA molecule, there is normally an amplification of a signal from RNA to protein. Because this amplification is lost when RNA is used as a reporter, it is important to consider ways to maximize the signal that can be produced by an RNA reporter. To increase the capacity of the RNA reporter to bind numerous labeled ligands, multiaptamers were constructed.

The modified version of the recursive directional ligation technique was successfully used to make several multiaptamer constructs of different lengths. Six repeats of the theophylline aptamer constructed from synthetic oligonucleotides were multiplied to create 18 repeats. The 18 repeat construct was then used to create constructs with 90 and 270 repeats of the theophylline aptamer. Following this method, other researchers in the Nilsen-Hamilton lab have created similar multiaptamer constructs of different lengths from the kanamycin, tobramycin and neomycin aptamers. The method of recursive
directional ligation utilized here is novel in its application to multiaptamer construction. As opposed to the original method, the restriction enzymes employed here can be applied for multiplication of any sequence.

The multiaptamer was cloned into an expression vector with a β-actin promoter and CMV enhancer. This vector was then introduced into CCL64 cells by transient transfection, and resulted in successful expression of the multiaptamer RNA. The expression begins within a few hours after transfection, with the highest rate of expression from 12 to 24 h after transfection. The quantity of RNA reached a maximum level around 24 h, followed by a slow and steady decrease. The multiaptamer RNA is quite stable, with a half life of approximately 43 h.

In order for an aptamer to be useful as either a genetic reporter or RNA-tag, it must be able to generate an unambiguous signal. This could be achieved in two ways: first, by inducing a fluorescent signal in a ligand that has low fluorescence by itself, or second, by sequestering a ligand that is either fluorescent or labeled with a radioisotope. Because theophylline has no inducible fluorescence, the aptamer must be able to efficiently sequester a labeled theophylline. Our results do not indicate that there is any difference in the level of tritium labeled theophylline taken up by cells transfected with the theophylline multiaptamer or another vector. It is possible that the aptamer was not functional because of the structure of the RNA transcript, or due to interactions with other components in the intracellular environment.
Another obstacle to using the theophylline aptamer in an IMAGEtag is evident from the results of efflux studies of theophylline, rhodamine-B and the theophylline-rhodamine-B conjugate GK38. It appears that all three compounds bind to endogenous cell components, because a significant amount remains in the cells after 1 h. This creates a high background signal in all cells. For GK38, about 55-60% of the compound is still present in the cells after 1 h. Fluorescent imaging by confocal microscopy revealed that the GK38 retained by the cell is localized to cytoplasm, with high concentrations in one area adjacent to the nucleus (possibly in the Golgi apparatus). The intracellular binding of GK38 could be caused by the theophylline segment of the molecule, the rhodamine-B segment of the molecule, or a combination of both. Efflux studies of rhodamine-B and theophylline revealed that 30-40% of rhodamine-B and 60-80% of theophylline remained in the cells after 1 h. Thus the binding property of GK38 is intermediate of its two molecular components. Although rhodamine-B could be exchanged for another fluorescent label, it is unlikely to cause a significant reduction in binding because theophylline appears to be a larger contributor to the cellular binding. This could limit the effectiveness of using theophylline and the theophylline aptamer in developing a reporter system.

A possible solution to the high background binding may be to utilize Förster resonance energy transfer (FRET) (Förster, 1948). In FRET studies, a donor fluorophore is excited by an external source, but instead of emitting a photon, the absorbed energy is transferred to an acceptor fluorophore. To achieve efficient FRET the two fluorophores have to be compatible so that the
emission spectrum of the donor overlaps with the excitation spectrum of the acceptor. Another requirement for FRET is that the donor and acceptor molecules are in close proximity (1-10nm). For this reason, FRET is often used to determine conformational changes and associations between labeled molecules.

In regard to the multiaptamers, a pair of FRET compatible aptamer ligands could be used to differentiate between ligands bound to the multiaptamers, and ligands bound to endogenous cell component. Due to the proximity requirement, ligands binding to adjacent aptamers in a multiaptamer transcript would be close enough for FRET to occur. Ligands that bind to other cellular components however, would be dispersed in a much larger volume and a very small fraction would be close enough to engage in FRET.
LITERATURE CITED


