NUCLEIC ACID-TEMPLATED ASSEMBLY OF THERAPEUTIC AGENTS

BY

SARAH BONSON KRUEGER

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Chemistry
in the Graduate College of the
University of Illinois Urbana-Champaign, 2022

Urbana, Illinois

Doctoral Committee:

Professor Steven C. Zimmerman, Chair
Professor Douglas A. Mitchell
Associate Professor Hong Jin
Assistant Professor Lisa Olshansky

ABSTRACT

Small molecule targeting of nucleic acids has come into focus as a therapeutic strategy for diseases such as myotonic dystrophy type 1 (DM1), a trinucleotide repeat disease characterized by its RNA gain-of-function mechanism. Methods of targeting nucleic acids range from the use of antisense oligonucleotides to the use of small molecules and oligomeric compounds to modulate symptoms. Of particular interest for targeting nucleic acid repeats is the use of multivalent targeting agents. Herein, we have explored the assembly of multivalent targeting agents using the nucleic acid target as a template for *in situ* synthesis. Using proximity driven azide-alkyne click, we developed a screening approach through which several hit compounds were found. As the proximity-induced click reaction is irreversible and requires long reaction times of at least 1 day before product is observed, we set out to develop a method for rapid, template-selected, reversible assembly of therapeutic agents *in situ* via aldehyde-amine condensation. After a proof-of-concept assay with compounds that contain two reactive aldehyde or amine groups, a library of rationally designed small molecule targeting agents containing aldehyde or amine functionality was synthesized and screened. The selective assembly of fragments on template was confirmed by MALDI-MS in the presence of DM1- and Huntington's Disease (HD)-relevant nucleic acid sequences. Of interest for both DM1 and HD, the resulting hit combinations of aldehyde and amine inhibited the formation of toxic RNA *in vitro* in a cooperative manner with low micromolar IC_{50} values and rescued mis-splicing in DM1 model cells. This reversible template-selected assembly is a promising strategy to achieve multivalent targeting and could be utilized to develop strong and selective binders for other disease-relevant nucleic acid targets.

This thesis is dedicated to my husband, Russ, and my family, friends, & mentors.

"Let us persevere in running the race that lies before us, while keeping our eyes fixed on Jesus."
-Hebrews 12:1-2

ACKNOWLEDGEMENTS

The work contained in this thesis would not have been possible without the supportive community that surrounded me throughout my scientific development. The people acknowledged below made it possible for me to present this thesis toward the completion of my Ph.D. I am so blessed to have many people to thank for their mentorship, friendship, fellowship, and support throughout this work. All glory be to God!

I would first like to express my deepest gratitude to my advisor, Prof. Steven Zimmerman. I am so grateful for the opportunity to work under your mentorship in my PhD studies. Your group has provided an atmosphere for my personal and professional growth, and I am incredibly thankful for the experiences I have had in your lab. I am eternally grateful for your patience with me as I continued to grow as a scientist and when projects were not working. What makes a great mentor is one who cares about their students not just as scientists, but also as people. You support a diverse group of scientists working on a diverse range of research questions, and this group has been an excellent home for me at Illinois. I remember that before I joined your group, I told you that I wanted to return to a PUI for my career and you pledged to support me in this endeavor. You have certainly lived out this promise, and I am so grateful for your support in allowing me to earn a teaching certificate, in taking part in your preparing future faculty course, in going to Hope College for the summer, and in finding a job. Thank you for accepting me into your group and supporting me over the last five years. I truly would not have been able to get my Ph.D. without you.

I would like to thank my committee members: Prof. Doug Mitchell, Prof. Hong Jin, Prof. Lisa Olshansky, and (formerly) Prof. Paul Hergenrother and Prof. Marty Burke. I very much appreciate your advice and feedback throughout my Ph.D. studies. I want to extend a big thank you to Prof. Mitchell for your guidance on my literature seminar and your support through my job applications.

I would like to thank all of my lab mates in the Zimmerman group, past and present. I feel so privileged to have had the opportunity to work alongside a group of the most intelligent, inclusive, kind, and caring group of individuals. Each one of you has had a profound impact on my personal and professional development, and I am grateful for every one of you. For this reason, I want to recognize each of you here by name: Dr. JuYeon Lee, Dr. Julio Serrano, Dr. Brenda Andrade, Brittany Walker, Dr. Junfeng Chen, Dr. Edzna Garcia, Dr. Lauren Hagler, Dr. Ephraim Morado, Dr. Yanhua (Henry) Xu, Dr. Ke Li, Dr. Thao Mee Xiong, Anand Poozhikunnel, Jazmin Aguilar-Romero, Nafisa Ibrahim, Amie Lanzendorf, Mara Paterson, Joel Roberts, and Catherine Jalomo. No matter how long or short we overlapped in the Z-group, you all have impacted me significantly. You are my collaborators, my mentors, my colleagues, and most of all my friends. You have taught me so much throughout my years in the Z-Group. I will miss our talks about science, about DEI efforts, and about whatever is on our minds at the time. I am so lucky and blessed to have been able to work with all of you, and I cannot wait to see all of the excellent things that you will continue to do in this world.

I also want to extend a special thank you to Dr. Long Luu. Although we did not overlap in the group, your work on the click project established us in the field of templated assembly and is the foundation upon which much of my Ph.D. work is built. I am deeply grateful for the work of Dr. Lien Nguyen, who helped to set our group up for biological studies and established our DM subgroup.

This work would also not have been possible without the help of my excellent undergraduate students in the Z-Group: Phil Kocheril, Niya Mitchell, and Hyoeun (Heather) Jeon. Each of you has shaped me as a scientist and mentor, and I am so grateful for the opportunity to work with such outstanding students. Your ideas and hard work have taken my projects (particularly

computational) to new levels that I would not have explored on my own. Thank you for investing your time in my projects and for entrusting your mentorship to Prof. Zimmerman and myself. As my first students, you helped me to grow as a mentor and I greatly appreciate your patience with me. I am thankful for your flexibility to work fully remotely when it became necessary during the pandemic, and I am so proud of all that each of you have accomplished despite the huge obstacles that you faced. I am inspired by your courage, fortitude, and desire to continually grow and learn. I look forward to hearing about your continued successes both in and out of the lab.

I want to thank my friends and colleagues at UIUC throughout the years. Thank you to all those involved in the Women Chemists Committee, the East Central Illinois American Chemical Society, various lecture and conference planning committees, and all those I have been privileged to work with through the department of chemistry. Working with and alongside you all has truly been my pleasure and has enhanced my Ph.D. experience. You inspire me with your leadership, dedication, and care for others. I also want to thank my colleagues in the Center for Innovation in Teaching and Learning, in particular Lucas Anderson, who throughout the years have helped me to develop as a teacher and helped to prepare me to take the next steps in my career.

I also need to express my gratitude to the excellent group of undergraduate students that I was privileged to work with at Hope College: Claire Benedict, Joe Cornell, Therese Joffre, Charlie Michels, Eric Salisbury, Erik Schoonover, Luke Shoemaker, Grace Stalions, and Jacob VanderRoest. When I arrived at Hope, I was a graduate student you had never met coming from a different state in the middle of a global pandemic, yet you did not hesitate to accept me into the group. You welcomed me to Hope College and into the Johnson lab with open arms and you taught me so much both in and out of the lab. You helped me to prepare for my career and made me a better scientist, mentor and person. You inspire me every day to be better for you and for my future

students. Your laughter, smiles, and personalities shine through to make the Johnson lab an awesome and productive environment. I hope you will never stop being yourselves and continue to have fun while growing and learning – life is too short to not enjoy it!

I owe a very huge thank you to Prof. Jeff Johnson at Hope College. Your Visiting Graduate Program truly changed my life, and I am so grateful for your willingness to bring me to your group and to allow me to work with your students. Being in your group fueled my motivation for finishing my Ph.D. and becoming a professor at a PUI. My summer at Hope was transformative and helped to prepare me for the next steps in my career. I am so grateful for your patience with me, your dedication to my scientific and professional development, and your constant commitment to answering my many questions. Your mentorship and guidance was instrumental in my job search process and getting through the end of my Ph.D. You are and will continue to be an inspiration and a role model for me. Thank you for teaching me about what it is like to be at a PUI and for your exemplary leadership. Your impact on student lives is apparent, and I aspire to be even half of the mentor and teacher that you are.

I need to thank several mentors, past and present. I have been so lucky to have Dr. Chrystal Bruce as a mentor to me through the Chemistry Women's Mentorship Network throughout my Ph.D. Chrystal, you have helped me through every step of my Ph.D., from literature seminar to applying for jobs, and I would not have made it through without your insight and encouragement. I was also incredibly blessed to have several excellent research mentors as an undergraduate student, who taught me how to do research and prepared me to be successful in my Ph.D. program, including Dr. Patti Kreke, Dr. Danny Miles, and Dr. Nathan Orloff. I am grateful for your continued support of me throughout my career.

I also want to thank the many excellent teachers who have contributed to my scientific training throughout the years. Each teacher who has had me in class has contributed to my development, and I appreciate your dedication to education. It was in college that I decided to pursue a Ph.D., in many thanks to the faculty at Mount St. Mary's University including Dr. Chris Bradley, Dr. Melanie Butler, Dr. Jonelle Hook, Dr. Patti Kreke, Dr. Christine McCauslin, Dr. Susan Mertins, Dr. Danny Miles, Dr. Garth Patterson, Dr. Mike Turner, Dr. Dana Ward, and all of the Mount faculty members.

My first chemistry course was in my sophomore year of high school, taught by Mr. Aaron Bubb. This is where I truly fell in love with chemistry and was the seed that began my chemistry career. Mr. Bubb, I would not be in chemistry today if not for your excellent teaching, your enthusiasm about chemistry, and your real-world examples that got us excited about chemistry. I had never watched CSI before your class, but your ability to help us connect chemistry to the show made me even more engaged in your course. I am continually inspired by your example as an incredible educator and hope you know that you had such a huge impact on my career because of your dedication to your students.

I would be remiss to not thank the people who helped me to get this Ph.D. through their day-to-day support. To my family, I cannot thank you enough for your unconditional love and support throughout this process. Although it was not always easy to understand what I was going through, you never stopped listening, praying, and loving me through the ups and downs of my Ph.D. To Mom, Dad, Monica, Bill, Hannah, Zack, Jamie, and all of my extended family, know that I am eternally grateful for each of you, and I would not be here today without you. I would also like to thank some of my closest friends, who stood by me or beside me throughout my Ph.D. Emily Mumford, Shannon Murphy, and Sam Weyant – thank you for your constant support and prayers!

Last but certainly not least, I need to extend a huge thank you to my husband. Russ, you have been by my side (from near and far) through every single day of this work. Your unending capacity to love and support me inspires me to be a better scientist, educator, and person. There are so many ways that you have helped me to be here today, from driving back and forth between Madison and Champaign-Urbana to listening as I talk through struggles and successes, and words cannot express how grateful I am to have you as my partner. I cannot wait to see what God has in store for the rest of our lives together!

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION TO NUCLEIC ACID TARGETING

1.1. NUCLEIC ACID TARGETING IN DRUG DISCOVERY

As traditional drug discovery efforts have largely focused on protein targeting, nucleic acid-targeting therapeutic agents have been comparatively under-explored. However, much of the genomic DNA that is transcribed into RNA in the human body is not translated into protein, and thus must have some other function or benefit to remain in the genome.[1] It is estimated that although about 95% of genomic DNA is transcribed to RNA, only about 5% of RNA is translated to protein.[1] Thus, DNA and RNA serve many more biological functions, including gene expression control. The ability to target the large part of the genome that is not expressed as protein could enable the treatment of diseases for which there is currently no known protein target as well as diseases that have "undruggable" protein targets. Though this approach is very attractive, it is challenging to develop specific nucleic acid binders, given that nucleic acids tend to have significantly less secondary and tertiary structure compared to protein targets.[2] Advances in understanding the role played by both coding and non-coding DNA and RNA in disease have brought nucleic acid targeting into focus as a potential therapeutic strategy.[3,4]

As nucleic acid therapeutics have shown promise, a particularly interesting example that has achieved clinical translation is the RNA vaccines for COVID-19, developed by Moderna and Pfizer, that utilize liquid nanoparticles to deliver mRNA encoding the SARS-CoV-2 spike protein to develop immunity.[5] Further, recent drug discovery efforts have enabled selective recognition of DNA and RNA sequences through many methods including antisense oligonucleotides,[6] CRISPR/Cas9 genome editing,[7,8] siRNA[9] and miRNA[10]-based approaches, and small-molecules.[11–13] The recent approval of an siRNA-based therapeutic for the treatment of polyneuropathy of hereditary transthyretin-mediated (hATTR) amyloidosis (Alnylam)[14] has

further sparked scientific interest in RNA-targeting agents. Many groups have explored the use of these and other DNA- and RNA-targeting strategies for a class of diseases caused by extended repeat sequences, or trinucleotide repeat expansion diseases (TREDs). Herein, we focus on myotonic dystrophy type 1 (DM1)[15–21] and Huntington's Disease (HD).[22–25] These diseases have well-studied pathobiologies mediated by expanded repeat sequences in both DNA and RNA, and thus provide an ideal model to study nucleic-acid targeting strategies.

1.2. MYOTONIC DYSTROPHY TYPE 1

DM1 is a neuromuscular disease caused by an expanded CTG repeat sequence in the 3'-untranslated region (3'-UTR) of the dystrophia myotonica protein kinase (*DMPK*) gene on chromosome 19.[26] Healthy individuals have 5-35 repeats, whereas individuals who manifest symptoms of the disease typically have 80 to more than 2,500 repeats.[27] The expanded (CTG·CAG) repeat sequence, $d(CTG{\cdot}CAG)^{exp}$, forms a DNA hairpin slip out, $d(CTG)^{exp}$, and is bidirectionally transcribed to form RNA with expanded (CUG) and (CAG) repeats, $r(CUG)^{exp}$ and $r(CAG)^{exp}$.[28] These expanded repeat transcripts result in a toxic RNA gain-of-function that leads to (i) sequestration of splicing regulators including muscleblind like 1 (MBNL1) that causes mis-regulation of splicing[29,30] as well as (ii) synthesis of undesirable homopeptides through repeat-associated non-ATG (RAN) translation.[31] Other mechanisms in the disease pathology include decreased expression of myocyte enhancer factor-2 (Mef2) leading to altered miRNA expression[32] and increased expression of CUG binding protein (CUG-BP)/Elav-like family member 1 (CELF1) leading to mis-regulation of pre-mRNAs.[33] In the primary mechanism of action, the $r(CUG)^{exp}$ formed by the transcription of $(CTG{\cdot}CAG)^{exp}$ forms a hairpin secondary structure[34,35] that sequesters the alternative splicing regulator muscleblind-like 1 (MBNL1), forming foci in the

nucleus of diseased cells (Figure 1.1).[36–38] This sequestration prevents proper splicing patterns in pre-mRNAs such as insulin receptor (IR), muscle-specific chloride ion channel (CLCN1), and cardiac troponin T (cTNT), causing insulin resistance, myotonia, and cardiac defects, respectively.[39–42] Given this disease pathobiology, small molecules that bind both the toxic RNA and the parent DNA repeat sequence could potentially inhibit MBNL1 sequestration and bidirectional transcription. Such selective binders could serve as a therapeutic strategy to ameliorate patient symptoms in DM1.

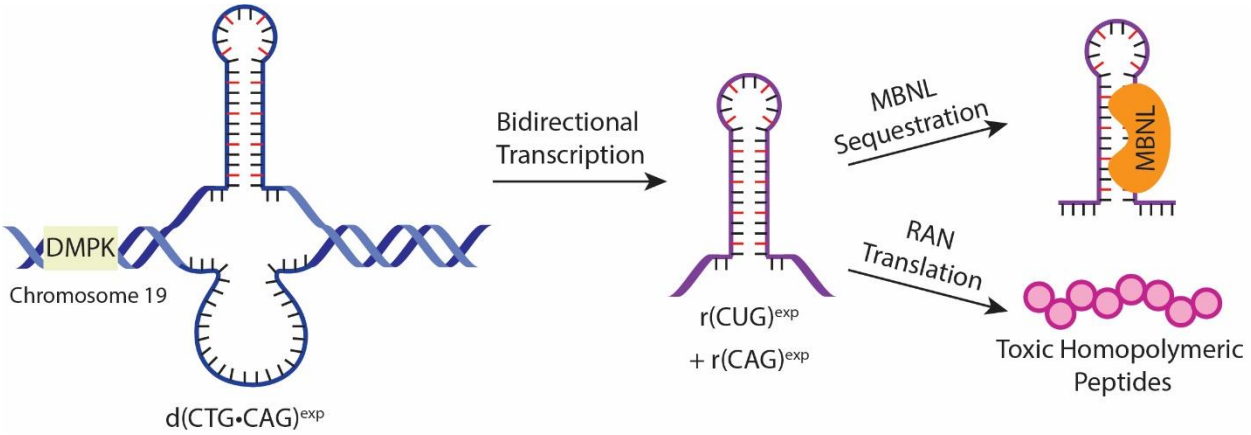


Figure 1.1. Myotonic Dystrophy Type 1 (DM1) Pathogenesis. Expanded d(CTG·CAG) repeats in the *DMPK* gene are bidirectionally transcribed to form r(CUG)^{exp} and r(CAG)^{exp}. The r(CUG)^{exp} can form hairpin secondary structures that sequester MBNL1 proteins and lead to improper splicing of pre-mRNAs. Both RNA transcripts can undergo RAN translation to form toxic homopolymeric peptides. The use of small-molecules or oligomers to specifically target RNA and DNA could alleviate symptoms of DM1 by preventing transcription of d(CTG·CAG)^{exp} and competitively binding the r(CUG)^{exp} hairpin structure to release MBNL1.

1.3. HUNTINGTON'S DISEASE

Huntington's Disease (HD) is another trinucleotide repeat disease that is caused by d(CTG·CAG)^{exp}. HD is a neurodegenerative disease that leads to striatal and cortical degradation that causes defects in motor and cognitive functions as well as psychiatric disturbance, weight loss, seizures, cardiac arrythmias, muscle atrophy (wasting), and diabetes.[43] Many of the symptoms overlap with DM1 and are caused by the same mechanisms: bidirectional transcription of

d(CTG·CAG)^{exp} to form r(CUG)^{exp} and r(CAG)^{exp}, r(CUG)^{exp} sequestration of splicing proteins such as MBNL1, and translation of r(CUG)^{exp} and r(CAG)^{exp} to form homopolymeric peptide sequences such as polyglutamine.[44] As in DM1, longer repeats have increased chance of strand slippage that leads to further expansion of the repeats; a healthy individual typically has about 9-35 repeats in the HTT gene, whereas diseased individuals carry 38 to about 100 repeats.[45,46] Although it is caused by the same repeat expansion as DM1, the mutation in HD occurs in a coding region within the huntingtin (HTT) gene on chromosome 4.[44,47] Thus, the repeat expansion leads to formation of mutant huntingtin protein with an expanded polyglutamine region that forms aggregates in the nucleus.[44] Reduction of HTT protein aggregates through various therapeutic strategies has been shown to improve cognitive and motor performance in HD mouse models.[43]

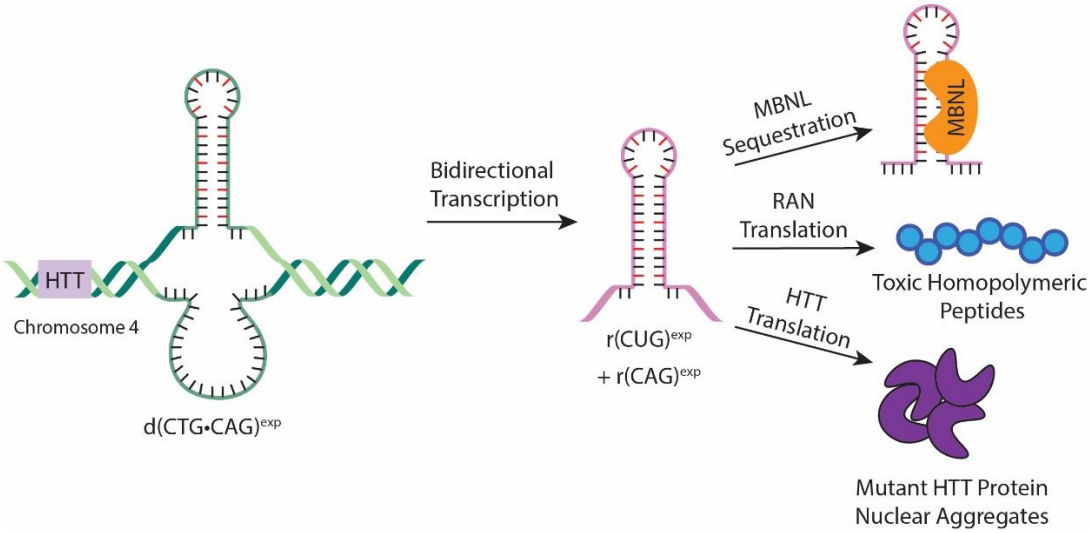


Figure 1.2. Huntington's Disease (HD) Pathogenesis. Expanded d(CTG·CAG) repeats in the HTT gene are bidirectionally transcribed to form r(CUG)^{exp} and r(CAG)^{exp}. The r(CUG)^{exp} can form hairpin secondary structures that sequester MBNL1 proteins and lead to improper splicing of pre-mRNAs. Both RNA transcripts can undergo RAN translation to form toxic homopolymeric peptides. Because HD occurs in a coding region, the HTT protein produced from the gene is mutant and causes the formation of nuclear aggregates. The use of small-molecules or oligomers to specifically target RNA and DNA could alleviate symptoms of HD by preventing transcription of d(CTG·CAG)^{exp}, competitively binding the r(CUG)^{exp} to release MBNL1, or binding r(CAG)^{exp} to prevent RAN translation and production of mutant HTT.

As in DM1, there are several additional mechanisms involved in the disease pathology of HD. As an example, increased production of mutant HTT protein results in downregulation of brain-

derived neurotrophic factor (BDNF), a protein that binds to neurotrophin receptors such as tyrosine receptor kinase B (TrkB) to support brain growth and survival as well as neuron differentiation.[43] Specifically, BDNF signaling pathways are thought to be involved in degradation of mutant HTT aggregates and necessary for proper cognitive development and function.[43] Efforts have been made to develop small molecules that act either as "BNDF-mimics" or directly and specifically activate TrkB signaling to mediate disease symptoms.[43] Although targeted activation of TrkB has shown some promise in HD therapeutic development efforts, one concern with this method is that TrkB levels are reduced in HD patients, and perhaps exist in insufficient quantities to allow full restoration of proper signaling and neurotrophic support.[43]

Another treatment strategy would be to prevent the formation of the mutant HTT protein all together by targeting inhibition at the DNA/RNA level rather than the protein level. DNA targeting in HD has been pursued through the development of targeted zinc finger proteins as well as CRISPR/Cas 9 targeting.[47] HD nucleic acid targeting has also been pursued by the Nakatani group through a series of naphthyridine-azaquinoline dimers that selectively bind $d(CAG)^{exp}$, inhibit DNA replication due to increased thermal stability of $d(CAG)_{10}$ hairpins, and can induce contractions or expansions of repeat regions.[45,48,49]

As previously mentioned, the expanded repeats in HD lie in a coding region. Importantly, the HTT gene codes for protein in both the sense and the antisense direction.[47] The sense protein (HTT) has unconfirmed function but is thought to be essential for development before birth.[50] It is hypothesized that knockout of HTT (e.g. through prevention of transcription) would not be more harmful than the disease in adult patients due to the negative impacts of the mutant protein.[50] It has also been hypothesized that the healthy HTT isoform may be involved in chemical signaling, transport, protein binding, and cell protection.[47] In the antisense direction, this gene codes for

transcription factor AS1 that regulates HTT levels; knockdown of AS1 has been shown to increase endogenous HTT.[47] RAN translation to form toxic homopolymeric peptides has also been observed in HD.[50]

1.4. OVERVIEW OF NUCLEIC ACID-TARGETING APPROACHES

We have explored several approaches for nucleic acid-targeting toward the development of therapeutic agents for DM1 and HD (Figure 1.3). These efforts include the development of novel small molecules and oligomers for selective targeting of trinucleotide repeat sequences, namely $d(CTG)^{exp}$ and $r(CUG)^{exp}$. The challenges that remain with these approaches include balancing potency with cell permeability. To achieve multivalent targeting while maintaining the promising cell permeability of small molecules, we envisioned the performing a ligation reaction with the nucleic acid target as a template. This would allow for dosing with cell-permeable small molecules that could form that larger, multivalent targeting agent *in situ*. Herein, two different ligation strategies for template-selected assembly of oligomeric agents have been developed and studied.

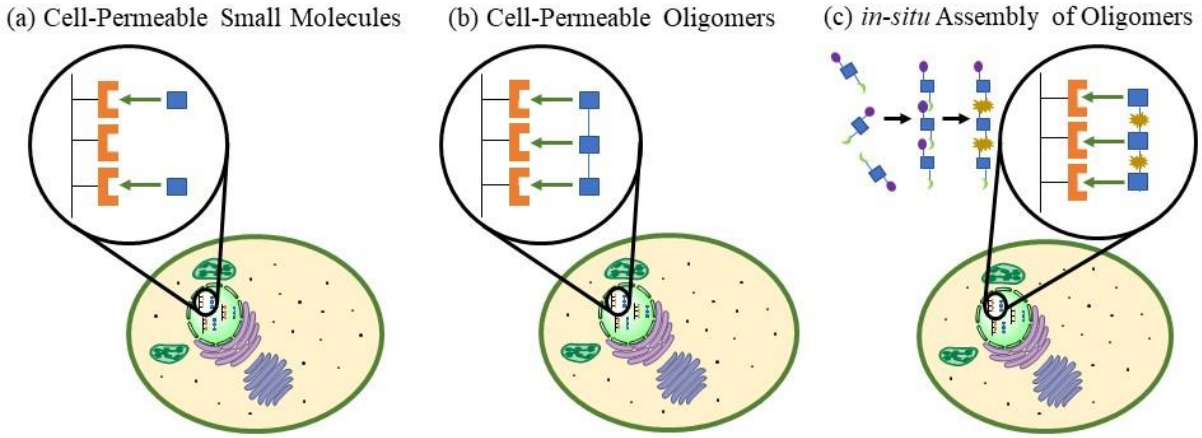
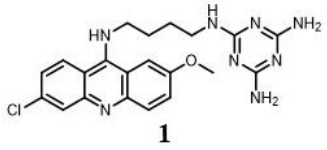


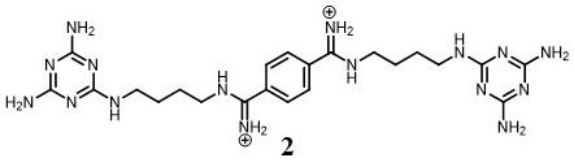
Figure 1.3. Overview of Proposed Strategies for Multivalent Targeting of DNA/RNA. (a) Small molecules can be used to target a characteristic site on the nucleic acid (e.g., recurring mismatches). (b) Cell-permeable oligomers such as **2** can be dosed directly to achieve multivalent targeting of repeat sequences. The challenge of this approach is synthetic accessibility and isolation of discrete oligomers (c) Use of *in situ*, nucleic acid-templated ligation chemistry to assemble DNA and RNA targeting molecules in cells. As reactive small molecules bind the DNA or RNA, they are brought into close enough proximity to undergo a ligation reaction and form a stronger binding agent after entering the cell. This approach improves cell permeability as well as ease of synthesis because smaller oligomers can be synthesized and later coupled *in situ* to form the desired oligomeric targeting agent.

1.4.1. EFFECTIVE SMALL MOLECULE TARGETING AGENTS

In the past 15 years, several small molecules have been developed for nucleic acid targeting and tested as potential therapeutic agents. The first report of a rationally designed small molecule that bound both DNA and RNA was by Zimmerman and coworkers in 2009.[51] This work was simultaneous with Miller and coworkers discovery of compounds that inhibit the interaction between r(CUG)^{exp} and MBNL through resin-bound dynamic combinatorial chemistry (RBDCC).[52] The Zimmerman group and others have previously reported several additional small molecules that effectively target DM1- and HD-relevant nucleic-acid sequences.[2,19,21,53–56] Although these potential therapeutics have shown much promise, none of these small molecules have made it to the clinic, suffering similar challenges to other noncoding RNA therapeutics.[57] The challenges in developing nucleic acid-targeting small molecules lie primarily in (i) achieving strong and specific binding (e.g. multivalent targeting) while (ii) maintaining cell permeability and (iii) low toxicity.



1



2

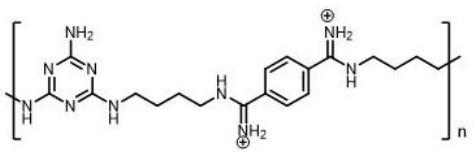
Our group has primarily focused on structure-based design of small molecules, including ligand **1** that contains an acridine intercalating agent linked to a melamine unit that acts as a T-T and U-U mismatch targeting agent in either a base triplet (Janus-wedge) interaction or a base flip out binding mode.[51,58,59] A new variant of this ligand, **2**, featured a bisamidinium major groove binder linked to two melamine targeting agents and exhibited lower toxicity and increased cell permeability.[60] Compound **2** selectively bound to r(CUG)^{exp} but showed no affinity to d(CTG)^{exp}.[60] Future generations of ligand, including one that

featured a small molecule RNase mimic that bound to r(CUG)$_{12}$ and d(CTG)$_{12}$, have shown promise to reduce r(CUG)^{exp} levels.[61]

Several other compounds have shown great potential as DM1 and HD therapeutic agents. Of particular interest in the context of this work is Actinomycin D (ActD), reported by the Berglund group as a transcription inhibitor that decreased the formation of r(CUG)^{exp}.[4] In addition to its binding to d(CTG·CAG)^{exp}, ActD is an FDA approved chemotherapeutic agent that can induce base flipping from a G-G mismatch in a (CGG) repeat commonly found in TREDs.[62,63] The Nakatani group has explored many compounds that bind the d(CAG)^{exp} or r(CAG)^{exp} sequences that are the causative agent in HD.[46,64–66] Originally reported as a G-G mismatch binding agent, their naphthyridine dimers induce slipping out in the (CAG) hairpin-like structure to cause a C-base to flip out adjacent to a G-G mismatch.[67] The group recently reported the ability of these compounds to induce contraction and expansion of repeat regions.[49,68] These molecules have established the potential of specific nucleic acid-targeting even for sequences that do not have strong secondary structure.

1.4.2. MULTIVALENT TARGETING AGENTS

Recent efforts in small molecule targeting of DNA and RNA secondary structures have suggested that multivalent targeting approaches lead to more effective treatments for myotonic dystrophy and other diseases.[69–72] This could be achieved through oligomeric targeting agents (Figure 1.3b). After finding success in targeting of r(CUG)^{exp} with small molecule **2**, Zimmerman and coworkers developed oligomeric mixture **3** that has shown enhanced promise as a therapeutic agent for DM1.[72] Attempts have been made by several members of the group to isolate discrete oligomers of this compound, including through solid phase synthesis in this work. A second series of discrete



3

oligomers was modeled in this work, consisting of bisguanidinium units in place of the bisamidinium units in **2** and **3**. Difficult synthesis and high cytotoxicity has limited the utility of these oligomers.

1.4.3. IN SITU, TEMPLATE-SELECTED SYNTHESIS

One strategy to overcome the synthetic barrier posed by development of discrete oligomeric agents while maintaining cell permeability would be to treat with monomeric units that could be assembled *in situ* on a nucleic acid template (Figure 1.3c). Rideout pioneered the idea of self-assembling therapeutics, in which two or more fragments are brought together to synthesize a more potent therapeutic.[73,74] In the case of DM1 therapeutics, it is hypothesized that the assembled structures have a higher affinity for $d(CTG)^{exp}$ and $r(CUG)^{exp}$ relative to their monomeric analogues. An additional benefit to this *in situ* assembly approach in the treatment of trinucleotide repeat diseases is that the oligomer synthesized *in situ* would have a size determined by length of the repeats and thereby maximize binding to that sequence. Thus, a personalized treatment could be realized by dosing with just a few monomers that would be used to synthesize an oligomer matching each patient's repeat length *in situ*.

One ligation reaction that has been explored for use in templated assembly of multivalent agents is the azide-alkyne 1,3-dipolar cycloaddition, or "click chemistry."[75] This strategy was first employed by Dervan and coworkers to assemble a polyamide-based therapeutic agent *in vitro*.[76] In addition to click chemistry, there are several other ligation reactions with potential to be used in a template-assembly strategy. This work has explored click chemistry as well as imine formation for *in situ* assembly of fragments on target. The latter, reversible ligation is attractive from a therapeutic perspective because the fragments are effectively recyclable. Cellular degradation of the target nucleic-acid or dissociation from the target would lead to hydrolysis of the imine bond

or exchange of the disulfide, thus degrading the oligomer to the monomeric fragments in the absence of template. These monomers, which are non-toxic, could go on to be re-assembled on a new template strand or diffuse out of the cell. Through this dynamic-covalent approach, recycling of monomers *in situ* could lead to extended therapeutic windows, lower required dosages, and decreased toxicity.

1.5. OVERVIEW AND IMPACT

DM1 and HD are two examples of repeat expansion diseases and provide an ideal model for exploring nucleic acid-targeting agents. This work explores the use of small molecules, oligomeric targeting agents, and self-assembling therapeutic agents for the treatment of DM1 and HD. Although focusing on these two diseases, the methods developed in this work could be applied to develop therapeutic agents for other diseases. This work has contributed to an increased understanding of how to develop strong and selective nucleic acid binders through (i) design and testing of novel targeting agents, (ii) exploration of the relationship between computational modeling approaches and experimental results, (iii) development of self-assembling therapeutics through two different ligation reactions, and (iv) exploration of an acid-amplifying polymeric drug delivery system. These approaches seek to directly address the challenge of achieving multivalent targeting with cell-permeable agents that have low cytotoxicity.

1.6. REFERENCES

(1) Pennisi, E. ENCODE Project Writes Eulogy for Junk DNA. *Science* **2012**, *337* (6099), 1159–1161.

(2) Warner, K. D.; Hajdin, C. E.; Weeks, K. M. Principles for Targeting RNA with Drug-like

Small Molecules. *Nat. Rev. Drug Discov.* **2018**, *17* (8), 547–558.

(3) Matsui, M.; Corey, D. R. Non-Coding RNAs as Drug Targets. *Nat. Rev. Drug Discov.* **2017**, *16* (3), 167–179.

(4) Siboni, R. B.; Nakamori, M.; Wagner, S. D.; Struck, A. J.; Coonrod, L. A.; Harriott, S. A.; Cass, D. M.; Tanner, M. K.; Berglund, J. A. Actinomycin D Specifically Reduces Expanded CUG Repeat RNA in Myotonic Dystrophy Models. *Cell Rep.* **2015**, *13* (11), 2386–2394.

(5) Cross, R. Moderna Provides Cursory Glance at First COVID-19 Vaccine Data. *C&EN News*. 2020.

(6) Ross, S. J.; Revenko, A. S.; Hanson, L. L.; Ellston, R.; Staniszewska, A.; Whalley, N.; Pandey, S. K.; Revill, M.; Rooney, C.; Buckett, L. K.; Klein, S. K.; Hudson, K.; Monia, B. P.; Zinda, M.; Blakey, D. C.; Lyne, P. D.; Macleod, A. R. Targeting KRAS-Dependent Tumors with AZD4785, a High-Affinity Therapeutic Antisense Oligonucleotide Inhibitor of KRAS. *Sci. Transl. Med.* **2017**, *9*, 1–13.

(7) Moreno, A. M.; Fu, X.; Zhu, J.; Katrekar, D.; Shih, Y. R. V.; Marlett, J.; Cabotaje, J.; Tat, J.; Naughton, J.; Lisowski, L.; Varghese, S.; Zhang, K.; Mali, P. In Situ Gene Therapy via AAV-CRISPR-Cas9-Mediated Targeted Gene Regulation. *Mol. Ther.* **2018**, *26* (7), 1818–1827.

(8) Cardinali, B.; Provenzano, C.; Izzo, M.; Battistini, J.; Golini, E.; Mandillo, S.; Scavizzi, F.; Raspa, M.; Voellenkle, C.; Perfetti, A.; Baci, D.; Martelli, F.; Cardiology, M. GENE THERAPY STRATEGIES FOR MYOTONIC DYSTROPHY TYPE 1. **2021**, 2021.

(9) Chi, X.; Gatti, P.; Papoian, T. Safety of Antisense Oligonucleotide and SiRNA-Based Therapeutics. *Drug Discov. Today* **2017**, *22* (5), 823–833.

(10) Bahrami, A.; Aledavood, A.; Anvari, K.; Hassanian, S. M.; Maftouh, M.; Yaghobzade, A.;

Salarzaee, O.; ShahidSales, S.; Avan, A. The Prognostic and Therapeutic Application of MicroRNAs in Breast Cancer: Tissue and Circulating MicroRNAs. *J. Cell. Physiol.* **2018**, *233* (2), 774–786.

(11) Donlic, A.; Morgan, B. S.; Xu, J. L.; Liu, A.; Roble, C.; Hargrove, A. E. Discovery of Small Molecule Ligands for MALAT1 by Tuning an RNA-Binding Scaffold. *Angew. Chemie - Int. Ed.* **2018**, *57* (40), 13242–13247.

(12) Liu, S.; Yang, Y.; Li, W.; Tian, X.; Cui, H.; Zhang, Q. Identification of Small-Molecule Ligands That Bind to MiR-21 as Potential Therapeutics for Endometriosis by Screening ZINC Database and in-Vitro Assays. *Gene* **2018**, *662* (April), 46–53.

(13) Matthes, F.; Massari, S.; Bochicchio, A.; Schorpp, K.; Schilling, J.; Weber, S.; Offermann, N.; Desantis, J.; Wanker, E.; Carloni, P.; Hadian, K.; Tabarrini, O.; Rossetti, G.; Krauss, S. Reducing Mutant Huntingtin Protein Expression in Living Cells by a Newly Identified RNA CAG Binder. *ACS Chem. Neurosci.* **2018**, *9* (6), 1399–1408.

(14) Alnylam Announces First-Ever FDA Approval of an RNAi Therapeutic, ONPATTRO™ (patisiran) for the Treatment of the Polyneuropathy of Hereditary Transthyretin-Mediated Amyloidosis in Adults | Alnylam Pharmaceuticals, Inc. http://investors.alnylam.com/news-releases/news-release-details/alnylam-announces-first-ever-fda-approval-rnai-therapeutic (accessed Jan 3, 2019).

(15) Gao, Z.; Cooper, T. A. Antisense Oligonucleotides: Rising Stars in Eliminating RNA Toxicity in Myotonic Dystrophy. *Hum. Gene Ther.* **2013**, *24* (5), 499–507.

(16) Rinaldi, C.; A Wood, M. J. Antisense Oligonucleotides: The next Frontier for Treatment of Neurological Disorders. *Nat. Publ. Gr.* **2017**.

(17) Provenzano, C.; Cappella, M.; Valaperta, R.; Cardani, R.; Meola, G.; Martelli, F.; Cardinali,

B.; Falcone, G. CRISPR/Cas9-Mediated Deletion of CTG Expansions Recovers Normal Phenotype in Myogenic Cells Derived from Myotonic Dystrophy 1 Patients. *Mol. Ther. - Nucleic Acids* **2017**, *9*, 337–348.

(18) Cerro-Herreros, E.; Sabater-Arcis, M.; Fernandez-Costa, J. M.; Moreno, N.; Perez-Alonso, M.; Llamusi, B.; Artero, R. MiR-23b and MiR-218 Silencing Increase Muscleblind-like Expression and Alleviate Myotonic Dystrophy Phenotypes in Mammalian Models. *Nat. Commun.* **2018**, *9* (1), 2482.

(19) Jenquin, J. R.; Coonrod, L. A.; Silverglate, Q. A.; Pellitier, N. A.; Hale, M. A.; Xia, G.; Nakamori, M.; Berglund, J. A. Furamidine Rescues Myotonic Dystrophy Type I Associated Mis-Splicing through Multiple Mechanisms. *ACS Chem. Biol.* **2018**, *13* (9), 2708–2718.

(20) Hsieh, W. C.; Bahal, R.; Thadke, S. A.; Bhatt, K.; Sobczak, K.; Thornton, C.; Ly, D. H. Design of a "Mini" Nucleic Acid Probe for Cooperative Binding of an RNA-Repeated Transcript Associated with Myotonic Dystrophy Type 1. *Biochemistry* **2018**, *57* (6), 907–911.

(21) Konieczny, P.; Selma-Soriano, E.; Rapisarda, A. S.; Fernandez-Costa, J. M.; Perez-Alonso, M.; Artero, R. Myotonic Dystrophy: Candidate Small Molecule Therapeutics. *Drug Discov. Today* **2017**, *22* (11), 1740–1748.

(22) Murmann, A. E.; Yu, J.; Opal, P.; Peter, M. E. Trinucleotide Repeat Expansion Diseases, RNAi, and Cancer. *Trends in Cancer* **2018**, *295*.

(23) Cleary, J. D.; Pattamatta, A.; Ranum, L. P. W. Repeat-Associated Non-ATG (RAN) Translation. *J. Biol. Chem.* **2018**, *293* (42), 16127–16141.

(24) Matthes, F.; Massari, S.; Bochicchio, A.; Schorpp, K.; Schilling, J.; Weber, S.; Offermann, N.; Desantis, J.; Wanker, E.; Carloni, P.; Hadian, K.; Tabarrini, O.; Rossetti, G.; Krauss, S.

Reducing Mutant Huntingtin Protein Expression in Living Cells by a Newly Identified RNA CAG Binder. *ACS Chem. Neurosci.* **2018**, *9* (6), 1399–1408.

(25) Grima, J. C. et al. Mutant Huntingtin Disrupts the Nuclear Pore Complex. *Neuron* **2017**, *94* (1), 93-107.e6.

(26) Mirkin, S. M. Expandable DNA Repeats and Human Disease. *Nature*. 2007.

(27) Brook, D. J. et al. Molecular Basis of Myotonic Dystrophy: Expansionof a Trinucleotide (CTG) Repeat at the 3' End of a Transcript Encoding a Protein Kinase Family Member. *Cell* **1992**, *68*, 799–808.

(28) Pearson, C. E. Slipped-Strand DNAs Formed by Long (CAG)(CTG) Repeats: Slipped-out Repeats and Slip-out Junctions. *Nucleic Acids Res.* **2002**, *30* (20), 4534–4547.

(29) Lin, X.; Miller, J. W.; Mankodi, A.; Kanadia, R. N.; Yuan, Y.; Moxley, R. T.; Swanson, M. S.; Thornton, C. A. Failure of MBNL1-Dependent Post-Natal Splicing Transitions in Myotonic Dystrophy. *Hum. Mol. Genet.* **2006**, *15* (13), 2087–2097.

(30) Charizanis, K. et al. Muscleblind-like 2-Mediated Alternative Splicing in the Developing Brain and Dysregulation in Myotonic Dystrophy. *Neuron* **2012**, *75* (3), 437–450.

(31) Zu, T. et al. Non-ATG-Initiated Translation Directed by Microsatellite Expansions. *Proc. Natl. Acad. Sci.* **2011**, *108* (1), 260–265.

(32) Kalsotra, A.; Singh, R. K.; Gurha, P.; Ward, A. J.; Creighton, C. J.; Cooper, T. A. The Mef2 Transcription Network Is Disrupted in Myotonic Dystrophy Heart Tissue, Dramatically Altering MiRNA and MRNA Expression. *Cell Rep.* **2014**, *6* (2), 336–345.

(33) Brinegar, A. E.; Cooper, T. A. Roles for RNA-Binding Proteins in Development and Disease. *Brain Research*. September 2016, pp 1–8.

(34) Pearson, C. E.; Sinden, R. R. Alternative Structures in Duplex DNA Formed within the

Trinucleotide Repeats of the Myotonic Dystrophy and Fragile X Loci. *Biochemistry* **1996**, *35* (15), 5041–5053.

(35) Kirchner, R.; Vogtherr, M.; Limmer, S.; Sprinzl, M. Secondary Structure Dimorphism and Interconversion between Hairpin and Duplex Form of Oligoribonucleotides. *Antisense Nucleic Acid Drug Dev.* **1998**, *8* (6), 507–516.

(36) Osborne, R. J.; Thornton, C. A. RNA-Dominant Diseases. *Hum. Mol. Genet.* **2006**, *15* (2), R162–R169.

(37) Fardaei, M.; Rogers, M. T.; Thorpe, H. M.; Larkin, K.; Hamshere, M. G.; Harper, P. S.; David Brook, J. *Three Proteins, MBNL, MBLL and MBXL, Co-Localize in Vivo with Nuclear Foci of Expanded-Repeat Transcripts in DM1 and DM2 Cells*; 2002; Vol. 11.

(38) Napierała, M.; Krzyzosiak, W. J. CUG Repeats Present in Myotonin Kinase RNA Form Metastable "slippery" Hairpins. *J. Biol. Chem.* **1997**, *272* (49), 31079–31085.

(39) Miller, J. W.; Urbinati, C. R.; Teng-umnuay, P.; Stenberg, M. G.; Byrne, B. J.; Thornton, C. A.; Swanson, M. S. Recruitment of Human Muscleblind Proteins to (CUG)n Expansions Associated with Myotonic Dystrophy. *EMBO J.* **2000**, *19* (17), 4439–4448.

(40) Savkur, R. S.; Philips, A. V.; Cooper, T. A. Aberrant Regulation of Insulin Receptor Alternative Splicing Is Associated with Insulin Resistance in Myotonic Dystrophy. *Nat. Genet.* **2001**, *29* (1), 40–47.

(41) Philips, A. V; Timchenko, L. T.; Cooper, T. A. Disruption of Splicing Regulated by a CUG-Binding Protein in Myotonic Dystrophy. *Science* **1998**, *280*, 737–741.

(42) Konieczny, P.; Stepniak-Konieczna, E.; Sobczak, K. MBNL Proteins and Their Target RNAs, Interaction and Splicing Regulation. *Nucleic Acids Res.* **2014**, *42* (17), 10873–10887.

(43) Simmons, D. A.; Belichenko, N. P.; Yang, T.; Condon, C.; Monbureau, M.; Shamloo, M.; Jing, D.; Massa, S. M.; Longo, F. M. A Small Molecule TrkB Ligand Reduces Motor Impairment and Neuropathology in R6/2 and BACHD Mouse Models of Huntington's Disease. *J. Neurosci.* **2013**, *33* (48), 18712–18727.

(44) Mirkin, S. M. Expandable DNA Repeats and Human Disease. *Nature* **2007**, *447* (7147), 932–940.

(45) Hagihara, M.; Nakatani, K. Inhibition of DNA Replication by a d(CAG) Repeat Binding Ligand. *Nucleic Acids Symp. Ser.* **2006**, *50*, 147–148.

(46) Li, J.; Sakata, A.; He, H.; Bai, L. P.; Murata, A.; Dohno, C.; Nakatani, K. Naphthyridine-Benzoazaquinolone: Evaluation of a Tricyclic System for the Binding to (CAG)n Repeat DNA and RNA. *Chem. - An Asian J.* **2016**, *11* (13), 1971–1981.

(47) Budworth, H.; McMurray, C. T. Bidirectional Transcription of Trinucleotide Repeats: Roles for Excision Repair. *DNA Repair (Amst).* **2013**, *12* (8), 672–684.

(48) Nakatani, K.; Hagihara, S.; Goto, Y.; Kobori, A.; Hagihara, M.; Hayashi, G.; Kyo, M.; Nomura, M.; Mishima, M.; Kojima, C. Small-Molecule Ligand Induces Nucleotide Flipping in (CAG)n Trinucleotide Repeats. *Nature Chemical Biology*. 2005, pp 39–43.

(49) Nakamori, M. et al. A Slipped-CAG DNA-Binding Small Molecule Induces Trinucleotide-Repeat Contractions in Vivo. *Nat. Genet.* **2020**, *52* (2), 146–159.

(50) McMurray, C. T. Mechanisms of Trinucleotide Repeat Instability during Human Development. *Nature Reviews Genetics*. November 1, 2010, pp 786–799.

(51) Arambula, J. F.; Ramisetty, S. R.; Baranger, A. M.; Zimmerman, S. C. A Simple Ligand That Selectively Targets CUG Trinucleotide Repeats and Inhibits MBNL Protein Binding. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (38), 16068–16073.

(52) Gareiss, P. C.; Sobczak, K.; Mcnaughton, B. R.; Palde, P. B.; Thornton, C. A.; Miller, B. L. Dynamic Combinatorial Selection of Molecules Capable of Inhibiting the (CUG) Repeat RNA-MBNL1 Interaction In Vitro: Discovery of Lead Compounds Targeting Myotonic Dystrophy (DM1). *J. Am. Chem. Soc.* **2008**, *130*, 16254–16261.

(53) Gonzalez, À. L.; Konieczny, P.; Llamusi, B.; Delgado-Pinar, E.; Borrell, J. I.; Teixidó, J.; García-España, E.; Pérez-Alonso, M.; Estrada-Tejedor, R.; Artero, R. In Silico Discovery of Substituted Pyrido[2,3-d] Pyrimidines and Pentamidine-like Compounds with Biological Activity in Myotonic Dystrophy Models. *PLoS One* **2017**, *12* (6).

(54) Hong, H.; Koon, A. C.; Chen, Z. S.; Wei, Y.; An, Y.; Li, W.; Lau, M. H. Y.; Lau, K. F.; Ngo, J. C. K.; Wong, C. H.; Au-Yeung, H. Y.; Zimmerman, S. C.; Chan, H. Y. E. AQAMAN, a Bisamidine-Based Inhibitor of Toxic Protein Inclusions in Neurons, Ameliorates Cytotoxicity in Polyglutamine Disease Models. *J. Biol. Chem.* **2019**, *294* (8), 2757–2770.

(55) Li, J.; Nakamori, M.; Matsumoto, J.; Murata, A.; Dohno, C.; Kiliszek, A.; Taylor, K.; Sobczak, K.; Nakatani, K. A Dimeric 2,9-Diamino-1,10-Phenanthroline Derivative Improves Alternative Splicing in Myotonic Dystrophy Type 1 Cell and Mouse Models. *Chem. - A Eur. J.* **2018**, 2–11.

(56) Onizuka, K.; Usami, A.; Yamaoki, Y.; Kobayashi, T.; Hazemi, M. E.; Chikuni, T.; Sato, N.; Sasaki, K.; Katahira, M.; Nagatsugi, F. Selective Alkylation of T-T Mismatched DNA Using Vinyldiaminotriazine-Acridine Conjugate. *Nucleic Acids Res.* **2018**, *46* (3), 1059–1068.

(57) Winkle, M.; El-Daly, S. M.; Fabbri, M.; Calin, G. A. Noncoding RNA Therapeutics — Challenges and Potential Solutions. *Nat. Rev. Drug Discov.* **2021**, *20* (8), 629–651.

(58) Wong, C. H.; Richardson, S. L.; Ho, Y. J.; Lucas, A. M. H.; Tuccinardi, T.; Baranger, A. M.; Zimmerman, S. C. Investigating the Binding Mode of an Inhibitor of the MBNL•RNA Complex in Myotonic Dystrophy Type 1 (DM1) Leads to the Unexpected Discovery of a DNA-Selective Binder. *ChemBioChem* **2012**, *13* (17), 2505–2509.

(59) Chien, C.-M.; Wu, P.-C.; Satange, R.; Chang, C.-C.; Lai, Z.-L.; Hagler, L. D.; Zimmerman, S. C.; Hou, M.-H. Structural Basis for Targeting T:T Mismatch with Triaminotriazine-Acridine Conjugate Induces a U-Shaped Head-to-Head Four-Way Junction in CTG Repeat DNA. *J. Am. Chem. Soc.* **2020**.

(60) Wong, C. H.; Nguyen, L.; Peh, J.; Luu, L. M.; Sanchez, J. S.; Richardson, S. L.; Tuccinardi, T.; Tsoi, H.; Chan, W. Y.; Chan, H. Y. E.; Baranger, A. M.; Hergenrother, P. J.; Zimmerman, S. C. Targeting Toxic RNAs That Cause Myotonic Dystrophy Type 1 (DM1) with a Bisamidinium Inhibitor. *J. Am. Chem. Soc.* **2014**, *136* (17), 6355–6361.

(61) Nguyen, L.; Luu, L. M.; Peng, S.; Serrano, J. F.; Chan, H. Y. E.; Zimmerman, S. C. Rationally Designed Small Molecules That Target Both the DNA and RNA Causing Myotonic Dystrophy Type 1. *J. Am. Chem. Soc.* **2015**, *137* (44), 14180–14189.

(62) Lian, C.; Robinson, H.; Wang, A. H. J. Structure of Actinomycin D Bound with (GAAGCTTC)2 and (GATGCTTC)2 and Its Binding to the (CAG)(n):(CTG)(n) Triplet Sequence as Determined by NMR Analysis. *J. Am. Chem. Soc.* **1996**, *118* (37), 8791–8801.

(63) Lo, Y. S.; Tseng, W. H.; Chuang, C. Y.; Hou, M. H. The Structural Basis of Actinomycin D-Binding Induces Nucleotide Flipping out, a Sharp Bend and a Left-Handed Twist in CGG Triplet Repeats. *Nucleic Acids Res.* **2013**, *41* (7), 4284–4294.

(64) Nakatani, K.; Hagihara, S.; Goto, Y.; Kobori, A.; Hagihara, M.; Hayashi, G.; Kyo, M.; Nomura, M.; Mishima, M.; Kojima, C. Small-Molecule Ligand Induces Nucleotide

Flipping in (Cag)n Trinucleotide Repeats. *Nat. Chem. Biol.* **2005**, *1* (1), 39–43.

(65) Murata, A.; Nakamori, M.; Nakatani, K. Modulating RNA Secondary and Tertiary Structures by Mismatch Binding Ligands. *Methods* **2019**, *167*, 78–91.

(66) Hasuike, Y.; Tanaka, H.; Gall-Duncan, T.; Mehkary, M.; Nakatani, K.; Pearson, C. E.; Tsuji, S.; Mochizuki, H.; Nakamori, M. CAG Repeat-Binding Small Molecule Improves Motor Coordination Impairment in a Mouse Model of Dentatorubral–Pallidoluysian Atrophy. *Neurobiol. Dis.* **2022**, *163*, 105604.

(67) Hagihara, S.; Kumasawa, H.; Goto, Y.; Hayashi, G.; Kobori, A.; Saito, I.; Nakatani, K. Detection of Guanine-Adenine Mismatches by Surface Plasmon Resonance Sensor Carrying Naphthyridine-Azaquinolone Hybrid on the Surface. *Nucleic Acids Res.* **2004**, *32* (1), 278–286.

(68) Hagihara, M.; Dohno, C.; Saito, K.; Sugimoto, K.; Hishinuma, Y.; Sohma, Y.; Shibata, T.; Nakatani, K. Short Tandem Repeat Contractions during in Vitro DNA Synthesis by Repeat-Binding Molecules. *Chem. Lett.* **2021**, *50* (11), 1851–1848.

(69) Krishnamurthy, V. M.; Estroff, L. A.; Whitesides, G. M. Multivalency in Ligand Design. In *Fragment-based Approaches in Drug Discovery*; 2006; Vol. 34, pp 11–53.

(70) Childs-Disney, J. L.; Yildirim, I.; Park, H.; Lohman, J. R.; Guan, L.; Tran, T.; Sarkar, P.; Schatz, G. C.; Disney, M. D. Structure of the Myotonic Dystrophy Type 2 RNA and Designed Small Molecules That Reduce Toxicity. *ACS Chem. Biol.* **2014**, *9* (2), 538–550.

(71) Bai, Y.; Nguyen, L.; Song, Z.; Peng, S.; Lee, J.; Zheng, N.; Kapoor, I.; Hagler, L. D.; Cai, K.; Cheng, J.; Chan, H. Y. E.; Zimmerman, S. C. Integrating Display and Delivery Functionality with a Cell Penetrating Peptide Mimic as a Scaffold for Intracellular Multivalent Multitargeting. *J. Am. Chem. Soc.* **2016**, *138* (30), 9498–9507.

(72) Lee, J.; Bai, Y.; Chembazhi, U. V; Peng, S.; Yum, K.; Luu, L. M.; Hagler, L. D.; Serrano, J. F.; Chan, H. Y. E.; Kalsotra, A.; Zimmerman, S. C. Intrinsically Cell-Penetrating Multivalent and Multitargeting Ligands for Myotonic Dystrophy Type 1. *Proc. Natl. Acad. Sci.* **2019**, *116* (18), 8709–8714.

(73) Rideout, D. Self-Assembling Cytotoxins. *Science* **1986**, *233* (4763), 561–563.

(74) Rideout, D.; Calogeropoulou, T.; Jaworski, J.; McCarthy, M. Synergism through Direct Covalent Bonding between Agents: A Strategy for Rational Design of Chemotherapeutic Combinations. *Biopolymers* **1990**, *29* (1), 247–262.

(75) Huisgen, R. 1,3-Dipolar Cycloadditions Past and Future. *Angew. Chem. Int. Ed* **1963**, *2* (10), 565–632.

(76) Poulin-Kerstein, A. T.; Dervan, P. B. DNA-Templated Dimerization of Hairpin Polyamides. *J. Am. Chem. Soc.* **2003**, *125* (51), 15811–15821.

CHAPTER 2: TEMPLATE-ASSISTED ASSEMBLY VIA CLICK CHEMISTRY

This chapter was adapted from the following publications:

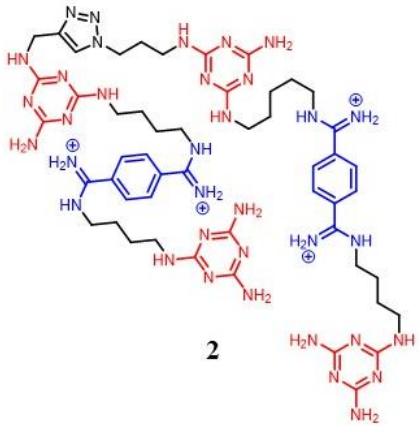
- Hagler, L.D., **Krueger, S.B.**, Luu, L.M., Lanzendorf, A.N., Mitchell, N.L., Vergara, J.I., Curet, D., and Zimmerman, S.C. Versatile Target-Guided Screen for Discovering Bidirectional Transcription Inhibitors of a Trinucleotide Repeat Disease. *ACS Med. Chem. Lett.* **2021** *12*(6) 935-940.
- Hagler, L.D., Luu, L.M., Tonelli, M., Lee, J., Hayes, S.M., **Bonson, S.E.**, Vergara, J.I., Butcher, S.E., and Zimmerman, S.C. Expanded DNA and RNA Trinucleotide Repeats in Myotonic Dystrophy Type 1 Select Their Own Multitarget, Sequence-Selective Inhibitors. *Biochemistry* **2020** *59*(37) 3463-3472.

2.1 INTRODUCTION

Although small molecules have been successful at targeting nucleic acids, it remains difficult to achieve specific binding.[1] This challenge can be overcome with multivalent targeting agents, for example oligomeric agents. However, difficult synthesis in addition to decreased cell permeability and increased toxicity of oligomeric agents have posed a barrier to their development to date. One strategy to address the challenge of balancing potency and cell permeability would be to utilize monomeric units that are functionalized with reactive groups to allow for assembly of multivalent agents *in situ* on the nucleic acid target. This therapeutic approach would allow for multivalent targeting of nucleic acids with improved affinity and cell permeability. Inspired by Rideout's self-assembling therapeutics, in which two or more fragments are brought together to synthesize a potent therapeutic,[2,3] we explored monomeric targeting agents functionalized with azide and alkyne moieties. Although click chemistry typically requires the use of a copper catalyst, we and others have hypothesized that a proximity-induced click reaction could occur if the reactive

groups were brought together on a nucleic acid template. Dervan's use of 1,3-dipolar cycloaddition, "click chemistry,"[4] to assemble a polyamide-based therapeutic agent[5] was the first demonstration of a proximity-induced click ligation on a nucleic acid template. The bioorthogonality of azides and alkynes is beneficial in reducing off-target reactivity in biological systems. Thus, this strategy could help to overcome the barrier of cell-permeability and contribute to development of optimal therapeutics for DM1.

1



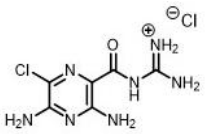
2

In the case of therapeutics for DM1 and other nucleotide repeat diseases, it is hypothesized that assembled, multivalent structures (from functionalized monomers) would have a higher affinity for d(CTG)^{exp} and r(CUG)^{exp} relative to their monomeric analogues. A highly potent inhibitor of the MBNL1-r(CUG)^{exp} interaction, **2**, was developed by Dr. Long Luu using click chemistry to form a heterodimer from substituted analogs of successful melamine-bisamidinium-melamine conjugate **1**, discussed in Chapter 1, containing azide and alkyne moieties.[6] Compound **2** inhibited MBNL binding to toxic RNA with ~1000 fold higher K_I. This result inspired further investigation of other dimeric, clickable molecules as potential therapeutics for DM1. Thus, a library of clickable targeting agents was designed by a previous member of the Zimmerman group, Dr. Long Luu, with the goal of screening for novel, optimized binders of d(CTG)^{exp} and r(CUG)^{exp}. This library was utilized to identify improved small molecules that target DM1-revelvant DNA and RNA sequences.[7–9] Using the repeat sequence as a template to bring azide and alkyne containing ligands into proximity, the 1,3-

dipolar cycloaddition ("click") reaction indeed proceeds without a copper catalyst. The goal of this project was to expand the library of clickable fragments with some new and promising scaffolds that might target d(CTG)^{exp} and r(CUG)^{exp}. The binding affinity, cytotoxicity, and *in vitro* transcription inhibition for several library members and hit compounds are reported herein.

2.2. RESULTS AND DISCUSSION

2.2.1. Expansion of Click Library



3

Amiloride hydrochloride, **3** (Midamor) is an anti-kaliuretic-diuretic agent that was approved by the FDA in 2002 and is marketed by Merck & Co.[10] It is currently used to treat hypertension and swelling caused by heart failure or cirrhosis.[10] When taken with diuretics, it can prevent potassium, hydrogen, and magnesium secretion by creating a negative potential across the luminal membrane.[10] This small molecule drug acts as an inhibitor of macropinocytosis and is excreted without being metabolized.[10,11] Amiloride has also been shown to intercalate DNA[12] and to bind HIV-1 transactivation response (TAR) RNA.[13] In 2017, Hargrove and coworkers reported the use of a variety of amiloride derivatives to target HIV-TAR RNA,[14] further establishing the potential of the amiloride scaffold as an RNA-targeting ligand. Previously, Teramae reported the "strong and selective" binding of amiloride to abasic sites opposite of T with a K_d of 150 nM and abasic sites opposite of U with a K_d value of 9.5 nM.[15,16] Although abasic sites are fundamentally different from mismatch sites, it is possible that these sites behave similarly in binding interactions with ligands. In particular, because one T- or U- base can be flipped out, the T-T or U-U mismatch site may mimick an abasic site,[17–24] thereby allowing selective recognition by amiloride. Taken together, the similarity of the

amiloride scaffold to previously developed triazine-based ligands and the established binding of amiloride to T- and U-abasic sites point to the potential for amiloride to target r(CUG)^{exp} as well as d(CTG)^{exp} (Figure 2.1).

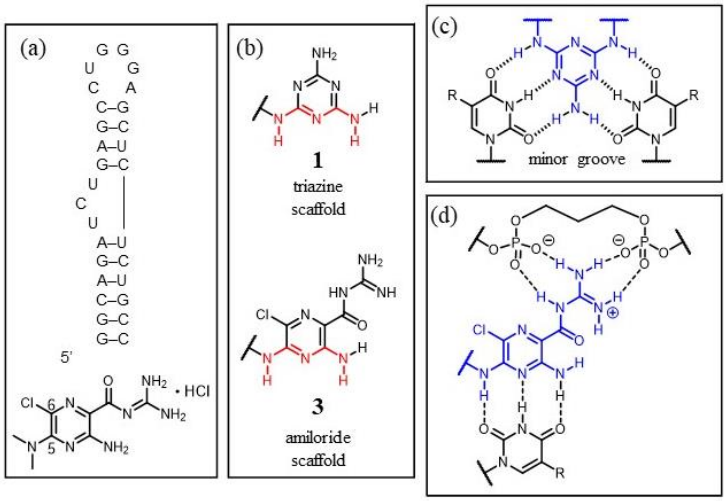
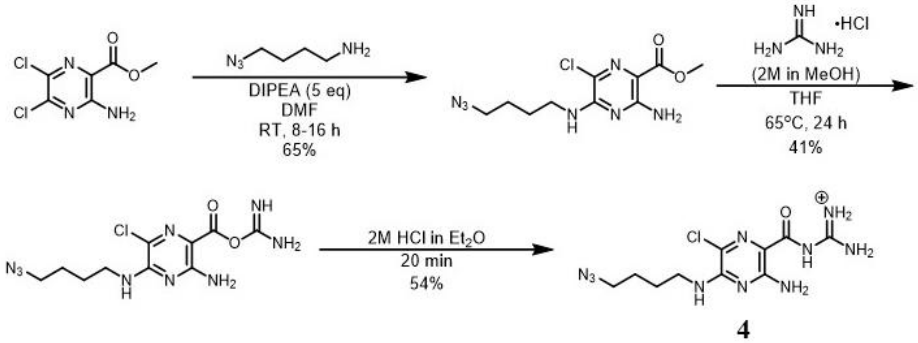


Figure 2.1. Amiloride shows promise as a potential d(CTG)^{exp} and r(CUG)^{exp} binding moiety. (a) Dimethyl amiloride was shown to bind to both the apical loop and the bulge of HIV-1 TAR RNA, but modifications at C5 and C6 shifted the binding affinity for the bulge region.[14] (b) The triazine-scaffold ligands and amiloride have similar potential for interaction with target sites (highlighted in red). (c) One of the potential interactions of interaction between ligand **1** and T-T (R = Me) or U-U (R = H) mismatches. (d) Teramae's proposed model of interaction, with amiloride hydrogen bonding to unpaired T- (R = Me) or U- (R = H) and the phosphate backbone at the abasic site.[15]

Given the potential for amiloride to act as a ligand for d(CTG)^{exp} and r(CUG)^{exp} sequences and the goal of expanding the click library, an amiloride derivative with a pendant alkyl azide, **4**, was synthesized by the route shown in Scheme 2.1. The stability of ligand **4** was monitored by ^{1}H NMR spectroscopy over 7 d and found to be stable in both aqueous solution and DMSO (Appendix A, Section A.2.1).



Scheme 2.1. Synthesis of amiloride derivative 4 for addition to click library. Methyl 3-amino-5,6-dichloropyrazine-2-carboxylate was coupled with 4-azidobutan-1-amine and the product was guanidinylated to make the functionalized amiloride to be tested for *in situ* DNA- and RNA- templated click reactions that would allow *in vivo* assembly of potential multivalent therapeutic agents. Details can be found in Section A.2.1.

Several other compounds were synthesized as potential members of the click library. Synthetic methods are detailed Section A.2.1. Activated alkynes, that contain an alkyne in the alpha position relative to a carbonyl group, seemed to be more reactive in this proximity-induced click reaction. However, the addition of activated alkyne functionality to hydrophobic fragments containing several amines proved particularly difficult. Thus, several alternative coupling partners (Figure 2.2a) were utilized in addition to the previously reported propiolic acid partner to make activated alkynes for the click library. [25–27] The coupling of two monomers was also attempted using a different route by performing the click reaction in an early step. This used propiolic acid to react with the monomer and would allow coupling with a monomer containing an activated ester instead of performing the click reaction with the two monomers (Figure 2.2b). This route would make products with slightly different linkers, it would provide an alternative to the challenging late-stage addition of activated alkynes. Although the direct coupling method proved ineffective, an alternative route was developed through TMS protection of the alkyne and is included in Section A.2.1. This work also re-synthesized monomeric click fragments for characterization by NMR and MS. Dimeric click products were also re-synthesized for characterization with Amie Lanzendorf. As demonstrated here, the click monomers and dimers proved extremely challenging to synthesize *ex vivo*.

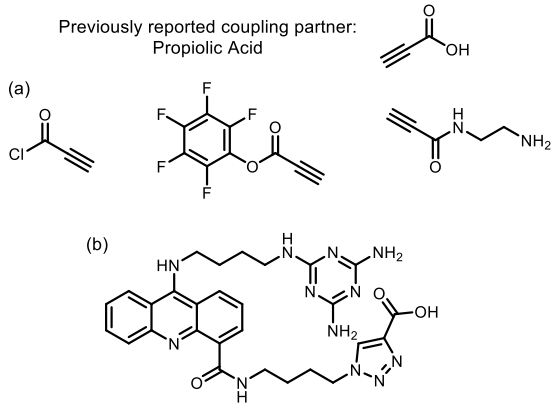


Figure 2.2. (a) Alternative coupling partners synthesized for use in making activated alkyne monomers from amine intermediates. (b) intermediate in route where triazole is attached and coupled directly instead of reacting two monomers.

2.2.2. Template-Assisted Click Screen (with Dr. Long Luu and Dr. Lauren Hagler)

The library of azide and alkyne fragments was tested in an *in situ* template-assisted click assay (Figure 2.3).[7] Although traditional azide-alkyne 1,3-dipolar cycloaddition reactions require a copper catalyst,[28] as noted above, the nucleic acid template could catalyze a proximity-induced, copper-free click reaction. This process would allow for assembly of targeting molecules on a template nucleic acid strand using the desired target as a catalyst, avoiding issues with cell permeability of larger molecules such as dimer **2**. After the building blocks have "clicked" to form the active agent selectively on target, they are expected to have stronger affinity for the target through the multivalent effect. Although product inhibition might be expected, the therapeutic goal is achieved by having the target site occupied.

Preliminary two-by-two click screens using d(CTG)^{exp} and r(CUG)^{exp} template strands showed promise in producing the template-assisted click products. The compounds (100 μM final

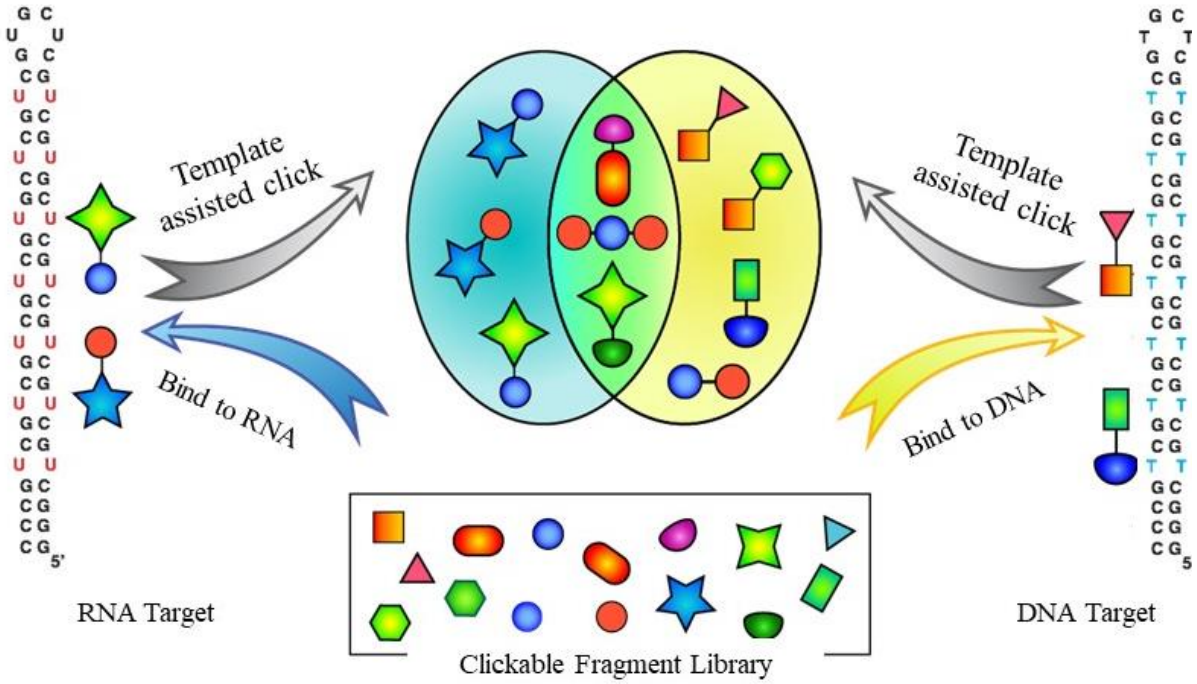


Figure 2.3. Template-Assisted Click Assay. The library of fragments was screened for combinations that click on the DNA and/or RNA targets.[8]

concentration of each) were incubated with the template sequence (10 μM final concentration) in buffer at 37 °C for up to 7 d. After degrading the template strand, reaction progress was monitored at 1, 3, and 7 d using MALDI-TOF MS and HPLC. Representative results with **5**, an alkyne derivative of ligand **1**, and the amiloride azide **4** are reported in Figure 2.4. After 1 d incubation with $d(CTG)_{16}$, a new peak was observed at *m*/*z* consistent with the **4**-**5** product. After 7 d incubation with $d(CTG)_{16}$, another new peak consistent with the **4**-**5**-**4** product was observed. After 7 d, the ratio of **4** : **5** : product in the mixture with $d(CTG)_{16}$ template was about 52 : 28 : 20 based on analytical HPLC. There was some uncertainty because the **4**-**5** and **4**-**5**-**4** products are thought to have very similar retention times and thus were integrated together. It is to be expected that amiloride **4** would be used up more quickly than **5** as the **4**-**5**-**4** product forms. Because the compound concentration was ten times that of the DNA template and the binding affinity of the **4**-**5** and **4**-**5**-**4** products is thought to be much higher than that of either **4** or **5** alone, product inhibition is expected to slow the turnover and thus limit the yield of this reaction. Upon incubation with a random DNA sequence, no new peaks were observed, supporting the conclusion that the template plays a role in the formation of product. When $r(CUG)_{90}$ was used as the template strand, a new peak was observed for the **4**-**5** product. The formation of click product on template suggests that amiloride may be a good binding partner for targeting DM1 relevant nucleic acid sequences.

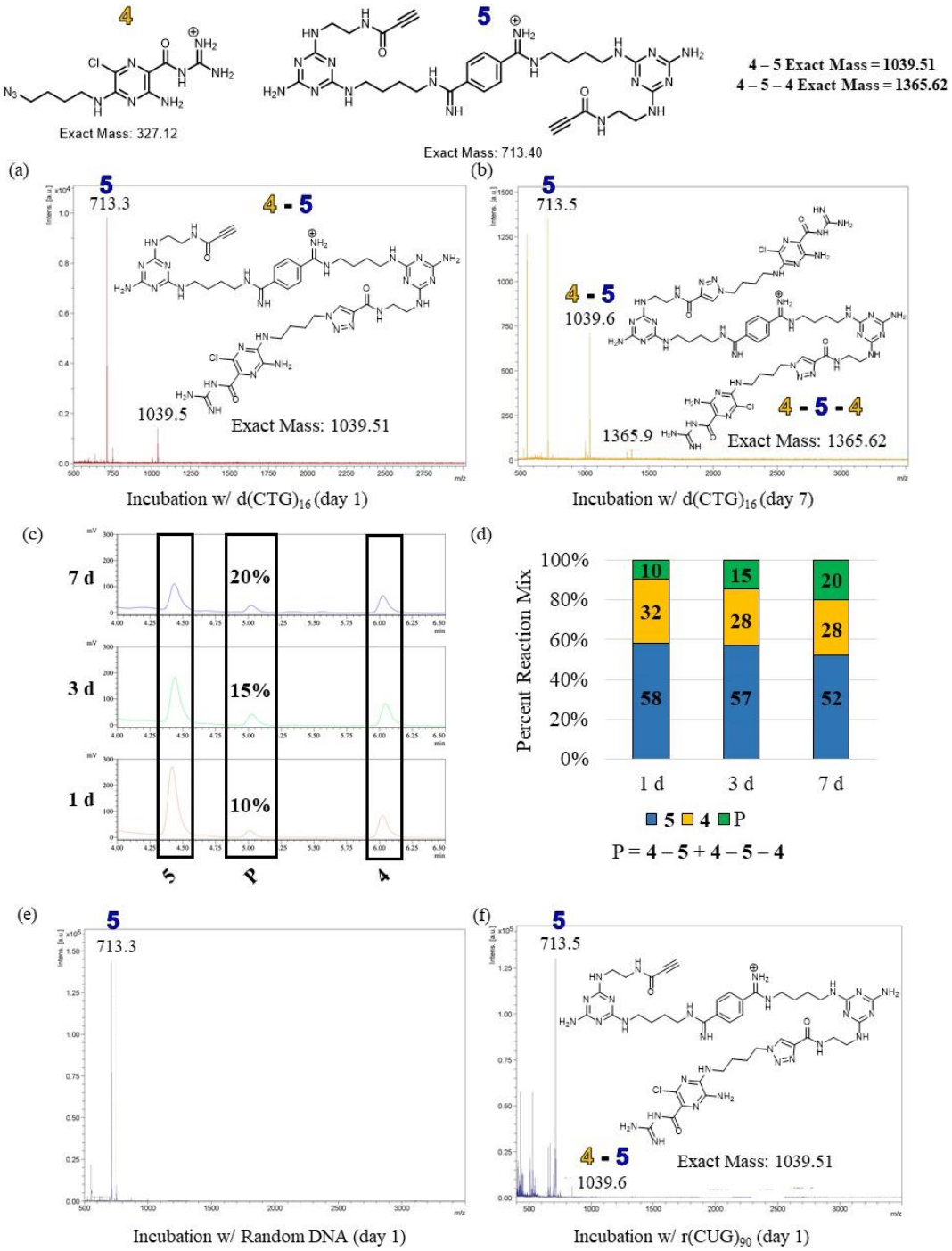
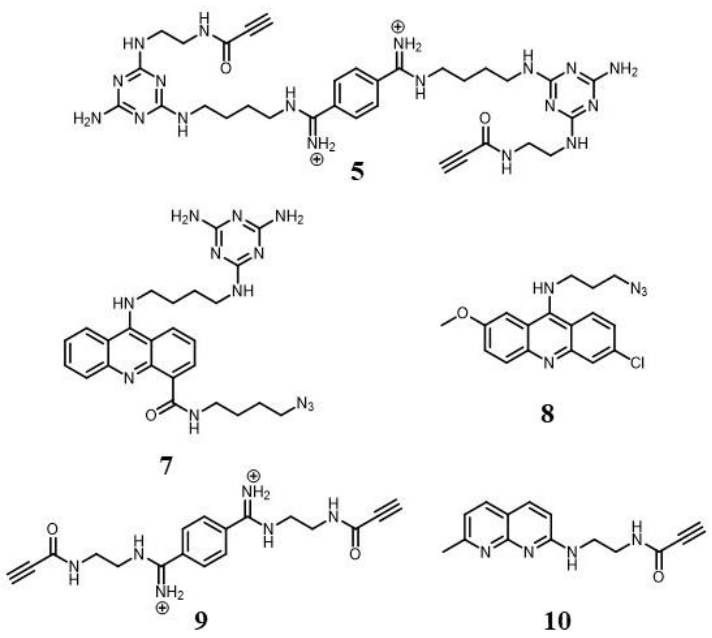


Figure 2.4. Click Screen Results: Alkyne 5 and Amiloride Azide 4. Nucleic acid templates were incubated with compounds in buffer at physiological pH and temperature for up to 7 d. Of note, amiloride **4** starting material is too small to be observed on MALDI. (a) After 1 d incubation with d(CTG)$_{16}$ target, the **4-5** product was observed. (b) After 7 d incubation with d(CTG)$_{16}$, both **4-5** and **4-5-4** products were observed in the MALDI. (c) HPLC traces for quantitation of reaction progress after 1, 3, and 7 d. (d) graphical representation of reaction mix calculated from areas under HPLC trace. (e) After 1 d incubation with random DNA duplex (5'-GGC TGG CTG GCT GGC TGG CGC-3'; 3'-CCG ACC GAC CGA CCG ACC GAC GCG-3') no new peaks were observed. (f) After 1 d incubation with r(CUG)$_{90}$ template, the **4-5** product was observed.

After having success in the preliminary 2x2 assays, we screened the rest of the library in a 2x2 fashion using MALDI-MS to detect click products. Through this method, we looked for new peaks consistent with the mass of dimeric or trimeric products. We also looked for products that formed selectively in the presence of the nucleic acid template and not in the presence of random DNA or in buffer alone. The most prominent hits observed by MALDI-MS are listed in Table 2.1 below. These hit compounds were prepared on larger scale further testing of their biological efficacy.



Compounds	Calculated Mass	Observed by MALDI-MS after 1 day incubation at 37 °C		
		$d(CTG)_{16}$ template (target)	random DNA template	buffer only
1+2	(d) 1227.4, (t) 1742.0	1227.8, 1742.0	n.d.	n.d.
3+2	868.4	858.5	n.d.	n.d.
8+2	(d) 867.0, (t) 1381.6	867.7, 1382.0	n.d.	n.d.

Table 2.1. Hits observed by MALDI-MS in the 2x2 click screen. (d) indicates dimeric product, (t) indicates trimer. n.d. indicates not detected. Masses are recorded after 1 d incubation at 37 °C in aqueous buffer consisting of 2 mM each of KCl, $CaCl_2$, $MgCl_2$, and Tris-HCl (pH 7).

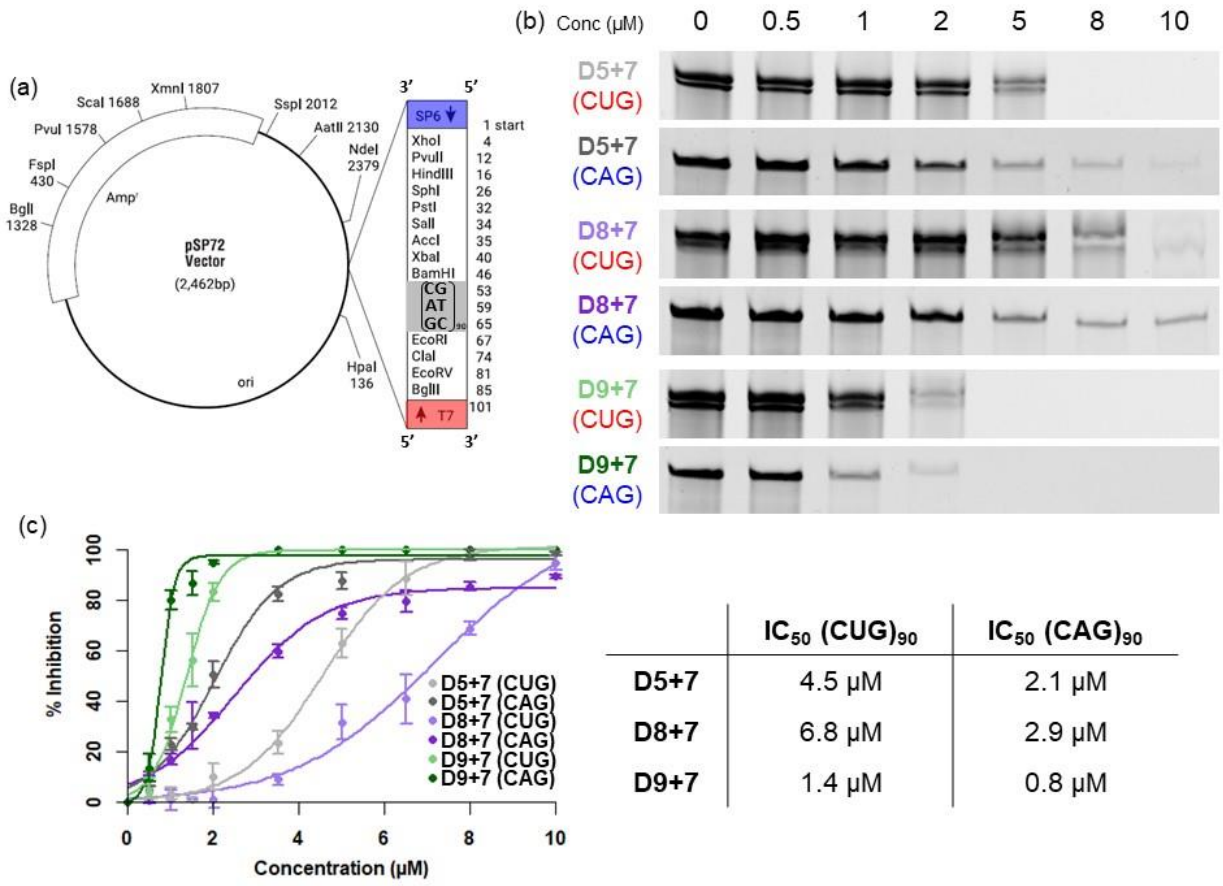
2.2.3. in vitro Bidirectional Transcription Inhibition

Following the library screen, four heterodimers of **7** and the corresponding monomers **5**, **8**, **9**, and **10** were evaluated for their ability to decrease the amount of $r(CUG)_{90}$ and $r(CAG)_{90}$ transcripts. Azide **7** was selected because its click products exhibited the most intense hits in the MALDI-MS screen. Three of the four dimers showed ca. 100% bidirectional transcription

inhibition at 10 μM whereas dimer **7+12** was much less effective.[9] Most of the corresponding monomeric compounds showed weak inhibition.

To further study the potency of the hit dimers in inhibiting transcription bidirectionally, IC_{50} values for the formation of both $r(CUG)_{90}$ and $r(CAG)_{90}$ were determined for dimeric compounds **5**+**7**, **8**+**7**, and **9**+**7** using the *in vitro* transcription inhibition assay (Figure 2.5). Interestingly, some of the inhibition curves were very steep, particularly for compound **9+7**, consistent with the inhibition observed for actinomycin D.[29] We hypothesize that the cause of the sharp decline in transcription over a narrow concentration window could be the result of either (i) cooperative binding of the compound to hairpin in the presence of T7 or SP6 polymerase or (ii) covalent attachment of the reactive alkyne compounds to the template.

For all three of these dimers, the observed IC_{50} value for $r(CAG)_{90}$ formation was lower compared to that for the formation of $r(CUG)_{90}$. This result can be explained by the compounds targeting d(CTG), thereby more effectively stalling polymerase on the sense strand. Of note, transcription of $d(CAG)_{90}$ (formation of $r(CUG)_{90}$) is still significantly inhibited, supporting our hypothesis that compound binding to one strand of the $d(CTG{\cdot}CAG)_{90}$ structure could inhibit the transcription of both strands through a stalling of the transcription complex. The inhibition of formation of both $r(CUG)_{90}$ and $r(CAG)_{90}$ at low micromolar concentrations suggests that these dimeric compounds may be effective against DM1.



	IC_{50} $(CUG)_{90}$	IC_{50} $(CAG)_{90}$
D5+7	4.5 μM	2.1 μM
D8+7	6.8 μM	2.9 μM
D9+7	1.4 μM	0.8 μM

Figure 2.5. *in vitro* bidirectional transcription inhibition of d(CTG·CAG)$_{90}$. (a) pSP72 plasmid with d(CTG·CAG)$_{90}$ spliced in between BamHI and EcoRI was used with T7 or SP6 polymerase was used to synthesize r(CUG)$_{90}$ and r(CAG)$_{90}$, respectively. (b) Sample gels from *in vitro* transcription inhibition assay. (c) IC_{50} curves and calculated IC_{50} values for hit dimers 1 + 2, 3 + 2, and 8 + 2 for inhibition of both r(CUG)90 and r(CAG)90 formation. Error is standard error of the mean, n =3

2.3. CONCLUSIONS

Synthesis of dimeric agents via template-assisted click chemistry provides an alternative approach that could help overcome the challenge of balancing potency with cell permeability. Herein we have outlined a method for the discovery of new ligands that bind to d(CTG)^{exp} and inhibit the production of the expanded transcripts, r(CUG)^{exp} and r(CAG)^{exp}, through screening of a clickable fragment library. This approach combines elements of fragment-based drug design (FBDD) and target guided screening (TGS) with a positive cross screen to discover multitargeting

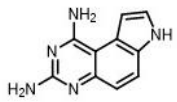
lead agents for DM1. Because the initial library did not lead to a new class of ligands, this work sought to expand the click library by adding new scaffolds such as amiloride. This broader discovery effort offered a higher level of diversity in the library and exploration of new scaffolds for targeting the DM1-relevant $d(CTG)^{exp}$ and $r(CUG)^{exp}$.

The DNA-binding hits identified in this screen inhibited transcription of $d(CTG{\cdot}CAG)_{90}$ bidirectionally, thus preventing the formation of both $r(CUG)_{90}$ and $r(CAG)_{90}$. Bidirectional transcription inhibition is an important therapeutic strategy for DM1 and other TREDs because of the potential not just to prevent formation of toxic $r(CUG)^{exp}$ and $r(CAG)^{exp}$, thus preventing toxic gain-of-function interactions between RNA and splicing proteins. Significantly, this strategy could also prevent the formation of toxic homopolymeric peptides by depleting the expanded transcripts and preventing RAN translation. Future efforts will examine whether these multivalent inhibitors are effective at reducing toxic RNA levels in cells. This target-guided, pairwise screening platform and the general DNA- or RNA-targeting nature of the molecules in the clickable fragment library allow for the discovery of new inhibitors for DM1 and the method may be extended to other diseases that have bidirectional RNA expression as a part of their pathogenesis including amyotrophic lateral sclerosis (ALS), frontotemporal dementia (FTD), and Huntington's disease (HD). Beyond the general approach described here for the discovery of more potent *in vitro* inhibitors and multitarget agents, this general protocol could be applied in cell culture. A cell-based approach would add the challenge of cellular and nuclear permeability, a critical aspect for drug development. More importantly, it would allow the prospect of a self-assembling therapeutic strategy using smaller molecules whose improved pharmacokinetic properties allow them to reach the target that facilitates the synthesis of the active agent. Such an approach would require a new method of product analysis and a faster and more biocompatible ligation reaction.

2.4. METHODS

2.4.1. Click Library Expansion: Additional Compounds

Derivatization of 7*H*-Pyrrolo[3,2-*f*]quinazoline-1,3-diamine

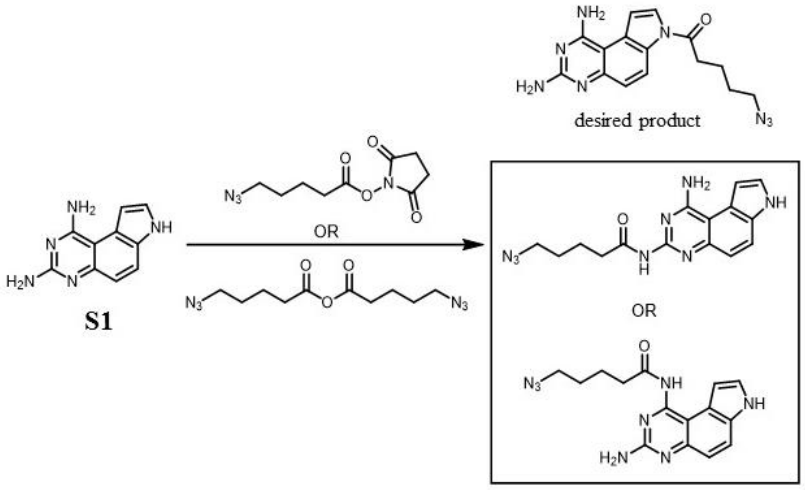


S1

7*H*-Pyrrolo[3,2-*f*]quinazoline-1,3-diamine, **S1**, was originally used as an antifolate,[30] then studied for its antimalarial properties,[31] and more recently been investigated for specific targeting of rapidly dividing cells for applications such as the treatment of breast cancer.[32] Although **S1** has come to be considered a privileged scaffold for biological targets, a potential problem with the use of this small molecule therapeutic is its inhibition of dihydrofolate reductase (DHFR), blocking the activity of folic acid which in turn inhibits cell division, DNA synthesis and repair mechanisms, and protein synthesis processes. Thus, any off-target effects would be toxic to healthy DNA and cells that encounter **S1**. Unpublished work by Eric Ho and Julio Serrano in the Zimmerman group indicated that **S1** has potential to bind the expanded repeat sequences relevant to DM1. Derivatization of this molecule by functionalization of the indole N-H or another position(s) on the scaffold could help to overcome the problem of DHFR inhibition by sterically or electronically blocking the interaction with the active site of the enzyme. To this end, attempted derivatization of **S1** was carried out in hopes of exploring the ligand's potential as a therapeutic agent for DM1.

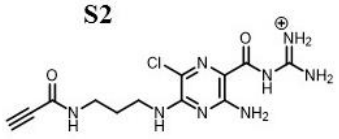
The first attempts at preparing a derivative of **S1**, involved acylation of the indole nitrogen with active esters containing a pendant azide. This modification would allow for addition of this molecule to the click library, and thus would contribute to the exploration of **S1** as a potential targeting molecule for $d(CTG)^{exp}$ and $r(CUG)^{exp}$. Various acylating agents were synthesized in attempt to achieve acylation of the indole nitrogen.[32] Unfortunately, the desired product was not obtained, but NMR analysis suggests that alkylation occurred at one of the aromatic amines

(Scheme 2.2). Further exploration of this acylation reaction, as well as Nuclear-Overhauser Effect (NOE) studies will be necessary to determine the exact position of acylation.



Scheme 2.2. Attempted Acylation of S1 at N7. The acylation did not form the desired product but occurred at either of the unsubstituted aromatic amines. Further studies will be necessary to definitively determine the position of substitution.

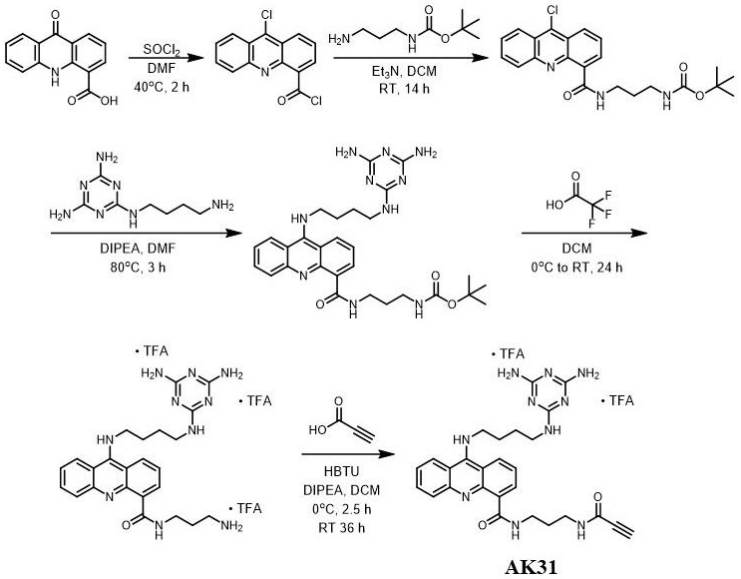
Amiloride Activated Alkyne Derivative



The synthesis of amiloride analogue **S2**, containing an activated alkyne for the click reaction, was attempted for further diversification of the click library, but not completed due to complications in the synthesis and generally low-yielding reactions. The challenge was the incompatibility of the activated alkyne with the guanidinylation step and the difficult purification of the by-product 1,3-diaminopropane from the initial coupling with amiloride (analogous to the first step in Scheme 2.2 above, but with *tert*-butyl (3-aminopropyl)carbamate in place of 4-azidobutan-1-amine).

Synthesis of AK31

An acridine-melamine conjugate that binds both r(CUG)^{exp} and d(CTG)^{exp} was previously synthesized with a pendant activated alkyne group, **AK31**.[8] The established route for the synthesis of **AK31** is shown in Scheme 2.3 (Long Luu Thesis p. xxxix-xli). Current attempts to produce **AK31** have been complicated by air and water sensitivity of the intermediate acridines. Thus far, the scheme below has been used to successfully yield up to 500 mg of the penultimate TFA salt, with attempts complicated by formation of inseparable mixtures of the product with various intermediates, decomposition products, and hydrolysis to the starting material acridone.



Scheme 2.3. Synthetic Route for AK31. Proposed synthesis of AK31 to further evaluate its activity toward d(CTG)^{exp} and r(CUG)^{exp} via screening of the click library and other experiments.

2.4.2. Synthetic Methods

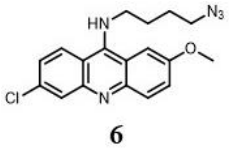
See Appendix A, Section A.2.1.

2.4.3. Click Screen (Run by Dr. Lauren Hagler)

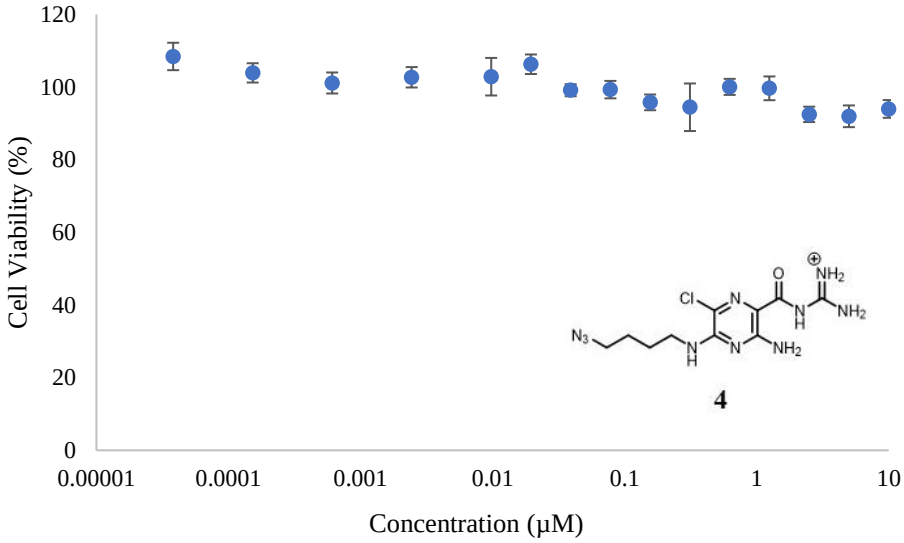
Each *in situ* click reaction was run in a 0.5 mL Eppendorf tube. The following is a general procedure performed with $d(CTG)_{16}$. To each tube was added 75 μL of nuclease-free water, 1 μL of 100 mM KCl, 1 μL of 100 mM $MgCl_2$, 1 μL of 100 mM Tris-HCl (pH 7.0), and 1 μL of 200 mM $CaCl_2$. The tubes were vortexed to mix. To each reaction tube was added 1 μL of 1 mM DNA (10 μM final concentration). The DNA was annealed by heating at 95 °C for 5 min and cooling over 90 min to room temperature. The appropriate azide and alkyne ligands were added to each reaction (10 μL each of 1 mM stock, 25% (v/v) DMSO, final concentration of each ligand: 100 μM, 2.5% (v/v) DMSO). The reactions were incubated at 37 °C for up to 7 d and monitored by MALDI-TOF mass spectrometry. A 10 μL aliquot was removed from the reaction tube and 2 μL of DNase I was added to degrade the template strand. The denaturing reaction was incubated at 37 °C for 30 min and submitted for MALDI-MS in the UIUC SCS mass spectrometry lab. After 1, 3, and 7 d, the reactions were denatured using DNase I and analyzed by HPLC (injection volume: 75 μL, 0-100% MeCN, 0.1% TFA, flow rate 1.5 mL/min). The percent reaction was determined as the ratio of the product peak over the combined area of the starting material and product peaks.

2.4.4. SRB Cytotoxicity Assays: Click Library

Before incubation with DM1 disease-relevant targets, a panel of sulforhodamine B (SRB) cytotoxicity assays[33] was performed for several compounds in the library, including amiloride **4**, alkyne **5**, and acridine **6** to determine an ideal working concentration for the assays (Figure 2.6-2.8). The general method is recorded in Appendix A, Section A.1.3. The acridine-based compounds in the library (such as **6**) were expected to be the most cytotoxic, as was observed in the assays.



6



4

Figure 2.6. Cytotoxicity Assays for amiloride azide 4. Performed according to the protocol detailed in Appendix A, Section A.1.3. Highest concentration: 100 µM. Error bars report standard error. Each point represents the average of at least 10 replicates.

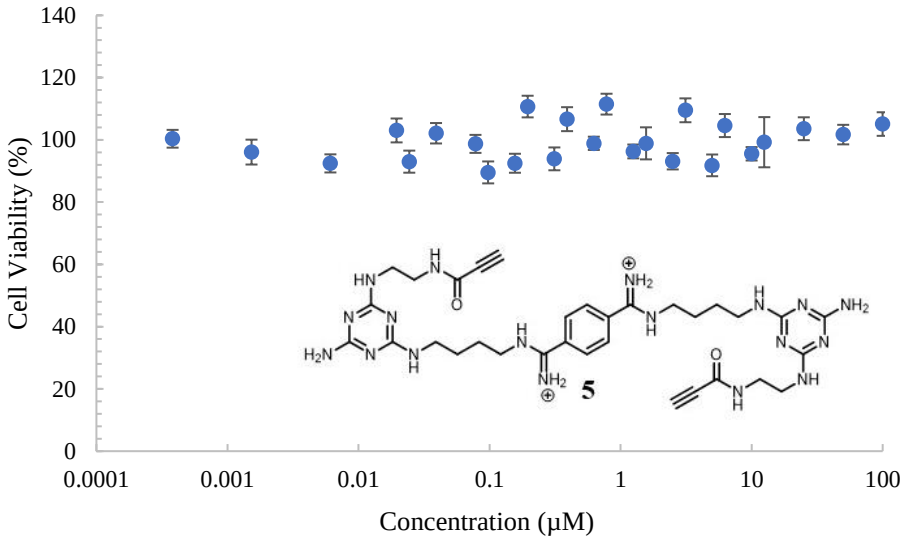


Figure 2.7. Ligand 6 (AK13) Cytotoxicity Assays. Performed according to the protocol detailed in Appendix A, Section A.1.3. Highest concentration: 100 µM. Error bars report standard error. Each point represents the average of at least 12 replicates.

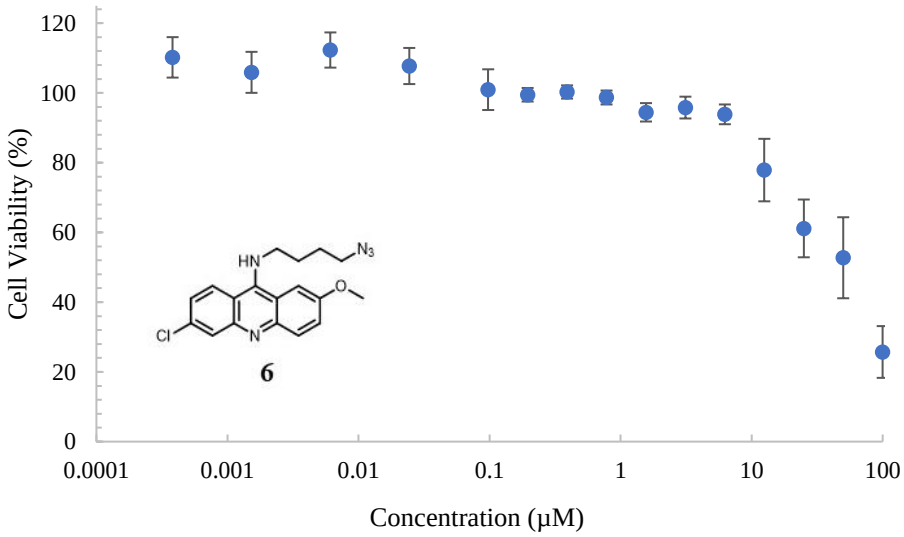


Figure 2.8. AZ35 Cytotoxicity Assays. Performed according to the protocol detailed in Appendix A, Section A.1.3. Highest concentration: 100 µM. Error bars report standard error. Each point represents the average of at least 10 replicates.

2.4.5. Isothermal Titration Calorimetry (ITC) Binding Studies: Click Monomers

ITC studies were performed with several compounds that produced hits in the click screen (Figure 2.9). Control ITCs (Figure 2.10) show that no background heat was generated. Of note, none of the ligands tested showed binding to DNA on their own. However, all ligands tested with ITC produced a heterodimeric click product in the presence of both d(CTG)^{exp} and r(CUG)^{exp} in the click screen. Although it was originally hypothesized that the small molecule fragments would first bind to the nucleic acid template and then undergo a proximity-induced click reaction, these binding studies suggest that other possibilities may need to be considered for the interaction at these concentrations. Because the click products were observed at compound concentration of 100 µM and the final compound concentration in the ITCs was only about 80 µM (a total of 272 µL of 500 µM solution was injected into a 1.42 mL cell), it is possible that the weak binding of the monomers was not observed in these ITC experiments. Another possibility is that the binding of

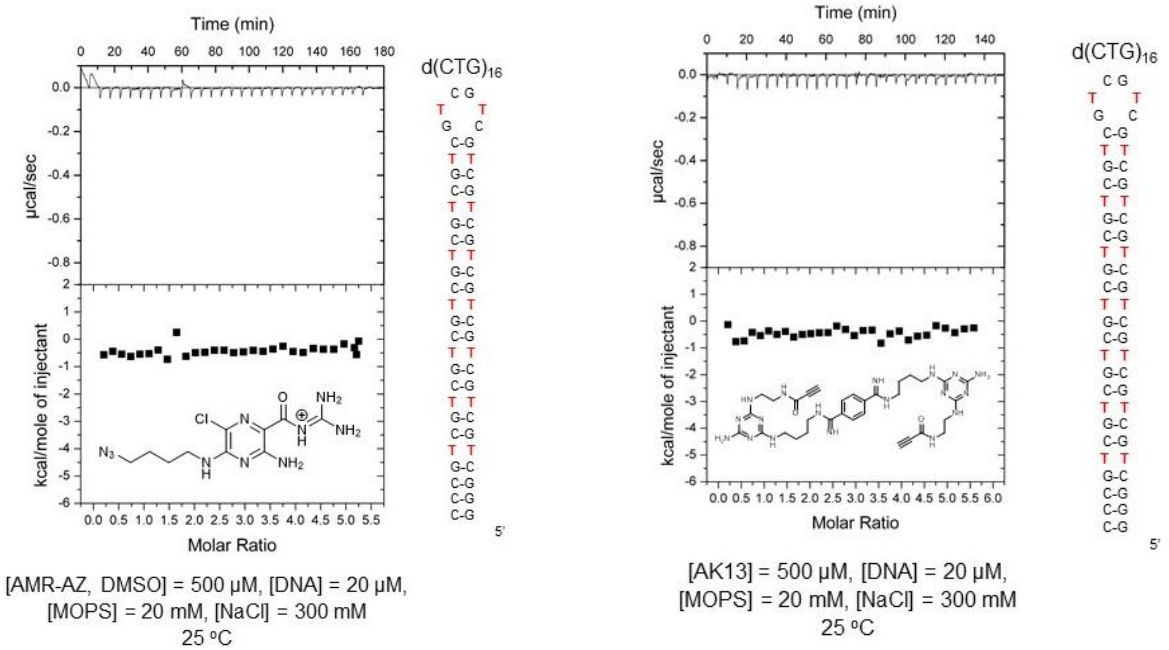


Figure 2.9. Representative ITC curve for interaction between d(CTG)$_{16}$ and click monomers 4 and 5. (right) No binding interaction was observed between amiloride **4** and d(CTG)$_{16}$. Of note, amiloride is a weak DNA intercalator (K_d = 1.3 mM), consistent with the lack of binding observed here.[12] (left) No binding interaction was observed. Three independent experiments were performed, and no binding was observed in any of the experiments.

these fragments is cooperative, and thus binding of each monomer is not observed. This hypothesis could be tested by running an ITC with both fragments (one ligand at constant concentration in d(CTG)^{exp} solution, the other titrated into that solution) to test if binding of one occurs in the presence of the other. Regardless of the mechanistic explanation, because click product has not been observed in the control reactions, the specific nucleic acid template appears to be playing some role in product formation.

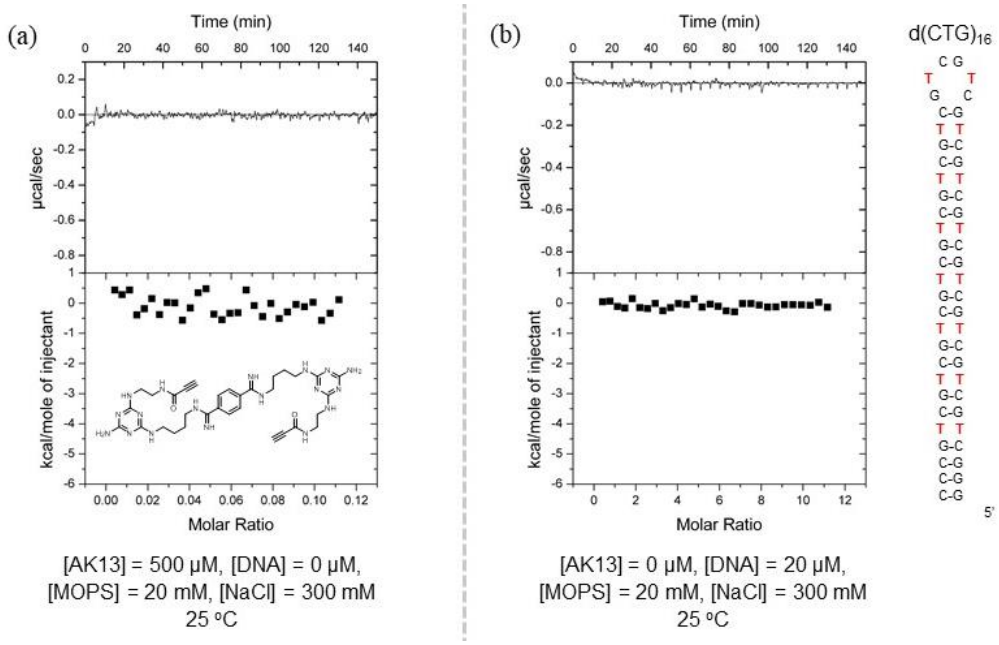


Figure 2.10. Representative ITC controls for ligand 5 (AK13). (a) Blank ITC titrating AK13 (**5**) into buffer solution. The same general method was followed, but biological grade water was added in place of the DNA in this experiment. (b) Blank ITC titrating buffer solution into d(CTG)$_{16}$. The same general method was followed, but biological grade water was added in place of the 10 mM aqueous ligand stock solution in this experiment.

Isothermal Titration Calorimetry: MicroCal VP-ITC

ITC Binding Experiments were performed on a MicroCal VP-ITC (MicroCal, Inc., Northampton, MA). The syringe for this instrument holds 250 µL and the reference and sample cells hold 1.24 mL each. The DNA oligos were purchased from Integrated DNA Technologies (IDT). To form duplex oligonucleotides, the complimentary single strands were mixed in equimolar amounts and

the DNA was annealed. All DNA was freshly annealed before each ITC experiment by heating at 90-95 °C for 5 min and slowly cooling back to 25 °C (1-2 h). The ITC was cleaned with water and methanol before each run, and the reference cell was changed before each new experiment. In preparation of necessary solutions, a master buffer solution was prepared and split amongst the reference cell solution, the DNA solution, and the ligand solution. Each ITC experiment consisted of 28 injections. The first injection was 2.0 µL over 5.1 s and all subsequent injections were 10 µL over 20.5 s. The initial delay was set to 300 s and a 400 s delay was included between each injection. All ITC experiments were run at 25 °C and stirring speed of 310 revolutions per min. The buffer for the reference, DNA, and ligand solutions was made from stock solutions of NaCl (5 M, aqueous) and MOPS (1 M, pH 7 ± 0.2) to a final concentration of 300 mM NaCl and 20 mM MOPS (pH 7) in water. No DMSO was used in these experiments. For ITC experiments with the clickable fragments, the ligand (syringe) concentration was 500 µM and the DNA (cell) concentration was 20 µM. The "High" feedback mode, "Fast Equilibration," and "Auto Equilibration" features were utilized in these experiments. The heat generated from each injection was determined using Origin 7.0 (MicroCal, Inc. Northampton, MA) to integrate each isotherm (µcal per s v. s). To account for heat of mixing and diffusive mixing at the tip of the syringe, the first data point from each ITC experiment was deleted when fitting the data to binding models. The automatically predicted baseline was inspected and manually adjusted to the base of each peak for all data points in the experiment. For these experiments, the Origin "one-site binding" algorithm (for one type of binding site) was used to predict the number of binding sites (N) and the dissociation constant (K_d) to be consistent with the other calculations in Julio's manuscript and also so that we could obtain an N value to satisfy the reviewer's desire for use of ITC to determine stoichiometry of binding.

2.5. REFERENCES

(1) Warner, K. D.; Hajdin, C. E.; Weeks, K. M. Principles for Targeting RNA with Drug-like Small Molecules. *Nat. Rev. Drug Discov.* **2018**, *17* (8), 547–558.

(2) Rideout, D. Self-Assembling Cytotoxins. *Science (80-.).* **1986**, *233* (4763), 561–563.

(3) Rideout, D.; Calogeropoulou, T.; Jaworski, J.; McCarthy, M. Synergism through Direct Covalent Bonding between Agents: A Strategy for Rational Design of Chemotherapeutic Combinations. *Biopolymers* **1990**, *29* (1), 247–262.

(4) Huisgen, R. 1,3-Dipolar Cycloadditions Past and Future. *Angew. Chem. Int. Ed* **1963**, *2* (10), 565–632.

(5) Poulin-Kerstein, A. T.; Dervan, P. B. DNA-Templated Dimerization of Hairpin Polyamides. *J. Am. Chem. Soc.* **2003**, *125* (51), 15811–15821.

(6) Luu, L. M.; Nguyen, L.; Peng, S.; Lee, J. Y.; Lee, H. Y.; Wong, C. H.; Hergenrother, P. J.; Chan, H. Y. E.; Zimmerman, S. C. A Potent Inhibitor of Protein Sequestration by Expanded Triplet (CUG) Repeats That Shows Phenotypic Improvements in a Drosophila Model of Myotonic Dystrophy. *ChemMedChem* **2016**, *11* (13), 1428–1435.

(7) Luu, L. M.; Zimmerman, S. C. Approaches to the Assembly of Potent Therapeutic Agents for the Treatment of Myotonic Dystrophy, 2016.

(8) Hagler, L. D.; Luu, L. M.; Tonelli, M.; Lee, J.; Hayes, S. M.; Bonson, S. E.; Vergara, J. I.; Butcher, S. E.; Zimmerman, S. C. Expanded DNA and RNA Trinucleotide Repeats in Myotonic Dystrophy Type 1 Select Their Own Multitarget, Sequence-Selective Inhibitors. *Biochemistry* **2020**, *59* (37), 3463–3472.

(9) Hagler, L. D.; Krueger, S. B.; Luu, L. M.; Lanzendorf, A. N.; Mitchell, N. L.; Vergara, J. I.; Curet, L. D.; Zimmerman, S. C. Versatile Target-Guided Screen for Discovering Bidirectional Transcription Inhibitors of a Trinucleotide Repeat Disease. *ACS Med. Chem. Lett.* **2021**, *12* (6), 935–940.

(10) Merck. Center for Drug Evaluation and Research: Midamor Final Printed Labeling. 2002.

(11) Koivusalo, M.; Welch, C.; Hayashi, H.; Scott, C. C.; Kim, M.; Alexander, T.; Touret, N.; Hahn, K. M.; Grinstein, S. Amiloride Inhibits Macropinocytosis by Lowering Submembranous PH and Preventing Rac1 and Cdc42 Signaling. *J. Cell Biol.* **2010**, *188* (4), 547–563.

(12) Besterman, J. M.; Elwell, L. P.; Blanchard, S. G.; Cory, M. Amiloride Intercalates into DNA and Inhibits DNA Topoisomerase II. *J. Biol. Chem.* **1987**, *262* (27), 13352–13358.

(13) Stelzer, A. C.; Frank, A. T.; Kratz, J. D.; Swanson, M. D.; Gonzalez-Hernandez, M. J.; Lee, J.; Andricioaei, I.; Markovitz, D. M.; Al-Hashimi, H. M. Discovery of Selective Bioactive Small Molecules by Targeting an RNA Dynamic Ensemble. *Nat. Chem. Biol.* **2011**, *7* (8), 553–559.

(14) Patwardhan, N. N.; Ganser, L. R.; Kapral, G. J.; Eubanks, C. S.; Lee, J.; Sathyamoorthy, B.; Al-Hashimi, H. M.; Hargrove, A. E. Amiloride as a New RNA-Binding Scaffold with Activity against HIV-1 TAR. *Medchemcomm* **2017**, *8* (5), 1022–1036.

(15) Zhao, C.; Dai, Q.; Seino, T.; Cui, Y. Y.; Nishizawa, S.; Teramae, N. Strong and Selective Binding of Amiloride to Thymine Base Opposite AP Sites in DNA Duplexes: Simultaneous Binding to DNA Phosphate Backbone. *Chem. Commun.* **2006**, No. 11, 1185–1187.

(16) Sato, Y.; Ichihashi, T.; Nishizawa, S.; Teramae, N. Strong and Selective Binding of Amiloride to an Abasic Site in RNA Duplexes: Thermodynamic Characterization and MicroRNA Detection. *Angew. Chemie - Int. Ed.* **2012**, *51* (26), 6369–6372.

(17) Roberts, R. J.; Cheng, X. Base Flipping. *Annu. Rev. Biochem.* **2002**, *67* (1), 181–198.

(18) Yin, Y.; Yang, L.; Zheng, G.; Gu, C.; Yi, C.; He, C.; Gao, Y. Q.; Zhao, X. S. Dynamics of Spontaneous Flipping of a Mismatched Base in DNA Duplex. *Proc. Natl. Acad. Sci.* **2014**, *111* (22), 8043–8048.

(19) Hollis, T.; Ichikawa, Y.; Ellenberger, T. DNA Bending and a Flip-out Mechanism for Base Excision by the Helix-Hairpin-Helix DNA Glycosylase, Escherichia Coli AlkA. *EMBO J.* **2000**, *19* (4), 758–766.

(20) Nakatani, K.; Hagihara, S.; Goto, Y.; Kobori, A.; Hagihara, M.; Hayashi, G.; Kyo, M.; Nomura, M.; Mishima, M.; Kojima, C. Small-Molecule Ligand Induces Nucleotide Flipping in (CAG)n Trinucleotide Repeats. *Nat. Chem. Biol.* **2005**, *1* (1), 39–43.

(21) Granzhan, A.; Kotera, N.; Teulade-Fichou, M. P. Finding Needles in a Basestack: Recognition of Mismatched Base Pairs in DNA by Small Molecules. *Chem. Soc. Rev.* **2014**, *43* (10), 3630–3665.

(22) Mukherjee, S.; Dohno, C.; Asano, K.; Nakatani, K. Cyclic Mismatch Binding Ligand CMBL4 Binds to the 5′-T-3′/5′-GG-3′ Site by Inducing the Flipping out of Thymine Base. *Nucleic Acids Res.* **2016**, *44* (15), 7090–7099.

(23) Liu, Y.; Wilson, S. H. DNA Base Excision Repair: A Mechanism of Trinucleotide Repeat Expansion. *Trends Biochem. Sci.* **2012**, *37* (4), 162–172.

(24) Lo, Y. S.; Tseng, W. H.; Chuang, C. Y.; Hou, M. H. The Structural Basis of Actinomycin D-Binding Induces Nucleotide Flipping out, a Sharp Bend and a Left-Handed Twist in CGG Triplet Repeats. *Nucleic Acids Res.* **2013**, *41* (7), 4284–4294.

(25) Magnani, J. L.; Peteroson, J. M.; Sarkar, A. K.; Vohra, T.; Baek, M.-G. Galactopyransoyl-Cyclohexyl Derivatives as E-Selectin Antagonists, 2018.

(26) Arstad, E. Process for Producing Radiohalgoenated Bioconjugates and Products Thereof, 2012.

(27) Srinivasachari, S.; Fichter, K. M.; Reineke, T. M. Polycationic β-Cyclodextrin "Click Clusters": Monodisperse and Versatile Scaffolds for Nucleic Acid Delivery. *J. Am. Chem. Soc.* **2008**, *130* (14), 4618–4627.

(28) Rostovtsev, V. V; Green, L. G.; Fokin, V. V; Barry Sharpless, K. A Stepwise Huisgen Cycloaddition Process: Copper (I) Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes. *Angew. Chem.* **2012**, *114* (14), 2708–2711.

(29) Siboni, R. B.; Nakamori, M.; Wagner, S. D.; Struck, A. J.; Coonrod, L. A.; Harriott, S. A.; Cass, D. M.; Tanner, M. K.; Berglund, J. A. Actinomycin D Specifically Reduces Expanded CUG Repeat RNA in Myotonic Dystrophy Models. *Cell Rep.* **2015**, *13* (11), 2386–2394.

(30) Ledig, K. W. 7-(SUBSTITUTED)-7H-PYRROLO(3,2-F]QUINAZOLINE-1,3-DIAMINES, 1978.

(31) Guan, J.; Zhang, Q.; O'neil, M.; Iii, N. O.; Ager, A.; Gerena, L.; Lin, A. J. Antimalarial Activities of New Pyrrolo[3,2-f]Quinazoline-1,3-Diamine Derivatives. *Antimicrob. Agents*

Chemother. **2005**, *49* (12), 4928–4933.

(32) Chen, J.; Kassenbrock, A.; Li, B. X.; Xiao, X. Discovery of a Potent Anti-Tumor Agent through Regioselective Mono-N-Acylation of 7H-Pyrrolo[3,2-f]Quinazoline-1,3-Diamine. *Medchemcomm* **2013**, *4* (9), 1275–1282.

(33) Vichai, V.; Kirtikara, K. Sulforhodamine B Colorimetric Assay for Cytotoxicity Screening. *Nat. Protoc.* **2006**, *1* (3), 1112–1116.

CHAPTER 3: REVERSIBLE TEMPLATE-SELECTED LIGATION

3.1 INTRODUCTION

Much of the genomic DNA that is transcribed into RNA in the human body is never translated into protein, and thus must have some other function or benefit to remain in the genome.[1] As a result of recent advances in understanding the roles of both coding and non-coding RNAs in disease, DNA- and RNA- targeting strategies have come into focus for development of therapeutics. Recent drug discovery efforts have enabled selective targeting of DNA and RNA sequences through many methods including antisense oligonucleotides,[2] CRISPR/Cas9 genome editing,[3] siRNA[4] and miRNA[5]-based approaches, and small-molecules.[6–8] Several groups have explored the use of these and other RNA-targeting strategies for trinucleotide repeat diseases, in particular myotonic dystrophy type 1 (DM1).[9–15] DM1 is characterized by RNA gain-of-function and thus is an ideal model to study nucleic acid targeting.

As described in Chapter 2, one ideal therapeutic strategy to target nucleic acids would be to dose with monomeric units that could be assembled *in situ* on the target nucleic acid template. Because multivalent agents tend to be more effective but also less cell permeable, *in situ* assembly provides an opportunity to overcome the challenge of cell permeability in multivalent targeting agents. Although click chemistry has been successful in forming dimers *in situ*, the products are formed irreversibly. An advantage of an irreversible approach is that product released from the nucleic acid target would stay assembled and could bind another unit of that target. However, a disadvantage is that the product is unable to diffuse out of the cell and thus could exhibit long-term cytotoxicity. Thus, it would be beneficial to develop a reversible assembly approach that would allow for recycling of monomer units which could be more easily cleared (Figure 3.1).

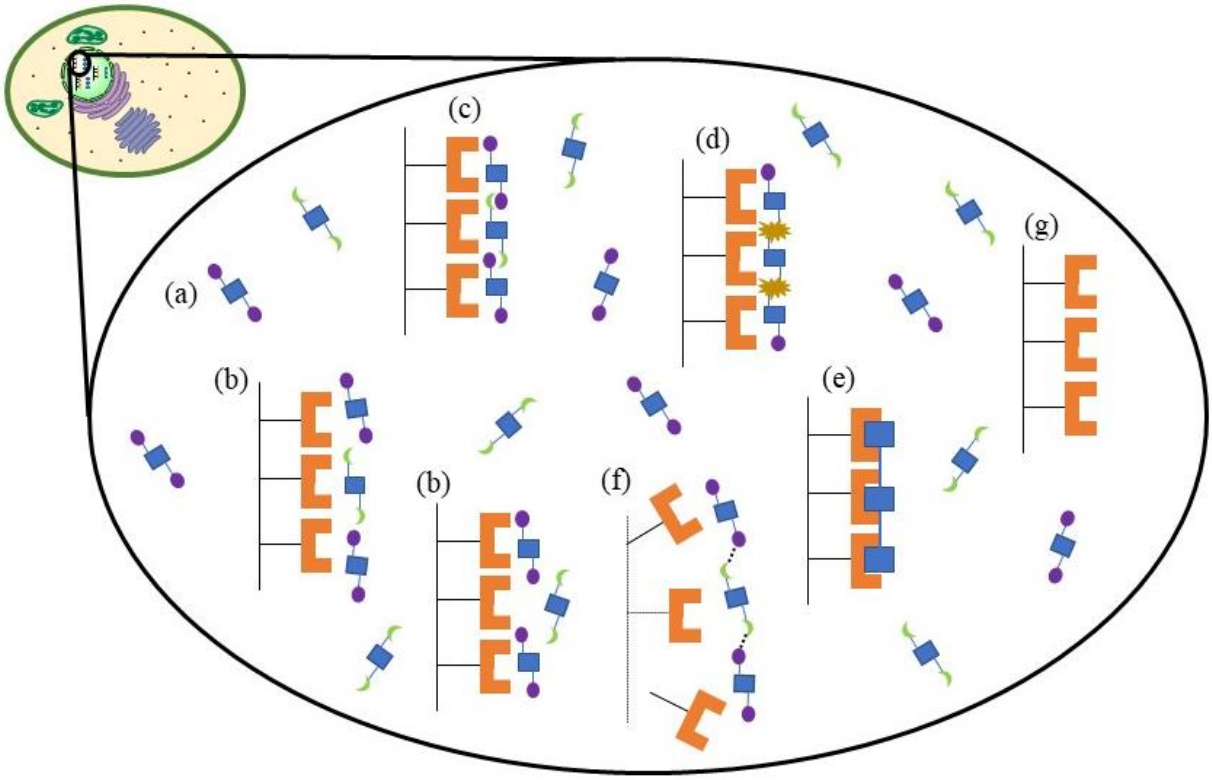


Figure 3.1. Dynamic Covalent Chemistry approach to reversibly form oligomers *in situ*. Orange receptors represent d(CTG)^{exp} or r(CUG)^{exp}, where each orange U-shape is a binding site. Blue boxes indicate small molecule monomers, and purple and green appendages indicate reactive moieties (e.g., aldehydes and amines). The idea behind this system is that: (a) many monomers are free in solution, but the concentration is dilute enough such that minimal ligation occurs off-template, (b) monomers are brought to an artificially high local concentration on the nucleic-acid template, (c) monomers come in close enough proximity to undergo a ligation reaction, (d) the reaction occurs, (e) the oligomeric, multivalent targeting agent sticks tightly to the template, (f) the nucleic acid template may be degraded or the product may dissociate, releasing monomers and (g) necessarily, some nucleic acid target remains unbound.

Herein, we employ imine condensation between an aldehyde and an amine *in situ* to reversibly assemble fragments on target (Figure 3.2). This is attractive from a therapeutic perspective because the fragments are effectively recyclable. Cellular degradation of the target nucleic acid or dissociation of the assembled product would lead to hydrolysis of the imine bond, thus degrading the oligomer to the original monomers in the absence of template (Figure 3.2). An additional benefit to this *in situ* assembly approach in the treatment of trinucleotide repeat diseases is that an oligomer synthesized *in situ* could have size determined by length of the repeat and thereby maximize binding to that sequence. Thus, a personalized treatment could be realized by dosing

with just a few monomers that could be used to synthesize an oligomer matching each patient's repeat length *in situ*. This strategy is particularly attractive for repeat diseases such as DM1.

DM1 is a neuromuscular disease caused by an expanded d(CTG·CAG)^{exp} repeat sequence in the 3'-untranslated region of the dystrophia myotonica protein kinase (*DMPK*) gene on chromosome 19[16] that is bidirectionally transcribed to form RNA with expanded repeats, i.e., r(CUG)^{exp} and r(CAG)^{exp}.[17] Healthy individuals have 5-35 repeats, whereas individuals who manifest symptoms of the disease typically have 80 to more than 2,500 repeats.[18] These expanded repeat transcripts result in a toxic RNA gain-of-function through splicing mis-regulation[19,20] and synthesis of toxic homopeptides through repeat-associated non-ATG (RAN) translation (Figure 3.3).[21]. By targeting the parent DNA repeat to inhibit bidirectional transcription and thus preventing the formation of toxic RNA, the effects of small molecule or oligomeric treatments could be utilized to ameliorate DM1 patient symptoms.

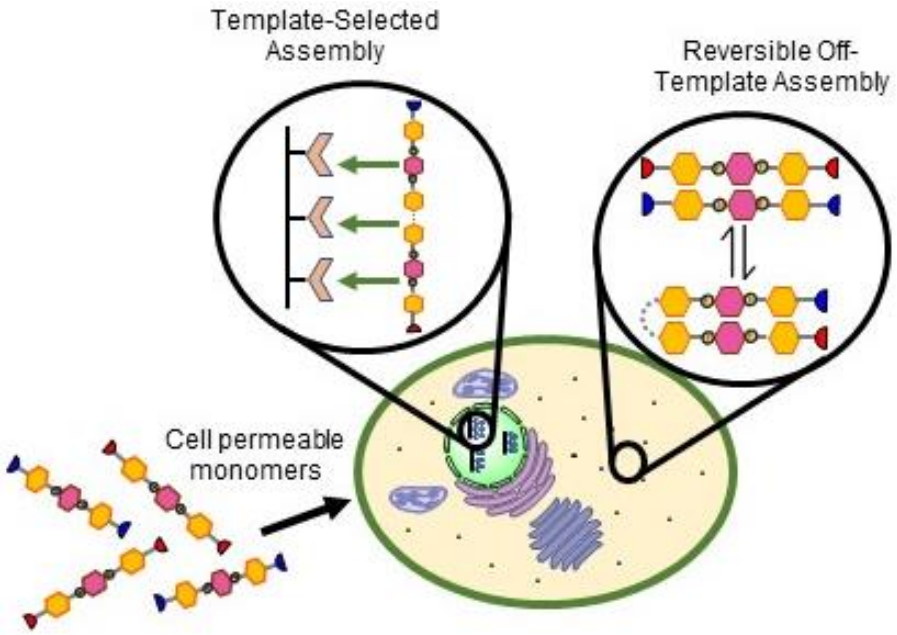


Figure 3.2. Dynamic Covalent Chemistry Template-Selected Assembly Approach. Cell-permeable monomers reversibly assemble on and off the template, allowing for template-selected amplification of the potent multivalent targeting agent.

Recent efforts in small molecule targeting of r(CUG)^{exp} secondary structures have suggested that multivalent targeting approaches might lead to more efficacious treatments for DM1 and other diseases.[22–24] Although the idea of using an oligomeric molecule to bind target DNA or RNA at multiple sites is promising, the challenges of this approach include difficult synthesis of discrete oligomers molecules and problems with cell permeability.[24] Therefore, the development of a multivalent targeting agent that is both synthetically feasible and cell-permeable would be beneficial for the treatment of DM1.[16,25]

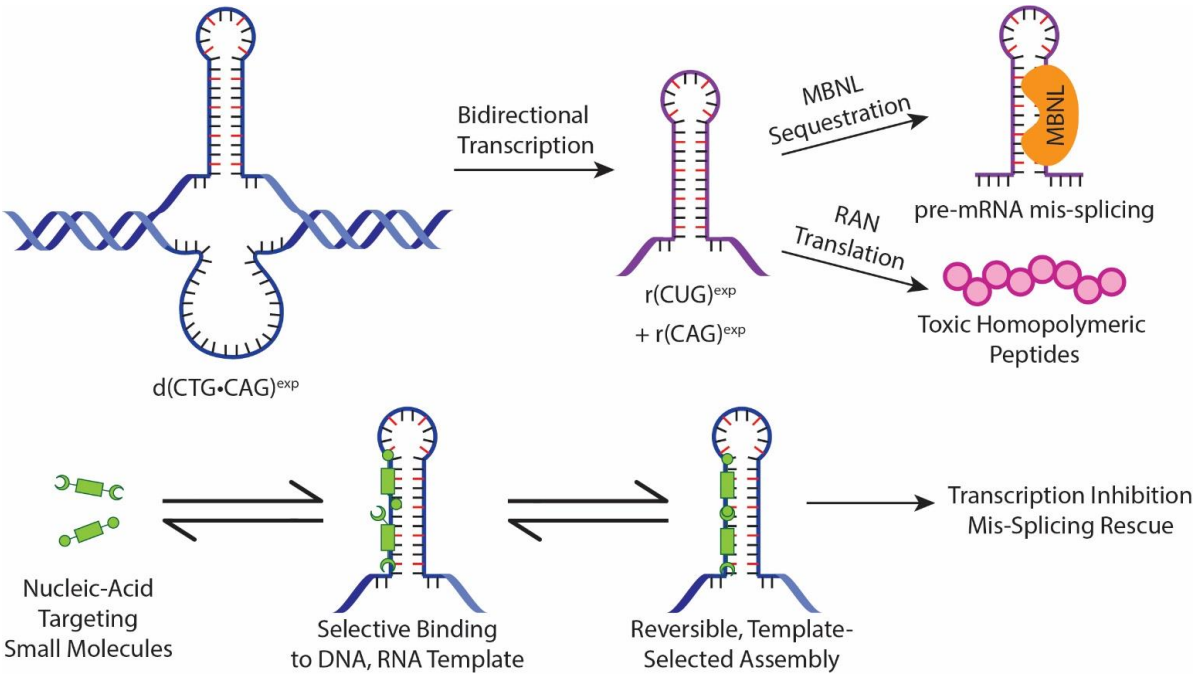


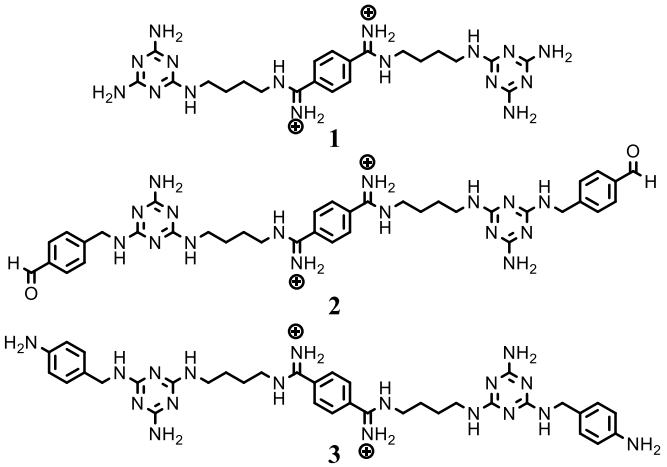
Figure 3.3. Myotonic Dystrophy Type 1 (DM1) Pathogenesis. d(CTG·CAG)^{exp} is bidirectionally transcribed to form r(CAG)^{exp} and r(CUG)^{exp}. The r(CUG)^{exp} can form hairpin secondary structures that sequester MBNL1 proteins and lead to improper splicing of pre-mRNAs. Both RNA transcripts can undergo RAN translation to yield toxic homopolymeric peptides such as polyglutamine. The use of reactive small-molecules to specifically target RNA and DNA and assemble on-template could alleviate symptoms of DM1 by preventing bidirectional transcription of d(CTG·CAG)^{exp} and/or competitively binding the r(CUG)^{exp} hairpin structure to release MBNL1 and prevent RAN translation.

3.2 RESULTS & DISCUSSION

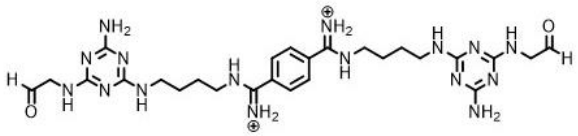
3.2.1 Compound Design

We previously reported that compound **1** selectively binds to r(CUG)^{exp} and shows low toxicity and high cell permeability.[26] A dimeric version of compound **1** proved more effective than other dimeric compounds (e.g., acridine dimers), but had low cell permeability.[27] Both **1** and its dimeric analogue selectively bound the RNA target with low cytotoxicity, but showed no affinity to the target DNA. Although RNA is a suitable target, a potentially more powerful therapeutic strategy would target upstream of the toxic RNA, at the DNA level, inhibiting the formation of the toxic RNA altogether. Thus, we sought to develop bisamidinium-based DNA targeting agents that would be more effective as therapeutic agents while taking advantage of the low cytotoxicity of the groove-binding scaffold.

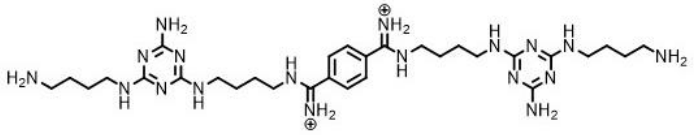
Based on the success of ligand **1** and its dimeric and oligomeric derivatives,[26–28] analogues using this scaffold were designed with benzaldehyde and aniline functionalities that may undergo a condensation reaction *in situ* to form imines. The triazine moieties serve as T-T and U-U mismatch recognition units and the bisamidinium groove-binding linker is designed to bind in the major groove of the RNA helix. In the docked structure of **1**, each ligand spans a total of 3 U-U mismatches. Monomers **2** and **3** were designed such that the aldehyde and amine could bind in adjacent sites on the DNA or RNA template and react, forming a more potent, multivalent targeting agent. Notably, oligomers of any length that match the repeat sequence could form.



The initial design of these compounds was also based on the scaffold of ligand **1**. Analogues were designed with aldehyde and free-amine functionalities that may undergo a condensation reaction *in situ*. Ligands **4** and **5** were designed for this dynamic covalent chemistry approach. Upon *in situ* assembly, these monomers could form an oligomer that could extend to match the length of the repeat sequence to which it is bound. To study the potential for binding and proximity-induced condensation, ligands **4** and **5** were docked on r(CUG)$_6$ (PDB ID: 3gm7)[29] in MOE. These ligands were docked individually on the sequence and the distance between reactive partners (the aldehyde carbonyl and the amine nitrogen) was measured to be 3.13 Å (Figure 3.10). A ligation was manually performed by linking the two ligands and the structure was again minimized. Based on this modeling study, the linker length would be sufficient if each ligand spans 3 U-U mismatches, which is the same mode of binding hypothesized for ligand **1**.

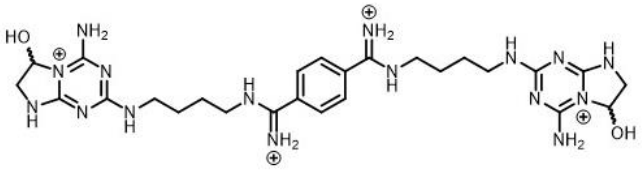


4



5

The synthetic routes to ligands **4** and **5** are detailed in Scheme 3.3 and 3.4, respectively. Interestingly, upon deprotection of the acetal precursor to obtain aldehyde **4**, hemiaminal **6** forms. This product has been extensively characterized and is believed to form as a result of attack of the internal triazine amine on the aldehyde carbonyl. A previously reported analogue[30] was synthesized and spectral evidence suggests that this structure also forms the cyclic product. Although the aldehyde reactivity is hindered in this form, it was thought that the hemiaminal might provide *in situ* protection



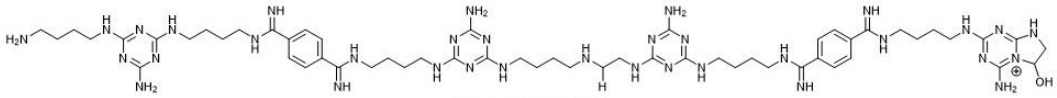
6

for the reactive aldehyde group, overcoming potential off-target reactions. For this template-assisted ligation to occur, the aldehyde form, **4**, would need to bind the nucleic acid target in proximity of diamine **5**, thereby shifting the hemiaminal-aldehyde equilibrium to allow additional proximity-induced ligations.

As a result of the formation of this hemiaminal, we wondered if **5** and **6** would be able to react on template. To address reactivity concerns with hemiaminal **6**, reactivity tests were performed 2,4-dinitrophenylhydrazine (Figure 3.11 and 3.12). A quantitative yield was observed for this reaction. Having established the ability of the hemiaminal to undergo condensation with highly reactive DNP-hydrazine, the reactivity was further explored in the context of *in situ* assembly. The analysis of product formation on a nucleic-acid template proved challenging, and the following assay was designed to monitor the formation of imine condensation products on a nucleic acid template. Ligand monomers were incubated with or without nucleic acid template in an aqueous buffer mixture. Aliquots of those reactions were removed after 1, 3, and 7 d. To these aliquots was added sodium cyanoborohydride ($NaCNBH_3$), a gentle reducing agent that should convert imine products to secondary amines that can be further characterized without reducing the aldehyde. To avoid sequestration of the product by the template, DNase I or RNase H, A, or I_f (NEB) were added to the aliquot to degrade the nucleic acid, ideally leaving the free oligomeric products or monomeric starting materials in the sample. MALDI-MS was used to screen for new peaks. For samples in which new peaks were observed, the aliquot would be passed through a silica spin column to remove excess $NaCNBH_3$ and analyzed via HPLC to quantify percent conversion. For this assay to be feasible, it is imperative that the DNase or RNase remains active in the presence of the reducing agent, $NaCNBH_3$. Thus, several control experiments were performed to confirm that these reagents are compatible. A sample of each DNA and RNA target was incubated with

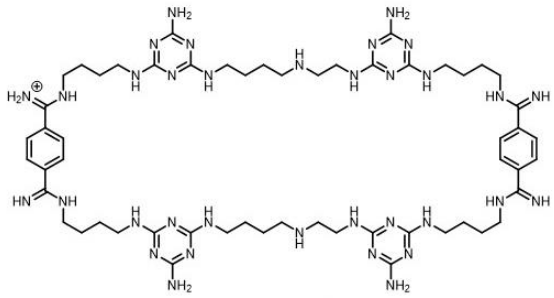
$NaCNBH_3$ with and without DNase I or RNase A, H, or I_f (each separately) at 37 °C for 20 min. The samples were analyzed by gel electrophoresis, and the DNase as well as all three RNases remained active for all sequences in the presence of $NaCNBH_3$ (Figure 3.13).

MALDI-MS screening was performed with several different conditions. Targets tested included $d(CTG)_{16}$, $d(CTG)_{32}$, $r(CUG)_{16}$, and $r(CUG)_{90}$ at concentrations of 10 μM or 20 μM. Compounds were tested at concentrations 10 μM, 20 μM, 25 μM, 50 μM, and 100 μM each. RNase H, RNase A, and RNase I_f were all tested in this assay. No new peaks from condensation products were observed at concentrations lower than 100 μM. At 100 μM, the only new peaks observed were from the dimeric product **7** (*m/z* 1256.58, 8) as well as a new peak at *m/z* 1239.9 (Figure 3.14). These peaks were observed both with and without the template, as early as 1 d. Although 1239.9 does not match the expected dimeric product, it does match the mass of macrocycle **8**. This macrocycle may form on the $r(CUG)^{exp}$ template, but because it was observed in both the control (no template) and the RNA-containing reactions, it is thought that the condensation reaction occurs off-template due to the higher concentration of monomers. The dimeric product **7** could undergo an intramolecular condensation reaction with the aldehyde to form macrocyclic product **8**, potentially explaining the presence of this new peak.



Molecular Weight: 1256.58

7



Exact Mass: 1239.82

8

These results suggest that the hemiaminal equilibrium may slow or inhibit the condensation reaction with the amine, preventing the desired self-assembly. 1,4-Dihydrazino-1,4-butanedione was also attempted as a condensation partner with the same conditions described above, but no oligomeric products were observed. It is possible that the distance between the free amine and the aldehyde carbonyl once bound to the template (about 3 Å) is too large to allow for the ligation reaction interaction. Thus, these compounds were not further studied as condensation partners and future efforts focused on **2** and **3**. Moving forward, other reversible ligation reactions that may be considered include noncovalent interactions,[31] disulfide exchange,[32] and native chemical ligation.[33]

3.2.2 Computational Study of Assembly of 2 and 3

After designing compounds **2** and **3** as analogues of **1**, we wondered if it was computationally feasible for these compounds to assemble. Thus, ligands **2** and **3** were docked onto the crystal structure of r(CUG)$_6$ (PDB ID: 3gm7)[29] in Molecular Operating Environment (MOE). These

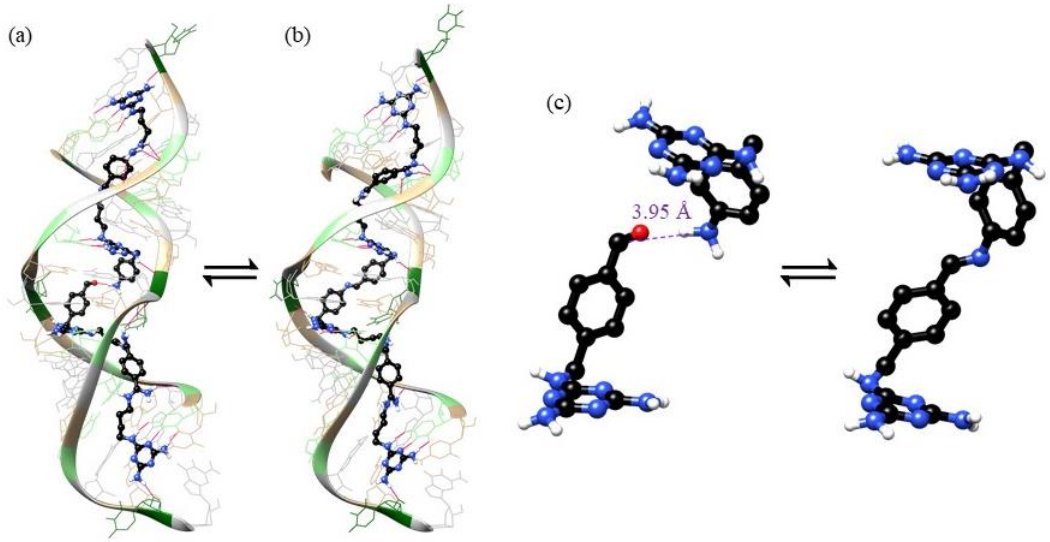


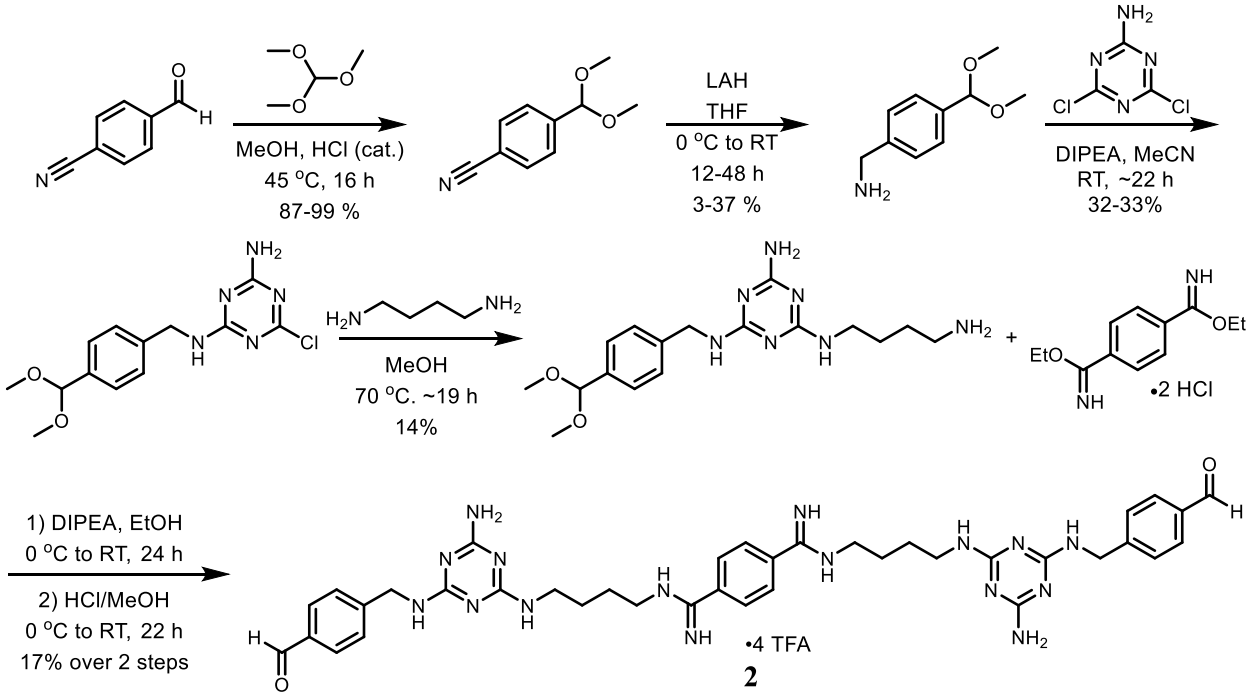
Figure 3.4. MOE docking studies for ligands 2 and 3. (a) Monomers **2** and **3** docked side by side. The fragments were manually ligated, and the structure minimized to form the dimeric product **2+3** on the r(CUG)$_6$ helix shown in (b). Uracil is green (light green are flipped in, dark green flipped out), thymine is tan and guanine is grey. Hydrogen bonds between ligand and RNA are shown in red. r(CUG)$_6$ PDB ID: 3gm7.[29] The distance between reactive atoms (from carbonyl carbon to aniline nitrogen) was measured to be 3.95 Å (purple dotted line in c.) Docking performed with Hyeoun (Heather) Jeon.

ligands were docked individually on the sequence with successive minimizations and the distance between reactive partners (the carbonyl carbon of the aldehyde and the aniline nitrogen) was measured to be 3.95 Å (Figure 3.4). Of note, a hydrogen bond between the carbonyl oxygen and the aniline nitrogen holds these moieties in close proximity. A ligation was performed manually by linking the two ligands, the structure was again minimized, and the distance between the reactive units was measured. Additional images of the docked structures are included in Figure 3.15. Based on this modeling study, the linker length and distance between reactive groups is sufficient for binding and ligation with each ligand spanning 3 U-U mismatches, consistent with the mode of binding hypothesized for ligand **1**.

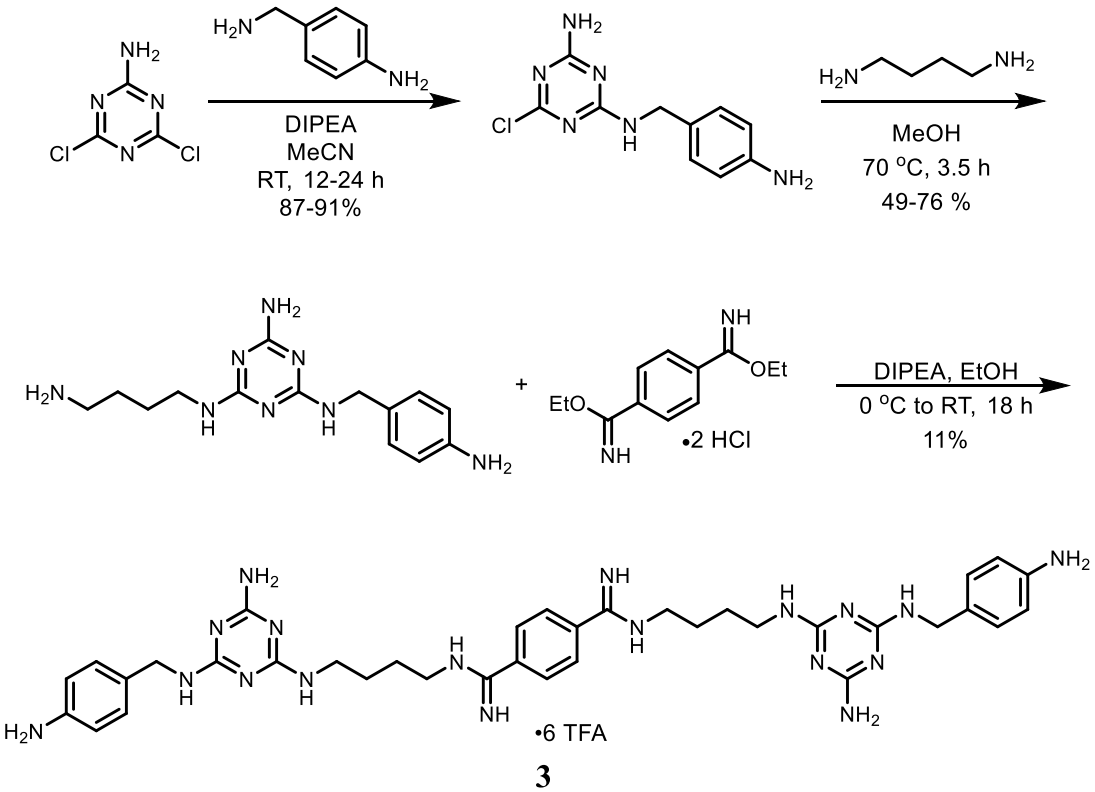
3.2.3 Synthesis of Compounds

Compounds **2** and **3** were synthesized according to the methods in Scheme 3.1 and Scheme 3.2, respectively. In the synthesis of **2**, 4-cyanobenzaldehyde was protected with trimethyl orthoformate and the nitrile group was reduced to a primary amine with lithium aluminum hydride. The resulting amine was reacted with melamine derivative 4,6-dicholoro-1,3,5-triazin-2-amine. The disubstituted triazine product was functionalized with diaminobutane. Reaction with diethyl terephthalimidate followed by deprotection yielded compound **2**.

Compound **3** was synthesized via reaction of 4-aminobenzylamine and 4,6-dicholoro-1,3,5-triazin-2-amine to form a disubstituted triazine that was functionalized with diaminobutane. Reaction with diethyl terephthalimidate yielded compound **3**. Both **2** and **3** were purified via preparatory HPLC in acetonitrile:water gradient with 0.1% TFA. Detailed methods and characterization data are included in the Appendix A, Section A.3.1.



Scheme 3.1. Synthesis of **2**. See Appendix A, Section A.3.1 for detailed methods.



Scheme 3.2. Synthesis of **3**. See Appendix A, Section A.3.1 for detailed methods.

3.2.4. MALDI-MS Analysis of Template-Selected Assembly

With compound **2** and **3** in hand, we wondered whether the benzaldehyde and aniline groups would react to assemble a dimeric and/or oligomeric product. Thus, monomers **2** and **3** were mixed in the presence or absence of template, incubated at 37 °C, and reduced with sodium cyanoborohydride. MALDI-MS analysis showed the presence of the monomers **2** and **3** in each sample as well as a peak consistent with fragmented dimer **2**+**3** in the presence of both d(CTG)$_{16}$ and r(CUG)$_{16}$ at *m*/*z* of 1193.8, but not in the absence of template (Figure 3.5). We observed

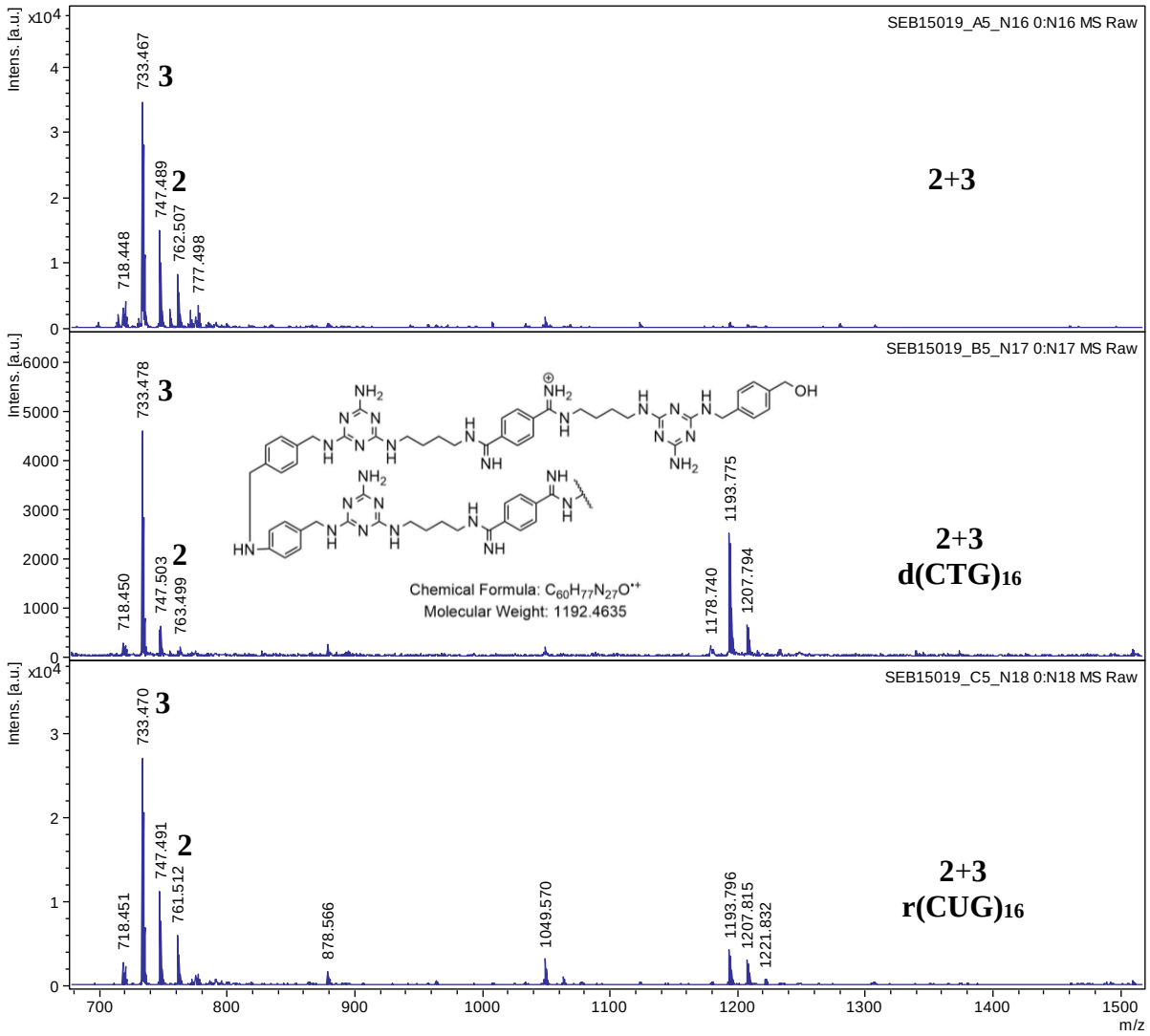


Figure 3.5. MALDI-MS analysis of **2+3** assembly on and off template with reduction by $NaCNBH_3$. Buffer: 2 mM each of KCl, $MgCl_2$, $CaCl_2$, and Tri-HCl, pH 7, incubated at 37 ºC for 75 min before reduction and analysis. 100 µM compounds, 10 µM d(CTG)$_{16}$ or r(CUG)$_{16}$. See procedure 3.4.1 for details.

formation of this product selectively on template after as few as 5 min (Figure 3.16). We were unable to characterize the reaction at a time faster than 5 min due to limitations of mixing the solution, sampling, reducing, and spotting for MALDI-MS. This assembly occurs not only selectively on the template, but also very quickly compared to previous assembly methods that take hours or even days to form product. Notably, any product that may form off-template in the aqueous *in situ* environment could be hydrolyzed, thus returning to the monomeric form. Therefore, we refer to this assembly as "template-selected." Due to partial reduction of the aldehyde by sodium cyanoborohydride, it was noted that in the MALDI-MS, monomer **2** and the dimeric product both exist in multiple reduction states. For example, monomer **2** could be present as the dialdehyde (M+H = 759.41), the aldehyde-alcohol (M+H = 761.42), or the diol (M+H = 763.44). The dimeric product could be present as an amine (reduced imine) with aldehyde end group (M+H = 1475.84) or an alcohol end group (M+H = 1477.86).

To determine if larger, oligomeric products could form, the compounds were incubated in the presence or absence of template at 37 °C for 24 h before MALDI-MS analysis. Trimer **2**+**3**+**2** and pentamer **2**+**3**+**2**+**3**+**2** were observed in the presence of d(CTG)$_{16}$, r(CUG)$_{16}$, and r(CUG)$_{90}$, but not in the absence of template (Figure 3.6). These longer oligomers support the potential to use this strategy to synthesize multivalent targeting agents *in situ*, selected by the nucleic acid template.

Because of the reactive functionalities on the monomers, we wondered if monomers **2** and **3** were reacting with the nucleic acid template to yield a covalent DNA modification. Thus, to check if monomer **2** or **3** reacted with the target DNA, MALDI-MS was used to screen for modification of d(CTG)$_3$ and a random duplex DNA target. No modification of these DNA sequences was observed (Figure 3.17), suggesting that interactions are non-covalent.

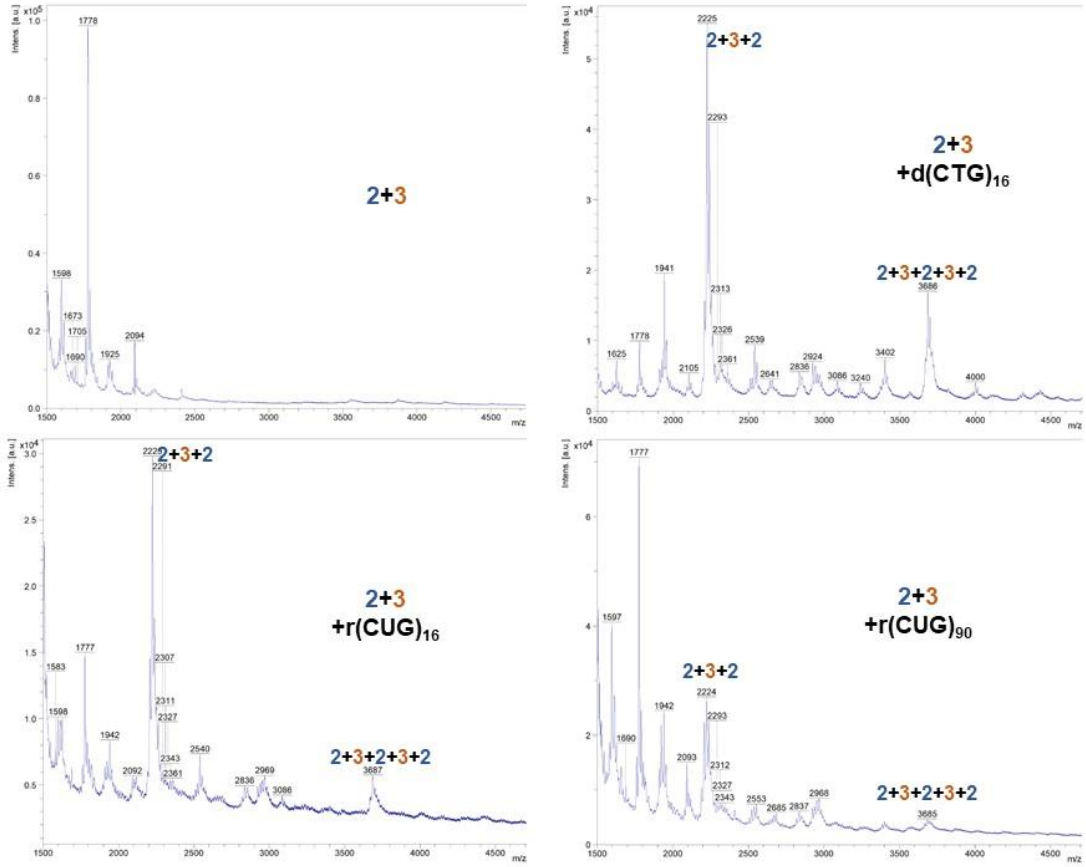


Figure 3.6. High MW MALDI-MS analysis of **2+3** assembly on and off template with reduction by $NaCNBH_3$. Buffer: 2 mM each of KCl, $MgCl_2$, $CaCl_2$, and Tri-HCl, pH 7, incubated at 37 °C for 24 h before reduction with sodium cyanoborohydride, degradation of template, and analysis. Compound concentration 100 μM, 10 μM $d(CTG)_{16}$ and $r(CUG)_{16}$ or 500 nM $r(CUG)_{90}$. See Procedure S2 for details.

3.2.5 HPLC Analysis of Template-Selected Assembly

After observing assembly by MALDI-MS, we utilized HPLC to screen for the appearance of new peaks that could be consistent with dimer or oligomer formation. As MALDI-MS is not quantitative, we were particularly interested in whether the trimer and the pentamer would be more prominent in the presence of $r(CUG)_{90}$ and the absence of template, as the length of the template may affect the length of the product that is formed.

For HPLC analysis, **2** and **3** were incubated in the presence or absence of template at 37 °C for 24 h before reduction with sodium cyanoborohydride and analysis by HPLC. A new peak appeared in the HPLC around 6.1 min, between the retention time of monomer **2** (6.6 min partially reduced, 6.4 min fully reduced) and monomer **3** (5.2 min) in the presence of all three templates ($d(CTG)_{16}$, $r(CUG)_{16}$, and $r(CUG)_{90}$) and the absence of template (Figure 3.18). Of note, two other new peaks formed around 7.0 and 7.5 min. Although the identity of these three new peaks has not yet been confirmed, the emergence of three new products is consistent with the observation of dimer, trimer, and pentamer in the MALDI-MS. Interestingly, the peak at 7.0 min made up 6.6% and 8.2% of the total integrated spectrum in the $r(CUG)_{16}$ and $r(CUG)_{90}$ experiments, respectively. A similar trend was observed for the peak at 7.5 min, making up 9.3% and 18.3%, respectively in the $r(CUG)_{16}$ and $r(CUG)_{90}$ experiments. The opposite trend is observed in the peak at 6.1 min, that makes up 15.5% and 10.0% in the $r(CUG)_{16}$ and $r(CUG)_{90}$ experiments, respectively. Thus, these peaks with higher retention time seem to be more prominent in the presence of the longer template and the peak at lower retention time seems to be more prominent in the presence of the shorter template. This could suggest that the peak at 6.1 minutes is the dimeric product that would form on the shorter template more favorably than the longer oligomers. Work to identify these peaks is ongoing, with efforts to synthesize the discrete dimeric and oligomeric products *ex vivo* and to isolate the HPLC peaks for direct characterization by MALDI-MS.

LC/MS was attempted to quantify the relative amounts of monomer, dimer, trimer, and larger oligomers in the solution. We expected that LC/MS would also allow for definitive assignment of the identity of new peaks observed in the HPLC. However, the results of the LC/MS were inconclusive due to the large amount of fragmentation and the high abundance of multiply charged peaks. Each new peak contained both **2** and **3** in the mass spectrum, consistent with an assembled

product that contains **2** and **3** but not conclusive in assigning which peak belonged to each discrete product. We are hopeful that switching to the gentler ionization source of MALDI-MS will allow for positive identification of the HPLC peaks.

In addition to supporting the formation of a new dimeric and oligomeric product as observed by MALDI-MS, these HPLC studies indicate that an amine product can be characterized from the template-selected assembly of an imine product. As expected, in a control assay with no reductant added, only monomers and no product were observed in the HPLC and only a very small peak consistent with the dimeric imine product was observed by MALDI-MS, suggesting that the imine product can be hydrolyzed in aqueous conditions to re-form the monomers and consistent with the fragility of imine fragmentation in MALDI-MS without reduction (Figure 3.20).

3.2.6. Isothermal Titration Calorimetry

ITC studies were performed to test the binding affinity of each monomer compared to the mixture of monomers, or transient dimeric/oligomeric compound with d(CTG)$_{16}$ and random duplex DNA. As expected, the calculated K_d value for the interaction between **2**+**3** and d(CTG)$_{16}$ was lower than the K_d for both monomeric units. The calculated K_d value for **2**+**3** was 5.6 ± 0.5 µM, whereas the K_d for **2** was calculated to be 81.6 ± 0.9 µM and no significant binding was observed for **3** (Figure 3.7, 3.22 a,b). Compound **2**, **3**, and **2**+**3** all showed no significant binding to a random duplex DNA sequence (Figure 3.22

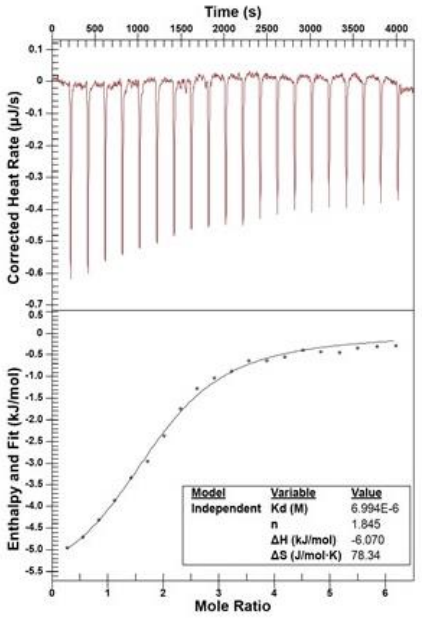


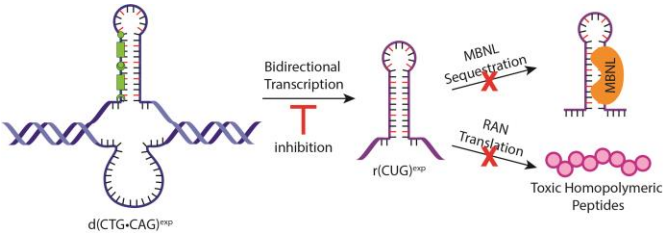
Figure 3.7. Isothermal Titration Calorimetry with **2+3**. [Cell (nucleic acid)] = 20 µM, [Syringe (ligand)] = 500 µM, Buffer = 300 mM NaCl, 20 mM MOPS. See Section 3.4.6 for experimental details.

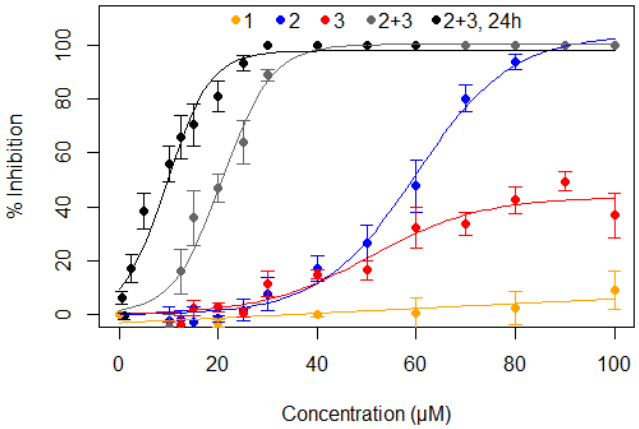
c-e). The thermodynamic data for these binding interactions is shown in Table 3.1. Of particular interest for the treatment of DM1, **2**+**3** also binds to r(CUG)$_{16}$ with a calculated K_d value of 4.3 ± 0.5 μM and **2**+**3** does not bind random ssRNA (Figure 3.23).

These binding studies suggest that the transient imine product binds more strongly to a d(CTG)$_{16}$ target than the monomers themselves and neither the monomers nor **2**+**3** bind to a random DNA duplex. Further, **2**+**3** can also bind to r(CUG)$_{16}$ as a potential dual-targeting agent for DM1 and does not bind to random ssRNA.

3.2.7. in vitro Transcription Inhibition.

Because we are interested in developing a DNA-targeting agent that would prevent the formation of toxic RNA and the associated disease pathobiology, we sought to test if monomers **2** and **3** could inhibit transcription of a DM1-relevant DNA template either independently or cooperatively through *in vitro* transcription inhibition (Figure 3.8). Although compound **1** shows no appreciable inhibition at these concentrations, analogues **2** and **3** both showed a higher level of inhibition than parent compound **1**, with **2** showing an





Ligand	Calc'd IC_{50}
2	59.2 ± 0.5 μM
3	>100 μM
2+3	20.8 ± 0.6 μM
2+3, 24 h	9.8 ± 0.5 μM

Figure 3.8. *in vitro* Transcription Inhibition Assay with T7 RNA Polymerase and pSP72 plasmid containing 90 (CTG·CAG) repeats. Results are reported as the average of at least 3 independent experiments are error is reported as standard error of the mean.

IC_{50} value of 59.2 ± 0.5 μM. Because the imine is transient by design, it is difficult to quantify its presence in solution. Thus, to test the mixture of **2**+**3**, equimolar amounts of the compounds were mixed in the treatment such that 20 μM indicates the presence of 20 μM of each monomer, or 20 μM of the dimeric imine that can form *in situ*. To compare this mixture (**2**+**3**) to the inhibition of the monomeric parts, an arithmetic sum of the individual monomer inhibition is shown on the graph in a green line. Notably, the calculated IC_{50} value for the mixture of **2**+**3** was 20.8 ± 0.6 μM (grey line). With an extended incubation period of 24 h, the IC_{50} value for **2**+**3** decreased to 9.8 ± 0.5 μM. For comparison, the arithmetic sum of the independent inhibition of **2** and **3** has a calculated IC_{50} value of 49.2 μM. This data is consistent with the K_d values observed via ITC.

As the assembly of the imine presumably occurs *in situ*, we wondered if providing a 24 h pre-incubation window during which **2** and **3** were incubated with the template before addition of the polymerase would allow for stronger inhibition. As the previous incubation window was 3.5 h, we hypothesized that providing longer incubation would allow for increased template-selected formation of imine and thus stronger inhibition. Indeed, The IC_{50} for **2**+**3** with 24 h incubation (black line) dropped to 9.6 ± 0.5 μM. Of note, the combination of **2**+**3** showed no inhibition in the transcription of a random duplex sequence (Figure 3.24). Together, these results suggest that the mixture of **2**+**3** is effective at inhibiting transcription *in vitro.*

Although a modest improvement was observed here for **2**+**3** compared to the monomeric units, we observed an almost 1,000-fold improvement with our groove-binding dimer compared to the monomeric ligand.[27] Importantly, the dimers studied here are formed reversibly *in situ*, compared to the previous approach in which the dimer was synthesized *ex vivo* and dosed as an assembled unit. As mentioned above, there are many benefits to reversible *in situ* assembly, but the challenge of bringing together the monomeric parts on the target may decrease the efficacy of the dimeric or

oligomeric assembled products. Further, it is important to note that these methods target two different species. Our previously reported groove binders targeted r(CUG)^{exp} but showed no affinity to d(CTG)^{exp}, whereas **2**+**3** binds both causative agents. Thus, the direct comparison of the groove-binding dimer and **2**+**3** is complicated by their different targets.

3.2.8. Insulin Receptor Mis-Splicing Rescue.

The biological activity of **2**+**3** was further investigated via a cellular mis-splicing assay. Prior to this assay, sulforhodamine B cytotoxicity assays suggested that the monomers and the combination of **2**+**3** were not toxic up to a concentration of 200 µM (Figure 3.25). In DM1 patients, sequestration of MBNL1 by r(CUG)^{exp} prevents proper splicing of the insulin receptor (IR) pre-mRNA, leading to insulin resistance due to aberrant splicing regulation.[34] Healthy individuals predominantly express isoform B, with exon 11 included, but DM1 patients predominantly express isoform A, a lower-signaling, non-muscle isoform. Detailed methods for this splicing rescue assay can be found in the SI. Briefly, HeLa cells were transfected

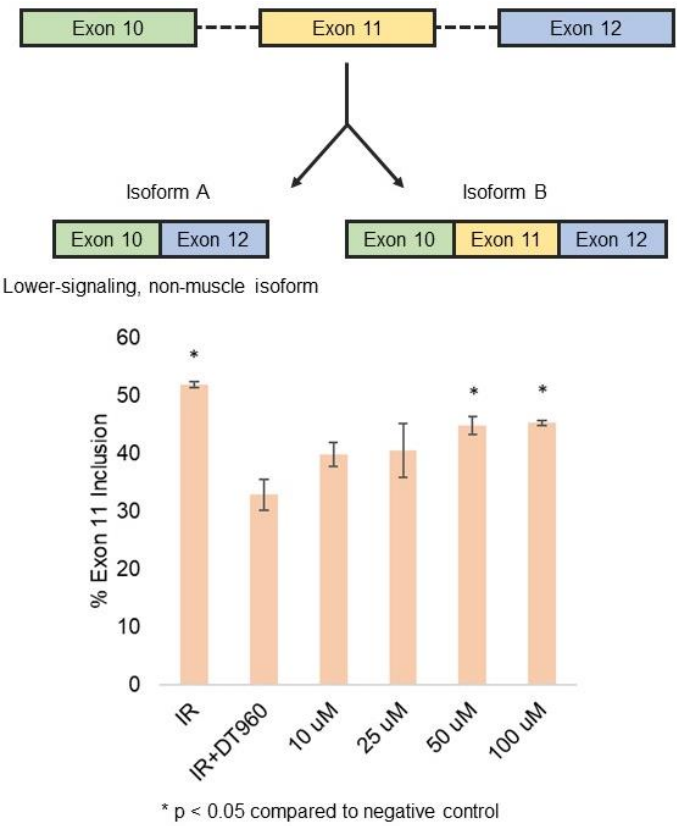


Figure 3.9. Insulin Receptor Mis-splicing Assay. Combination of ligands **2** and **3** rescued mis-splicing in the insulin receptor minigene in a dose-dependent manner. Results are reported as the average of at least 3 independent replicates. Error bars indicate standard error. * indicates $p < 0.05$ compared to the negative control.

with the IR minigene with or without the DT960 minigene, which carries 960 d(CTG·CAG) repeats. Cells were treated with varying concentrations of compound ranging from 10 μM to 100 μM and compared to the positive control (IR without DT960, healthy) and the negative control (IR with DT960, diseased). A dose-dependent splicing rescue was observed (Figure 3.9, 3.26), where 65% splicing recovery was achieved at 100 μM. Of note, monomers **2** and **3** on their own did not achieve any significant mis-splicing rescue up to 100 μM (Figure 3.27). As IR mis-splicing is one of the most difficult to rescue,[35] this strong splicing improvement coupled with the *in vitro* transcription inhibition suggests that the formation of a transient imine product *in situ* could be a viable strategy to treat DM1 or other trinucleotide repeat diseases.

3.3 CONCLUSIONS

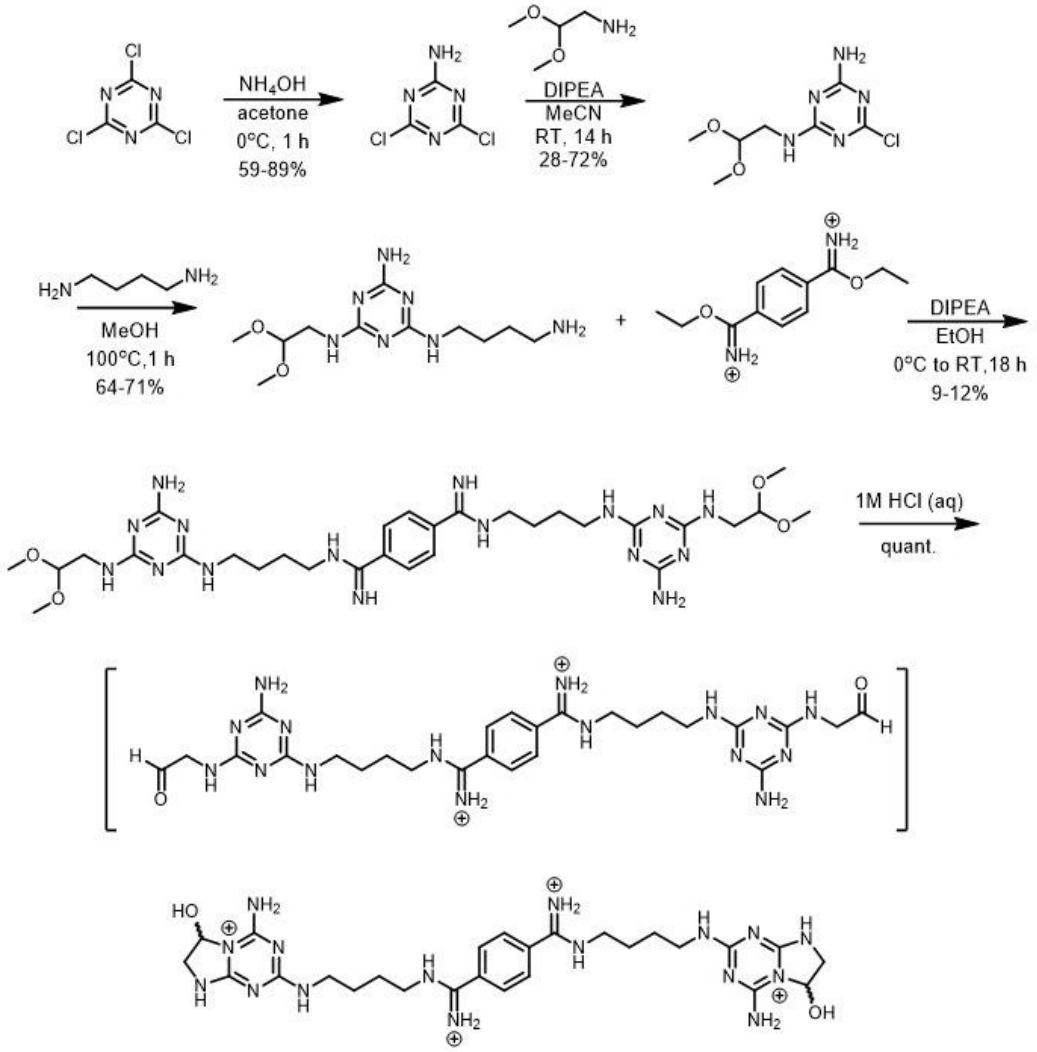
Herein we reported evidence for a template-selected, reversible conjugation of nucleic acid-targeting benzaldehyde **2** and aniline **3**. Computational modeling studies suggest that these compounds can be brought into close proximity when bound to adjacent sites on a nucleic acid template and can link together via a condensation reaction to form a more potent targeting agent. As determined by HPLC and MALDI-MS, monomers **2** and **3** can react to form an imine dimer **2**+**3** as well as oligomeric products that are hydrolyzed in aqueous conditions. The putative imine product bound $d(CTG)_{16}$ and $r(CUG)_{16}$ with a low micromolar K_d, inhibited transcription of $d(CTG{\cdot}CAG)_{90}$ to form $r(CUG)_{90}$, and partially rescued IR mis-splicing. We anticipate that this reversible assembly strategy could be broadly applied in drug discovery efforts to reversibly synthesize multivalent targeting agents *in situ*.

3.4. METHODS

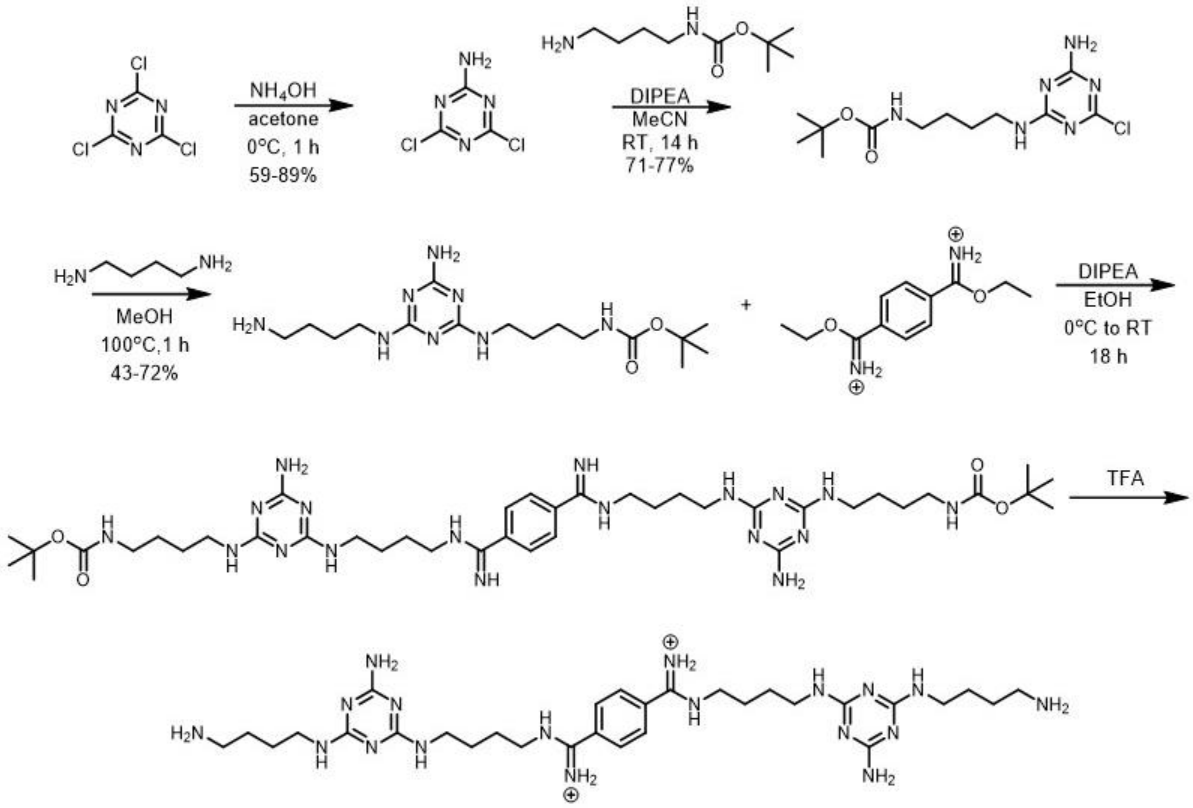
3.4.1. Computational Docking Method

The crystal structure of r[$(CUG)_6$]$_2$ was obtained from the Protein Data Bank (code 3gm7)[29] and extended to r[$(CUG)_7$]$_2$ using the Nucleic Acid Builder in Molecular Operating Environment (MOE, v2018.01). Uracil (U) bases (from U-U mismatches denoted 1, 3, 4, and 6) were manually rotated out of the helix to simulate base flipping. Each melamine was docked with 3 hydrogen bonds with the Watson-Crick-Franklin face of the undisturbed U and upon minimization, an additional hydrogen bond was observed with the phosphate backbone of the flipped U base. Bisamidinium units were placed in the major groove of the helix and sequentially minimized. The melamine and bisamidinium units were linked with an alkyl group consisting of four methylene units and minimized, with melamines at U-U mismatch 1 and 3 linked to a bisamidinium and those at U-U mismatch 4 and 6 linked, to represent monomers **2** and **3**. Upon minimization, the aniline and benzaldehyde functionalities were built onto internal melamine at uracil 3 and 4, respectively. The structure was minimized to obtain the final docked structure of the monomers. The distance between the carbonyl carbon and aniline nitrogen was measured using USCF Chimera. A manual ligation was performed by linking the carbonyl carbon to the aniline nitrogen atom and by deleting the oxygen atom that would be lost as a water molecule. The docked structure was minimized to yield the final docked structure of the imine dimer. Final docked structures were exported as protein data bank (pdb) files. Images were rendered and hydrogen bonds (H-bonds) were quantified[36] in UCSF Chimera (v1.15.1).[37] Docking studies performed with Heather Jeon.

3.4.2. Supplemental Data for Initial Compound Design



Scheme 3.3. Synthetic route to hemiaminal 6. Notably, upon deprotection of the acetal, the cyclized aldehyde product was formed. For experimental details, see Appendix A, Section A.4.1.

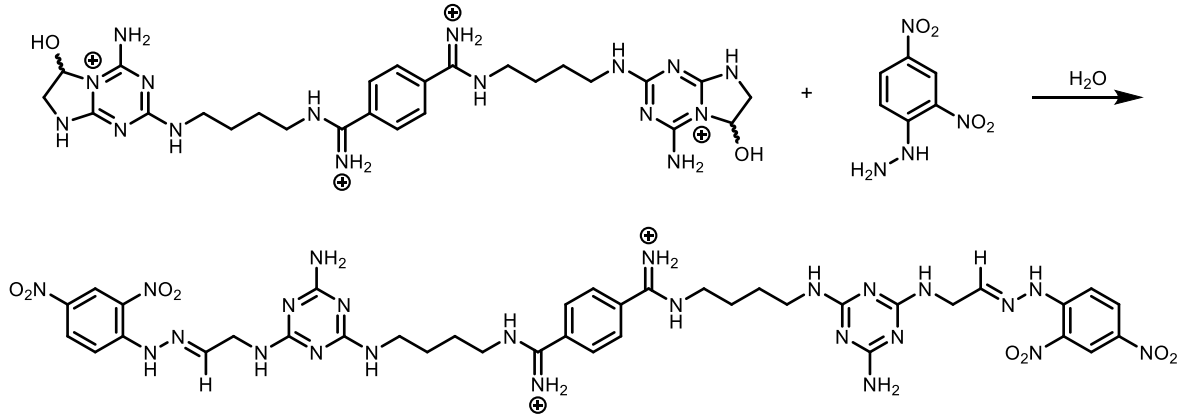


Scheme 3.4. Synthetic route to ligand 5. For experimental details, see Appendix A, Section A.4.1.

Hemiaminal Reactivity Tests

Brady's Reagent[38] (1.03 g 2,4-dinitrophenylhydrazine in 5.0 mL conc. H_2SO_4, 8.25 mL water, 23.75 mL ethanol) was used to test for the presence of the aldehyde in solution. In the positive control, about 2 mg of solid terephthalaldehyde was added to a test tube followed by 1 mL 95% ethanol and 1 mL Brady's Reagent to form a thick orange product. In the native control, 1 mL 95% ethanol was mixed with 1 mL Brady's Reagent, forming a transparent orange solution. In the test reaction, about 2 mg of solid product was mixed with 1 mL 95% ethanol and 1 mL Brady's Reagent to form a cloudy orange solution. The tubes were imaged after 5-min incubation at room temperature, see Figure 3.11. This result suggests that some aldehyde is present in the natural equilibrium. A 100 mM aqueous stock solution of the compound was prepared by mixing 18.43 mg and 244.9 µL molecular biology grade water for further tests.

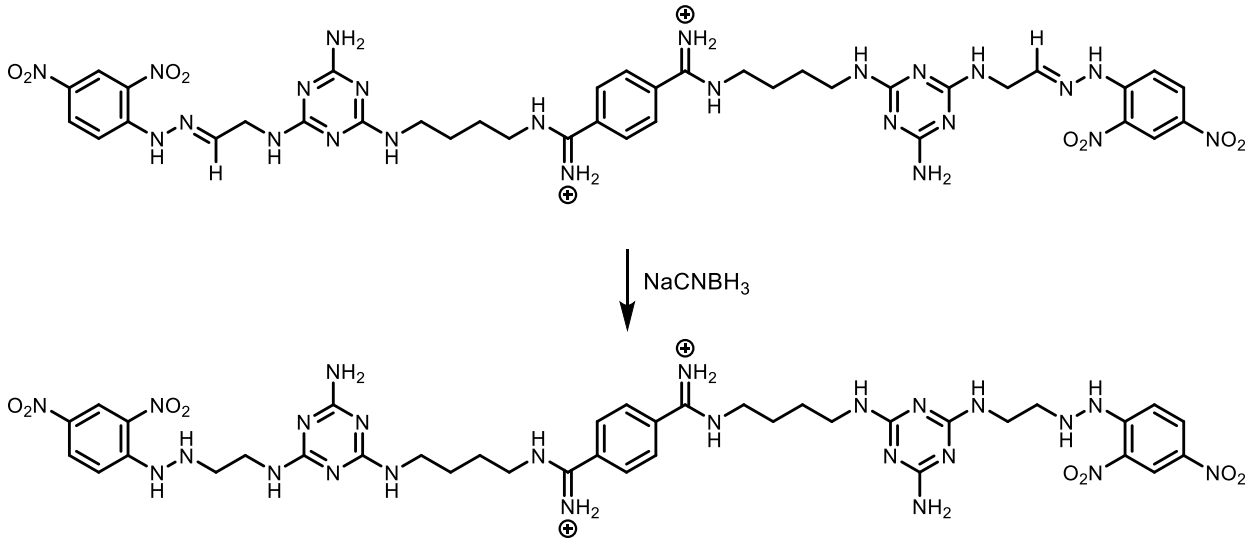
A second test was performed to evaluate the degree of conversion in this reaction. To the cyclized dialdehyde ligand **7** in DMSO was added 2 drops of 2,4-dinitrophenylhydrazine (DNP) in DMSO. The mixture was shaken, and the NMR spectra were recorded after about 5 h and again after about 5 d. The ^{1}H NMR spectra support the formation of the hydrazone product (Figure 3.12). A quantitative yield was obtained for this reaction, and HR ESI-MS supports formation of the product.



1,4-Phenylenebis(((4-((4-amino-6-(((E)-2-(2-(2,4-dinitrophenyl)hydrazineylidene)ethyl)amino)-1,3,5-triazin-2-yl)amino)butyl)amino)methaniminium). This reaction was performed as a reactivity test for the cyclized "pseudo-protected" aldehyde product. To a 1.5 mL Eppendorf tube was added 5 µL 100 mM (0.5 µmol) aqueous stock solution of ligand **9**, 12.5 µL 100 mM (1.25 µmol) DMSO stock of 2,4-dinitrophenylhydrazine, and 482.5 µL molecular biology grade water. The tube was vortexed, centrifuged, and incubated in 37 °C shaker for 24 h. The light orange solution was frozen and dried on a lyophilizer to yield 0.5 mg (100%) of the title compound as an orange to red solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 11.51 (m, 2H), 10.14 (m, 2H), 10.01 (s, 2H), 9.73 (m, 2H), 9.28 (m, 2H), 8.86 (m, 2H), 8.63 – 8.36 (m, 4H), 8.29 (dd, J = 10.05, 2.01, 2H), 8.10 – 7.84 (m, 8H), 7.68 (d, J = 9.74, 2H), 4.23 (bs, 4H), 1.63 (brm, 8H). Note: 4H missing, likely buried under water signal. HR-ESI MS (*m/z*) calcd for [M+H^{+}] 967.40; found 967.3991, calcd for

[M+2H$^+$]/2 484.20; found 484.2046. Elemental Composition 967.4009 $C_{38}H_{47}N_{24}O_8$ (desired M+H$^+$ $C_{38}H_{47}N_{24}O_8$).

Control $NaCNBH_3$ reduction of hydrazone



1,4-Phenylenebis(((4-((4-amino-6-((2-(2-(2,4-dinitrophenyl)hydrazineyl)ethyl)amino)-1,3,5-triazin-2-yl)amino)butyl)amino)methaniminium). The sample of 0.5 mg 1,4-phenylenebis(((4-((4-amino-6-(((E)-2-(2-(2,4-dinitrophenyl)hydrazineylidene)ethyl)amino)-1,3,5-triazin-2-yl)amino)butyl)amino)methaniminium) was dissolved in 400 µL DMSO and 15 µL 100 mM (0.015 mmol) aqueous $NaCNBH_3$ was added. The Eppendorf tube was incubated on a shaker at 37 °C for 15 h. The yield of the reduction was not quantified, because the goal was to see if the reduced product could be seen on MALDI-MS. HR-ESI MS (*m/z*) calcd for [M+H$^+$] 967.40; found 967.4028, calcd for [M+2H$^+$]/2 484.20; found 484.2004. It remains unclear why the imine is still observed in the ESI-MS, because the theoretical yield of imine product is 0.50 µmol and the amount of reducing agent added was 15 µmol, which should be plenty to reduce all imine bonds. The expected mass of the reduced product [M+H$^+$] is 971.43 and [M+2H$^+$]/2 is 486.22.

Analytical agarose gel to confirm successful DNase cleavage in presence of $NaCNBH_3$. An agarose gel was prepared by mixing 0.5 g agarose, was 5 mL 10X TBE buffer, and 45 mL DI water in a 50 mL beaker, microwaving 15 s at a time until all agarose dissolved, adding 1 µL ethidium bromide, and swirling to mix. The gel cast by slowly pouring the clear solution into the mold, removing bubbles, sliding the comb into the liquid, and allowing to set for about 30 min. The buffer solution for running the gel was made with 25 mL 10X TBE Buffer, 225 mL DI water, and 5 µL ethidium bromide. DNA samples were prepared by adding 2 µL of 24.3 µM 5'-GGG $(CTG)_{16}$ CCC-3' solution (54 bases) to each of 4 Eppendorf tubes. Two tubes were treated with $NaCNBH_3$ (2 µL 100 mM solution) and/or DNase I (NEB, 1 µL 2,000 U/mL). Of note, the final concentration of NaCNBH3 used in this experiment (33.3 µM) is higher than used in the screening experiments described here. In the untreated tubes, these reagents were substituted with the same volume of molecular biology grade water. After addition of DNase, the tubes treated with DNase were incubated for 10 min on ice (0 °C), 10 min at 37 °C, and 10 min at 95 °C to inactivate the DNase. Prior to loading the samples, 1 µL 6X purple DNA loading dye was added to each sample. The gel was loaded with 6 µL of 100 bp DNA ladder (NEB) and 6 µL of each sample. The gel was run at 150 V for 42 min and imaged on a Bio-Rad Gel Doc System. This gel experiment demonstrates that DNase is active in the presence of $NaCNBH_3$ and thus the two reagents are compatible. This test was repeated for other targets and RNase H, I_f, and A, and similar results were obtained – all cleavage enzymes remained active in the presence of $NaCNBH_3$. Data included in Figure 3.13.

3.4.3. Synthetic Methods

See Appendix A, Section A.3.1 for synthesis of **2** and **3**.

3.4.4. Synthesis of r(CUG)$_{90}$ Template via in vitro Transcription

For general procedures, see Appendix A, Section A.1.11.

3.4.5. MALDI-MS and HPLC Assays

Nucleic Acid Sequences

d(CTG)$_{16}$ = 5' – GGG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CCC – 3' (purchased from IDT, MW = 16,554.6 g/mol)

r(CUG)$_{16}$ = 5' – GGG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CCC – 3' (purchased from IDT, MW = 17,194.1 g/mol)

r(CUG)$_{90}$ = see section 3.4.4. (above) for sequence

Procedure 3.4.1. MALDI-MS Assay: 2+3 assembly with varied templates. Compounds (100 µM each) and DNA or RNA templates (10 µM for d(CTG)$_{16}$ and r(CUG)$_{16}$ and 500 nM for r(CUG)$_{90}$) were incubated in aqueous buffer consisting of 2 mM each of KCl, $MgCl_2$, $CaCl_2$, and Tri-HCl, pH 7 in a final volume of 50 µL. The solutions were annealed in buffer before addition of compounds on a thermocycler at 95 °C for 5 min and allowed to cool by 0.1 °C /s to 25 °C. The samples were incubated at 37 °C for 3 h before reduction with 2.5 µL of 10 mM $NaCNBH_3$. Following overnight incubation at 37 °C (for reduction step), 1 µL of either DNase I, RNase H, or molecular biology grade water was added to each tube, dependent upon the presence of DNA or RNA template (or lack thereof). The solutions were incubated at 37 °C for 15 min, a 1 µL aliquot was removed for MALDI-MS (run in 1 µL DHB matrix). See Figure 3.5, 3.16 for data.

Procedure 3.4.2. High MW MALDI-MS Assay: 2+3 assembly with varied templates Compounds (100 µM each) and DNA or RNA templates (10 µM for d(CTG)$_{16}$ and r(CUG)$_{16}$ and 500 nM for r(CUG)$_{90}$) were incubated in aqueous buffer consisting of 2 mM each of KCl, $MgCl_2$,

$CaCl_2$, and Tri-HCl, pH 7 in a final volume of 50 µL. The solutions were annealed in buffer before addition of compounds on a thermocycler at 95 °C for 5 min and allowed to cool by 0.1 °C /s to 25 °C. The samples were incubated at 37 °C for 24 h before reduction with 2.5 µL of 100 mM $NaCNBH_3$. Following overnight incubation at 37 °C (for reduction step), a 1 µL aliquot was removed for MALDI-MS (run in 1 µL DHB matrix). See Figure 3.6 for data.

Procedure 3.4.3. MALDI-MS Assay: Monomers 2 and 3 with DNA Templates. Compounds (100 µM) and DNA templates (100 µM for $d(CTG)_3$ and random duplex DNA) were incubated in aqueous buffer consisting of 2 mM each of KCl, $MgCl_2$, $CaCl_2$, and Tri-HCl, pH 7 in a final volume of 10 µL. The solutions were annealed in buffer before addition of compounds on a thermocycler at 95 °C for 5 min and allowed to cool by 0.1 °C/s to 25 °C. The samples were incubated at 37 °C for 3 h and reduced with 1 µL of 10 mM $NaCNBH_3$. After 3 h incubation at 37 °C, a 1 µL aliquot was removed for MALDI-MS (run in 1 µL DHB matrix for low MW and 3-HPA for high MW). See Figure 3.17 for data. Experiment performed by Amie Lanzendorf.

Procedure 3.4.4. HPLC Assay: 2+3 assembly with varied templates. Compounds (100 µM each) and DNA or RNA (10 µM for $d(CTG)_{16}$, $r(CUG)_{16}$, and 500 nM for $r(CUG)_{90}$) were incubated in aqueous buffer consisting of 2 mM each of KCl, $MgCl_2$, $CaCl_2$, and Tri-HCl, pH 7 in a final volume of 50 µL. The solutions were annealed in buffer before addition of compounds on a thermocycler at 95 °C for 5 min and allowed to cool by 0.1 °C/s to 25 °C. The samples were incubated at 37 °C for 24 h before reduction with 2.5 µL of 100 mM $NaCNBH_3$. Following overnight incubation at 37 °C (for reduction step), 1 µL of either DNase I, RNase I_f, or molecular biology grade water was added to each tube, dependent upon the presence of DNA or RNA template (or lack thereof). The solutions were incubated at 37 °C for 30 min, a 1 µL aliquot was removed for MALDI-MS (run in 1 µL DHB matrix) and the rest of the samples were run on

analytical HPLC with gradient acetonitrile in water, 0.1% TFA 0-50% over 10 min, 50% for 2.5 min, 50-100% over 2.5 min. See Figure 3.18 for data.

Procedure 3.4.5. HPLC and MALDI-MS Assay: 2+3 assembly without reduction. Compounds (100 μM each) and DNA or RNA templates (10 μM for $d(CTG)_{16}$ and $r(CUG)_{16}$ and 500 nM for $r(CUG)_{90}$) were incubated in aqueous buffer consisting of 2 mM each of KCl, $MgCl_2$, $CaCl_2$, and Tri-HCl, pH 7 in a final volume of 50 μL. The solutions were annealed in buffer before addition of compounds on a thermocycler at 95 °C for 5 min and allowed to cool by 0.1 °C/s to 25 °C. The samples were incubated at 37 °C for 24 h and a 1 μL aliquot was removed for MALDI-MS (run in 1 μL DHB matrix) before the rest of the samples were run on analytical HPLC with gradient acetonitrile in water, 0.1% TFA 0-50% over 10 min, 50% for 2.5 min, 50-100% over 2.5 min. See Figure 3.20 for data.

3.4.6. Isothermal Titration Calorimetry: TA Affinity ITC

ITC Binding Experiments were performed on TA Affinity ITC low volume (TA Instruments, New Castle, DE) at the University of Illinois at Urbana-Champaign Microanalysis Laboratory. The syringe holds up to 250 μL and the reference and sample cells hold 190 μL each. All DNA/RNA solutions were freshly annealed by heating at 95 °C for 5 min and slowly cooling to 25 °C. The reference cell was filled with DI water for all experiments and spectra are reported with a a blank (ligand into buffer) subtracted to account for heat-of-dilution. The ITC was cleaned with 1% Contrad solution, Millipore water, and methanol and vacuum dried before and after each experiment. Each ITC experiment consisted of 20 injections of 2.00 ± 0.01 μL. The instrument was set to auto-equilibrate before each experiment with maximum equilibration delay of 300 s followed by a 200 s initial baseline. Injections were 200 s apart. All ITC experiments were run at

25 °C and stirring speed of 125 revolutions per minute. The buffer for the DNA and ligand solutions was made from stock solutions of NaCl (5 M, aqueous) and MOPS (1 M, pH 7 ± 0.2) to a final concentration of 300 mM NaCl and 20 mM MOPS (pH 7) in water. Ligand stock solutions were 10 mM in molecular biology grade water and nucleic acid stock solutions were 1 mM or 2 mM in molecular biology grade water. The nucleic acid (cell) concentration was 20 μM and the ligand concentration (syringe) was 500 μM. In the case of **2**+**3**, syringe concentration was 500 μM in both **2** and **3**. Data was analyzed using NanoAnalyze software (TA Instruments, New Castle, DE) to integrate each isotherm (μJ per s v. s). The automatically predicted baseline was visually inspected for errors and adjusted as necessary and the auto-adjust feature was utilized to set integration regions. The blank titration (ligand into buffer) in Figure 3.21 was subtracted from each spectrum as the average across all injections. The independent algorithm (for one type of binding site) was used to calculate the dissociation constant (K_d) where applicable. For curve fitting, the final heat generated was normalized to ~0 kJ/mol to allow the model algorithm to properly fit the curve. Each experiment was run in triplicate and the K_d values reported as average of the three experiments. Error is reported as standard error of the mean.

Oligonucleotide sequences

$d(CTG)_{16}$ = 5’-GGG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CCC-3’

$r(CUG)_{16}$ = 5’-GGG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CCC-3’

Random DNA Duplex = 5'-GCC ATC ACG GAT CAC GTC-3' // 3’-CGG TAG TGC CTA GTG CAG-5’

Random ssRNA = 5’-ACC AGA ACG AUG UCC AAG-3’

3.4.7. in vitro Transcription Inhibition Assay with T7 RNA Polymerase[39]

For general procedures, see Appendix A, Section A.1.8.

T7-Random Duplex Sequence:

5’-TAATACGACTCACTATAGGGCGACGTTAAGTACATTACGCTGTCATAGGCGGCGTTCAGG-3’

3’-ATTATGCTGAGTGATATCCCGCTGCAATTCATGTAATGCGACAGTATCCGCCGCAAGTCC-5’

3.4.8. Insulin Receptor (IR) Mis-splicing Rescue Assay[40]

For procedure, see Appendix A, Section A.1.9.

3.4.9. Sulforhodamine B Cytotoxicity Assay[41]

For procedure, see Appendix A, Section A.1.3.

3.5. SUPPLEMENTAL FIGURES

3.5.1. Preliminary Design of Monomers

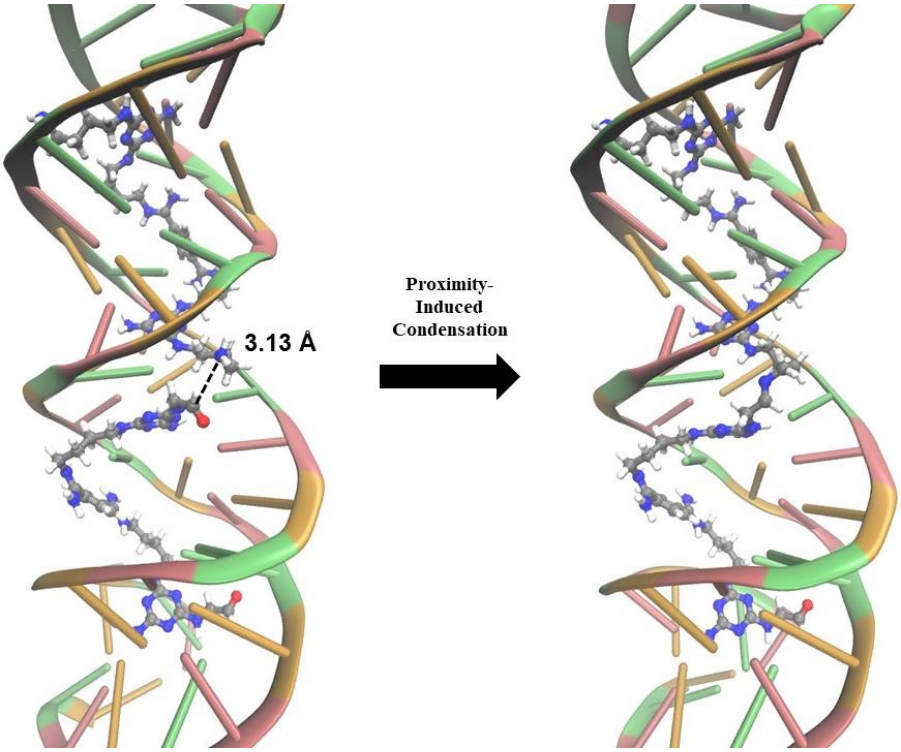


Figure 3.10. MOE docking studies for ligands 4 and 5. Uracil is green, cytosine is pink, and guanine is orange. Hydrogen bonds are depicted as red dotted lines. The triazine moieties serve as U-U mismatch recognition units and the bisamidinium groove-binding linker is docked in the major groove of the RNA helix. In this docked structure, each ligand spans a total of 3 U-U mismatches. The distance between reactive partners was measured to be 3.13 Å. The fragments were manually ligated, and the structure minimized to show the dimeric product on the $(CUG)_6$ helix.

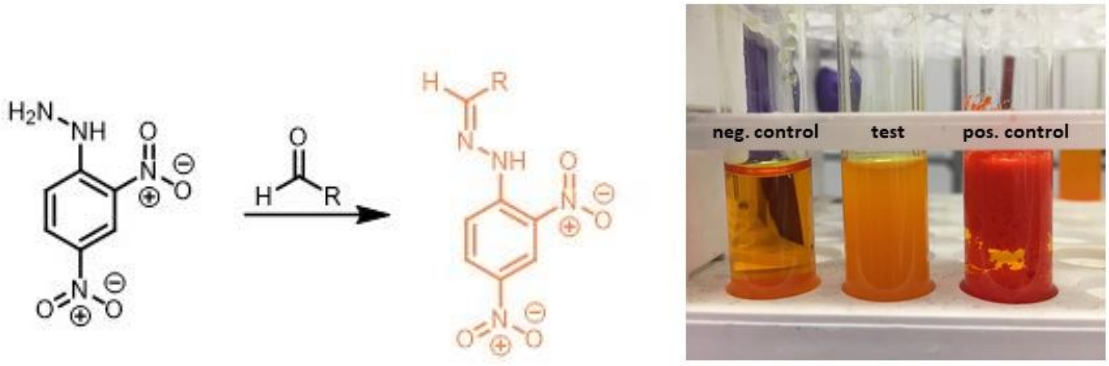


Figure 3.11. Brady's test for presence of aldehyde product. Negative control on the left, test solution (with masked dialdehyde) in center, positive control (with terephthalaldehyde) on the right. The cloudy solution indicates that some reaction is occurring. Of note, the center solution is much less concentrated than the positive control.

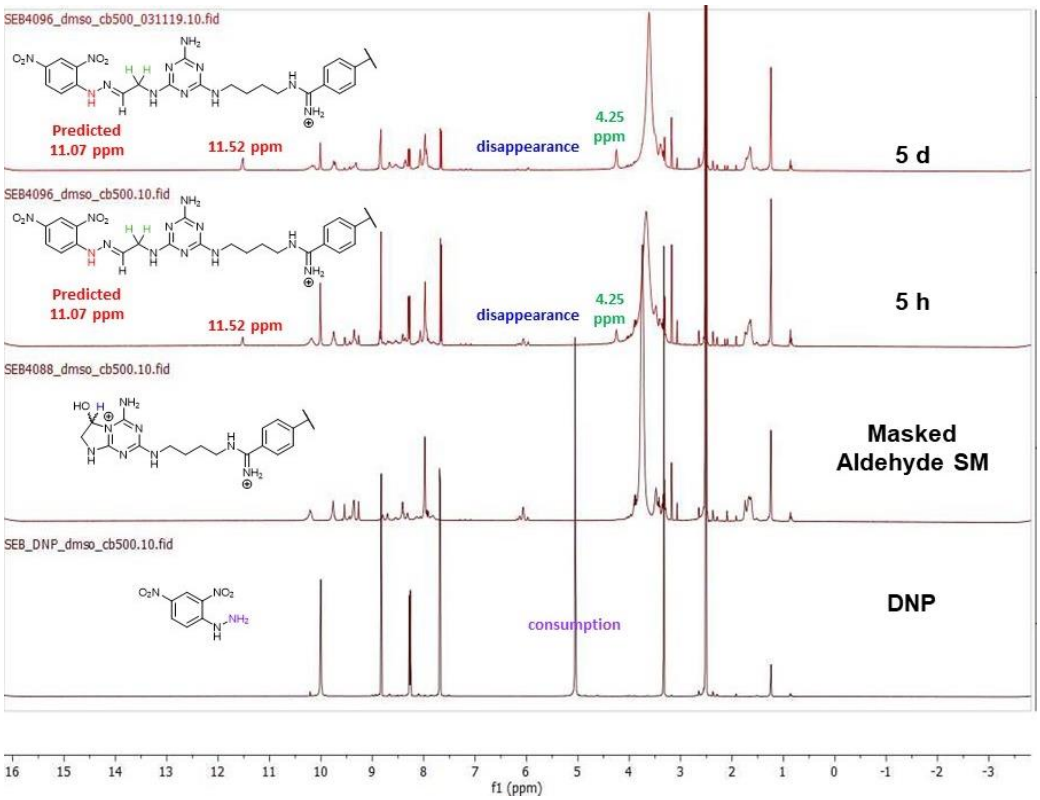


Figure 3.12. ^{1}H NMR time-study of ligand 10 + DNP-hydrazine in DMSO. Bottom to top: DNP-hydrazine starting material, hemiaminal **9** starting material, reaction mixture after 5 h, reaction mixture after 5 d. The formation of the desired product is supported by the appearance of a new peak at 4.25 ppm (green), presumably resulting from the disappearance of the diastereotopic methylene protons and the reappearance of these protons as a single peak after condensation. The NMR also demonstrates the disappearance of the methine proton in the 5-membered imidazole-like ring (blue). Additionally, the hydrazine NH_2 protons (~5 ppm, purple) were completely consumed and a new peak at 11.52 ppm (red) was observed, indicative of hydrazone formation.

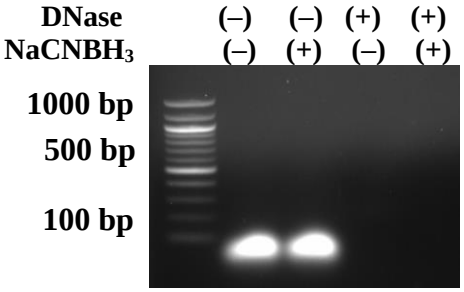


Figure 3.13. Analytical agarose gel of to monitor activity of DNase in presence of $NaCNBH_3$. Lane 1 contains NEB 100 bp DNA ladder. Lane 2-5 contain $d(CTG)_{16}$ treated with $NaCNBH_3$, DNase, neither, or both.

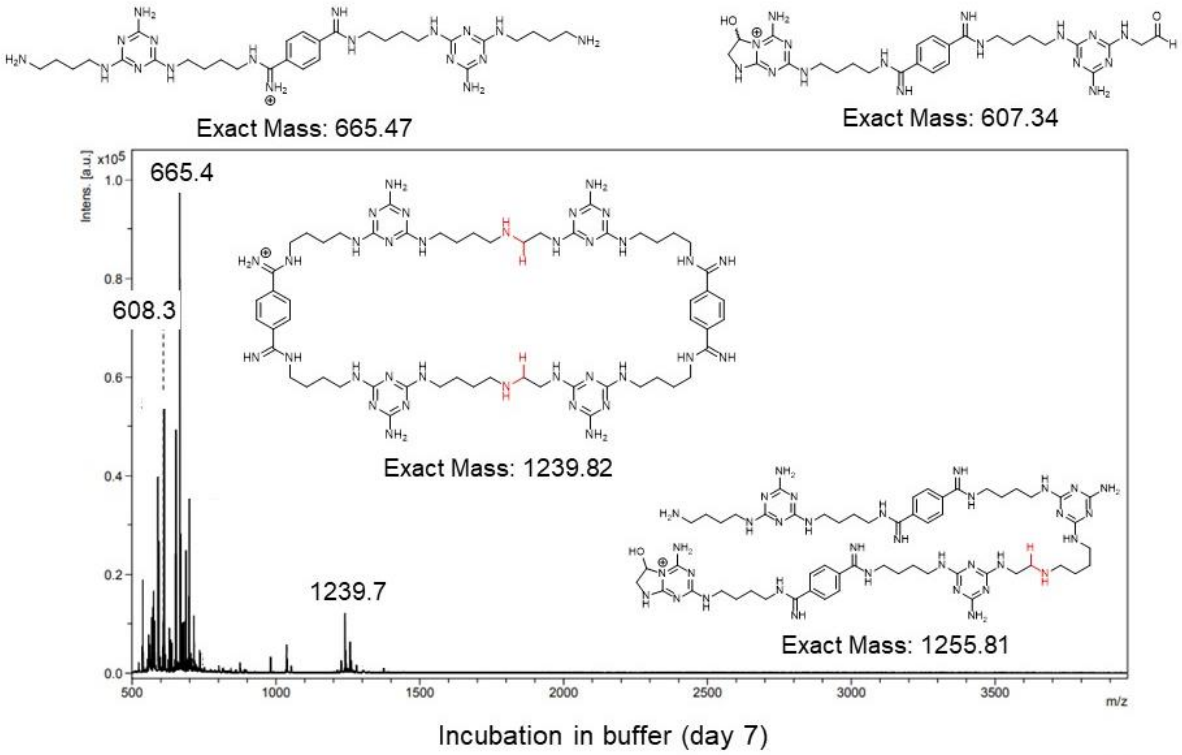


Figure 3.14. MALDI-MS Results from initial condensation screen with 5 and 6. Macrocyclic products were observed.

3.5.2. Computational Modeling

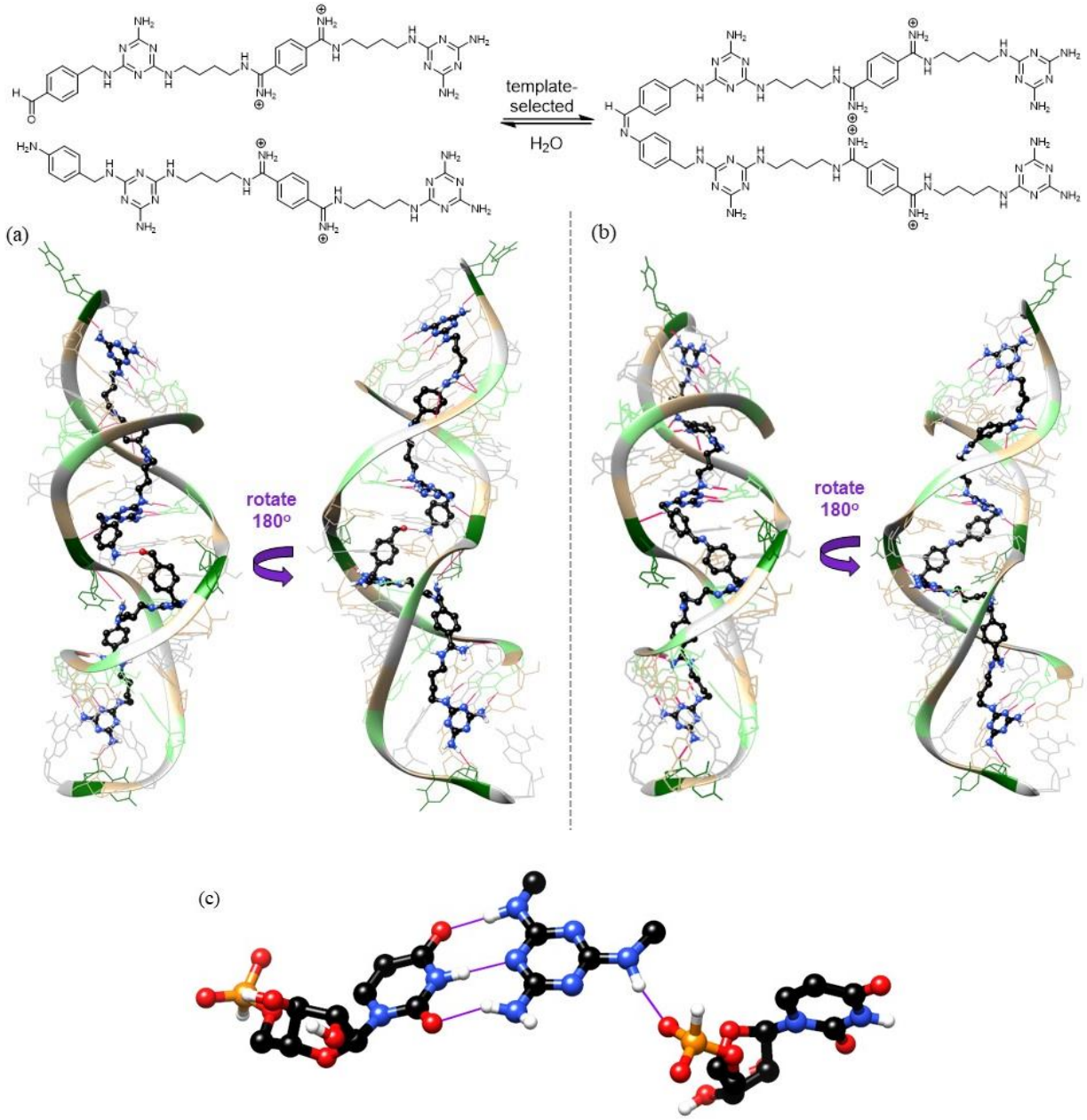


Figure 3.15. Molecular Docking Study for 2+3 of r(CUG)$_6$. (a) Monomers **2** and **3** docked individually with sequential minimizations. (b) Manually linked imine dimer **2**+**3** after minimization. (c) Representative view of melamine – uracil hydrogen bonding interactions with one uracil flipped out of the helix. The melamine forms 3 hydrogen bonds with the intact uracil and one hydrogen bond with the phosphate backbone where the mismatched uracil is flipped out of the helix. Ligands are shown in ball and stick model with atoms colored by heteroatom and carbon black. Uracil is green (light green are flipped in, dark green flipped out), thymine is tan and guanine is grey. Hydrogen bonds between ligand and RNA are shown in red. r(CUG)$_6$ PDB ID: 3gm7.[29]

3.5.3. MALDI-MS and HPLC Assays

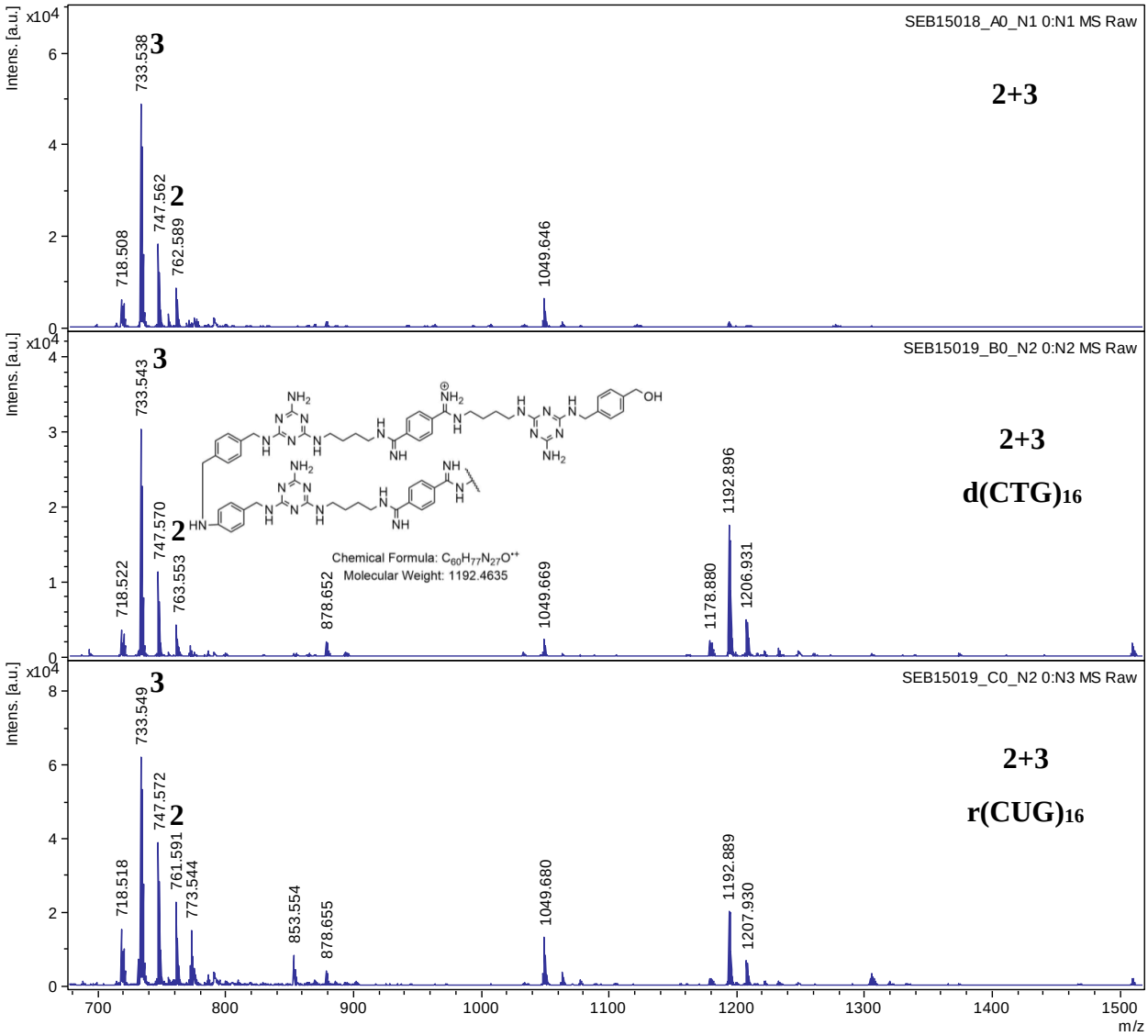


Figure 3.16. MALDI-MS analysis of **2**+**3** assembly on and off template with reduction by $NaCNBH_3$ after 5 min incubation. Buffer: 2 mM each of KCl, $MgCl_2$, $CaCl_2$, and Tri-HCl, pH 7, incubated at 37 °C for 5-75 min before reduction with sodium cyanoborohydride and analysis. Compound concentration 100 µM, 10 µM $d(CTG)_{16}$ or $r(CUG)_{16}$. Notably, the product and monomer **2** exhibit several peaks across a range of *m/z* values, consistent with these compounds existing in multiple states of reduction. See Procedure 3.4.1 for details.

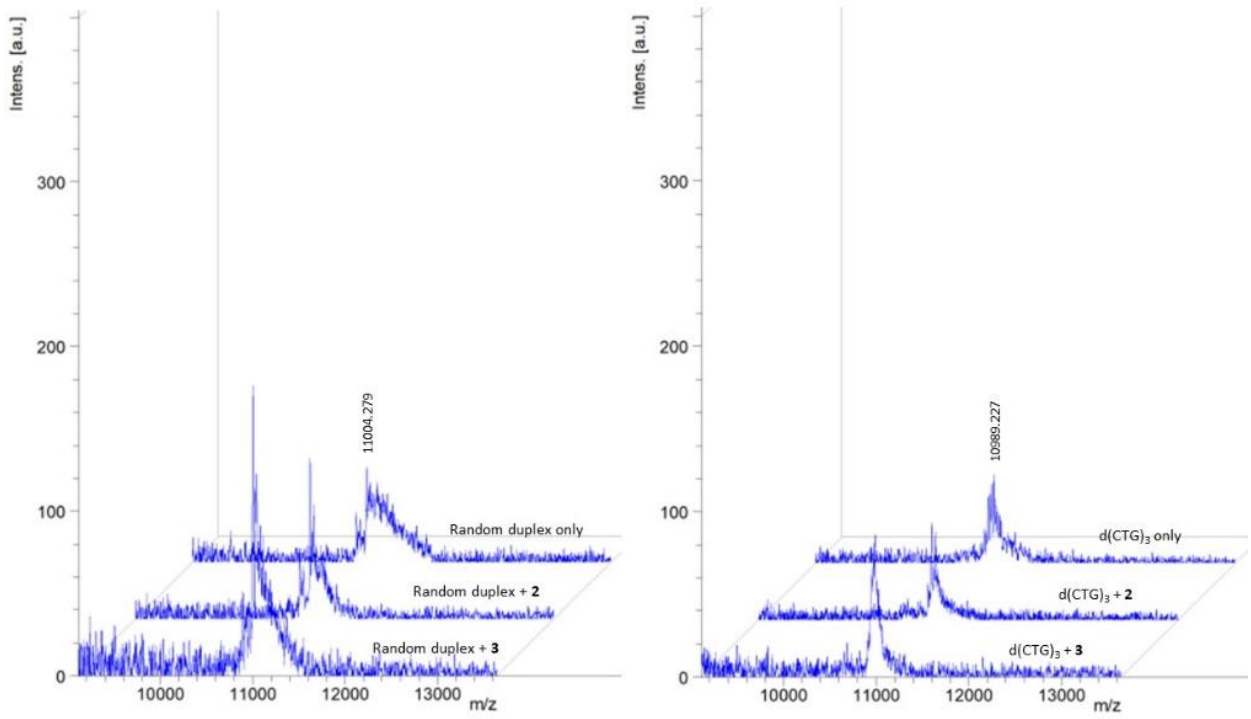


Figure 3.17. MALDI-MS analysis of monomers **2** and **3** with $d(CTG)_3$ and random duplex DNA. Buffer: 2 mM each of KCl, $MgCl_2$, $CaCl_2$, and Tri-HCl, pH 7, incubated at 37 °C for 3 h before reduction with sodium cyanoborohydride for 3 h and analysis. 100 μM compounds, 100 μM $d(CTG)_3$ or random duplex DNA. See Procedure 3.4.3 for details. Experiment performed by Amie Lanzendorf.

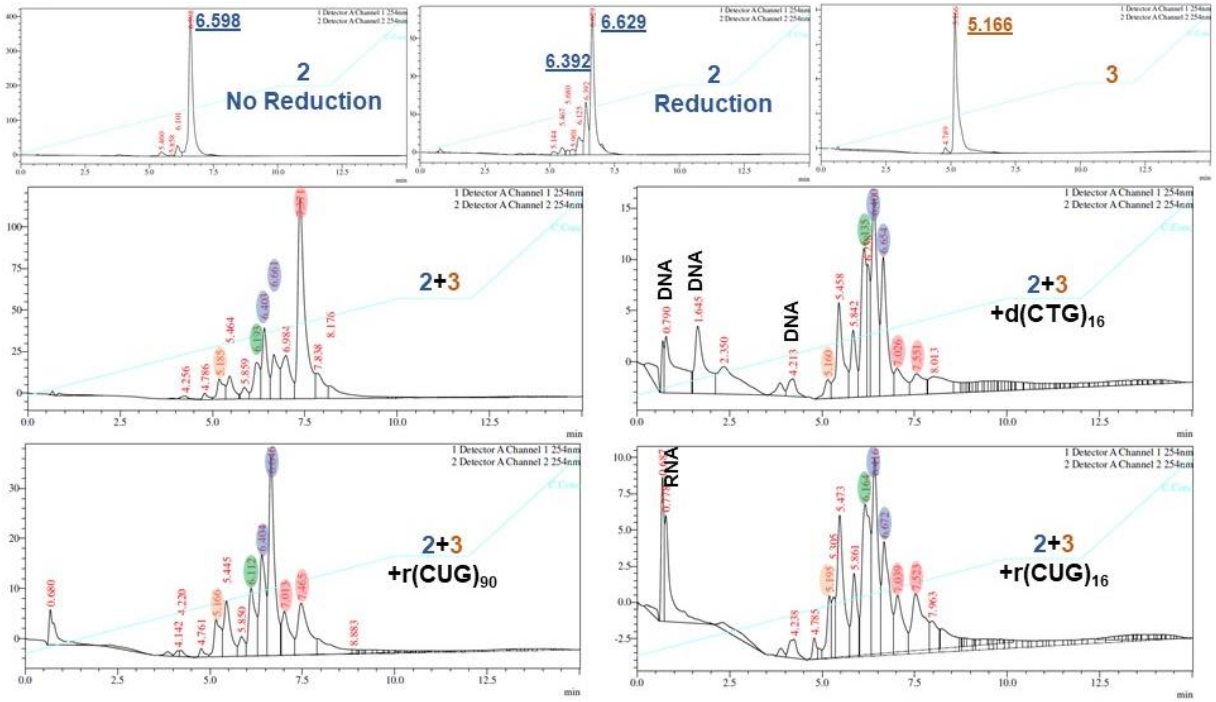


Figure 3.18. HPLC analysis of **2**, **3**, and **2+3** on with no template or d(CTG)$_{16}$, r(CUG)$_{16}$, and r(CUG)$_{90}$ templates. Blue highlights consistent with compound **2**, orange **3**, green dimer **2+3**, and red trimer and pentamer. Buffer: 2 mM each of KCl, $MgCl_2$, $CaCl_2$, and Tri-HCl, pH 7, incubated at 37 °C for 24 h (HPLC) before reduction with sodium cyanoborohydride and degradation of templates. Compound concentration 100 µM, RNA concentration 10 µM for d(CTG)$_{16}$, r(CUG)$_{16}$, and 500 nM for r(CUG)$_{90}$. See Procedure S4 for details.

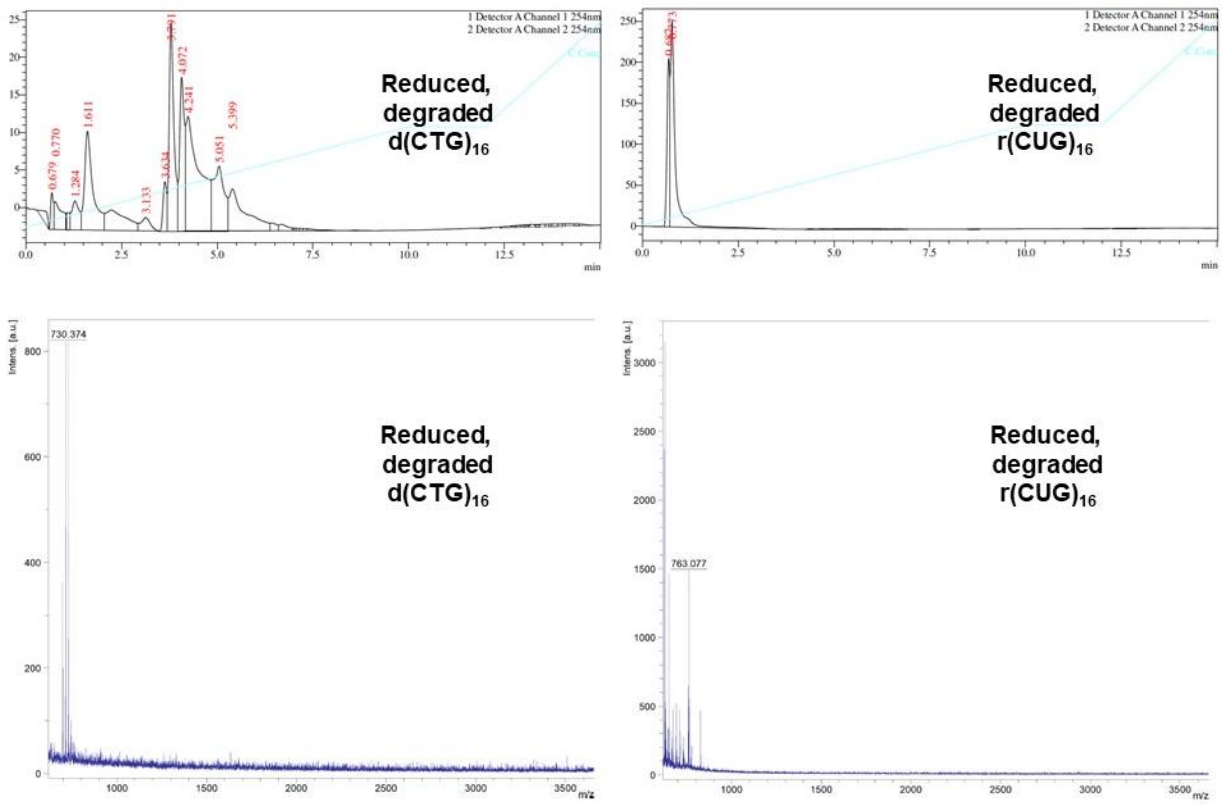


Figure 3.19. HPLC and MALDI controls: analysis of degraded $d(CTG)_{16}$ and $r(CUG)_{16}$ templates. DNA or RNA (10 μM) were incubated in aqueous buffer consisting of 2 mM each of KCl, $MgCl_2$, $CaCl_2$, and Tri-HCl, pH 7 in a final volume of 50 μL. The solutions were annealed in buffer before addition of compounds on a thermocycler at 95 °C for 5 min and allowed to cool by 0.1 °C/s to 25 °C. The samples were reduced with 2.5 μL of 100 mM sodium cyanoborohydride. Following overnight incubation at 37 °C (for reduction step), 1 μL of either DNase I or RNase I_f was added to each tube, dependent upon the presence of DNA or RNA template. The solutions were incubated at 37 °C for 30 min, a 1 μL aliquot was removed for MALDI-MS (run in 1 μL DHB matrix) and the rest of the samples were run on analytical HPLC with gradient acetonitrile in water, 0.1% TFA 0-50% over 10 min, 50% for 2.5 min, 50-100% over 2.5 min.

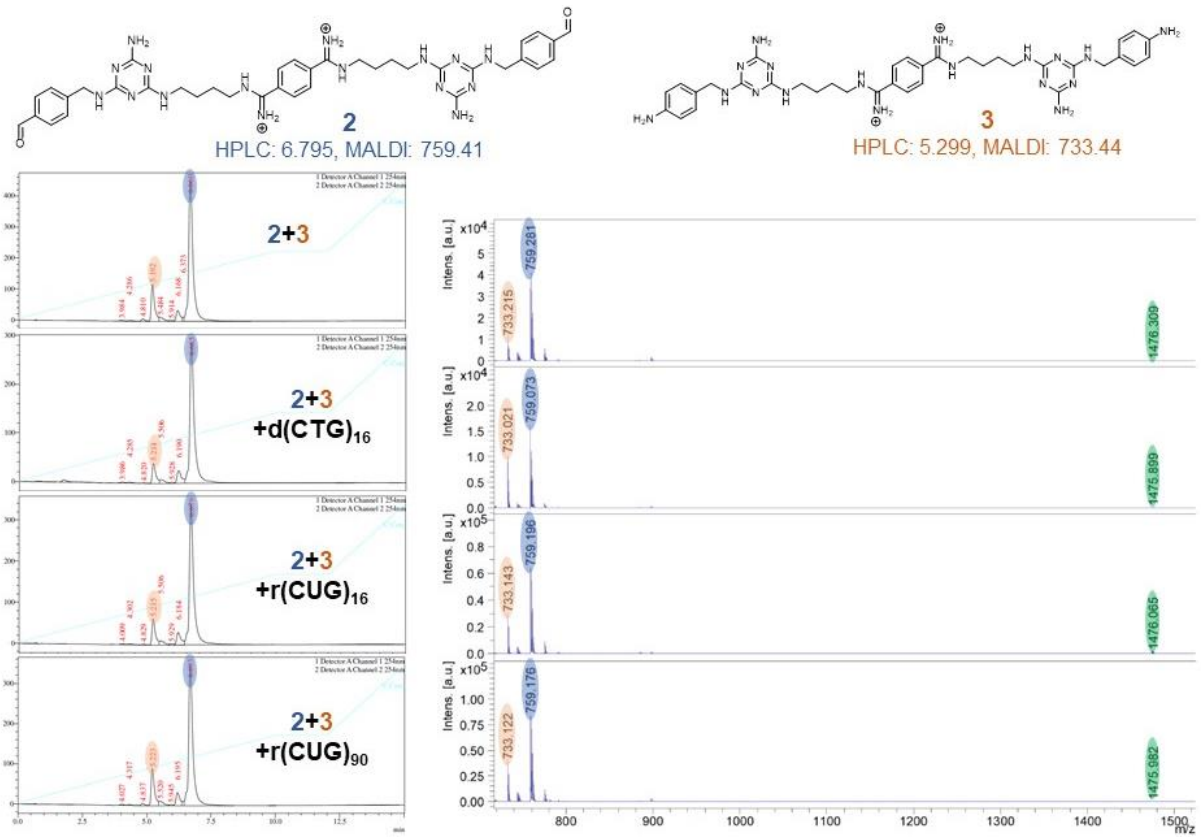


Figure 3.20. HPLC and MALDI-MS analysis of **2**+**3** assembly on $d(CTG)_{16}$ and $r(CUG)_{16}$ without reduction of the imine. Buffer: 2 mM each of KCl, $MgCl_2$, $CaCl_2$, and Tri-HCl, pH 7, incubated at 37 °C for 24 h before analysis. No reduction was performed. Compound concentration 100 µM, 10 µM $d(CTG)_{16}$ and $r(CUG)_{16}$ or 500 nM $r(CUG)_{90}$. See Procedure 3.4.5 for details.

3.5.4. Isothermal Titration Calorimetry

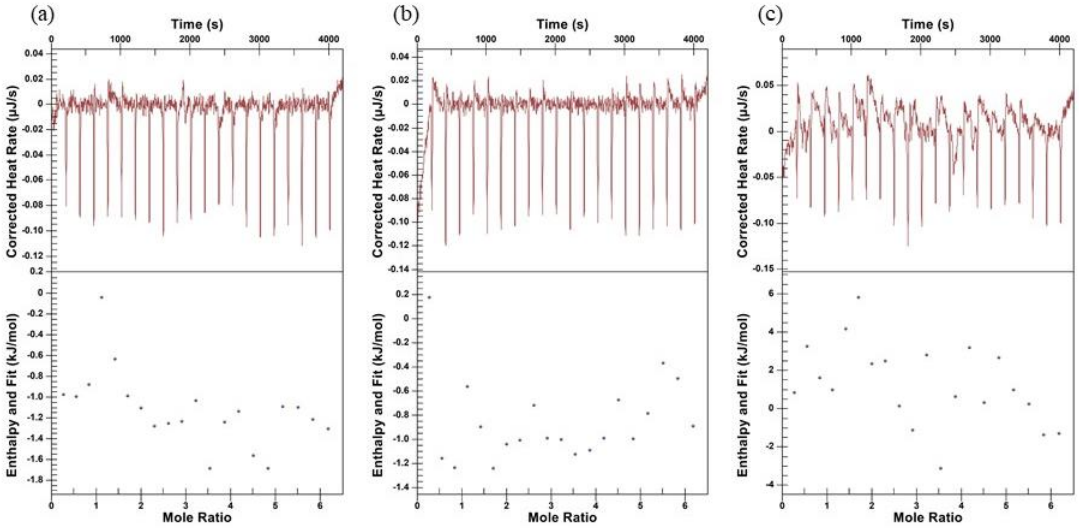


Figure 3.21. ITC blanks, with ligand titrated into buffer for (a) **2**, (b) **3**, and (c) **2**+**3**. [Cell (nucleic acid)] = 0 µM, [Syringe (ligand)] = 500 µM, Buffer = 300 mM NaCl, 20 mM MOPS. These blank controls were subtracted from each titration into DNA or RNA to account for heat of dilution.

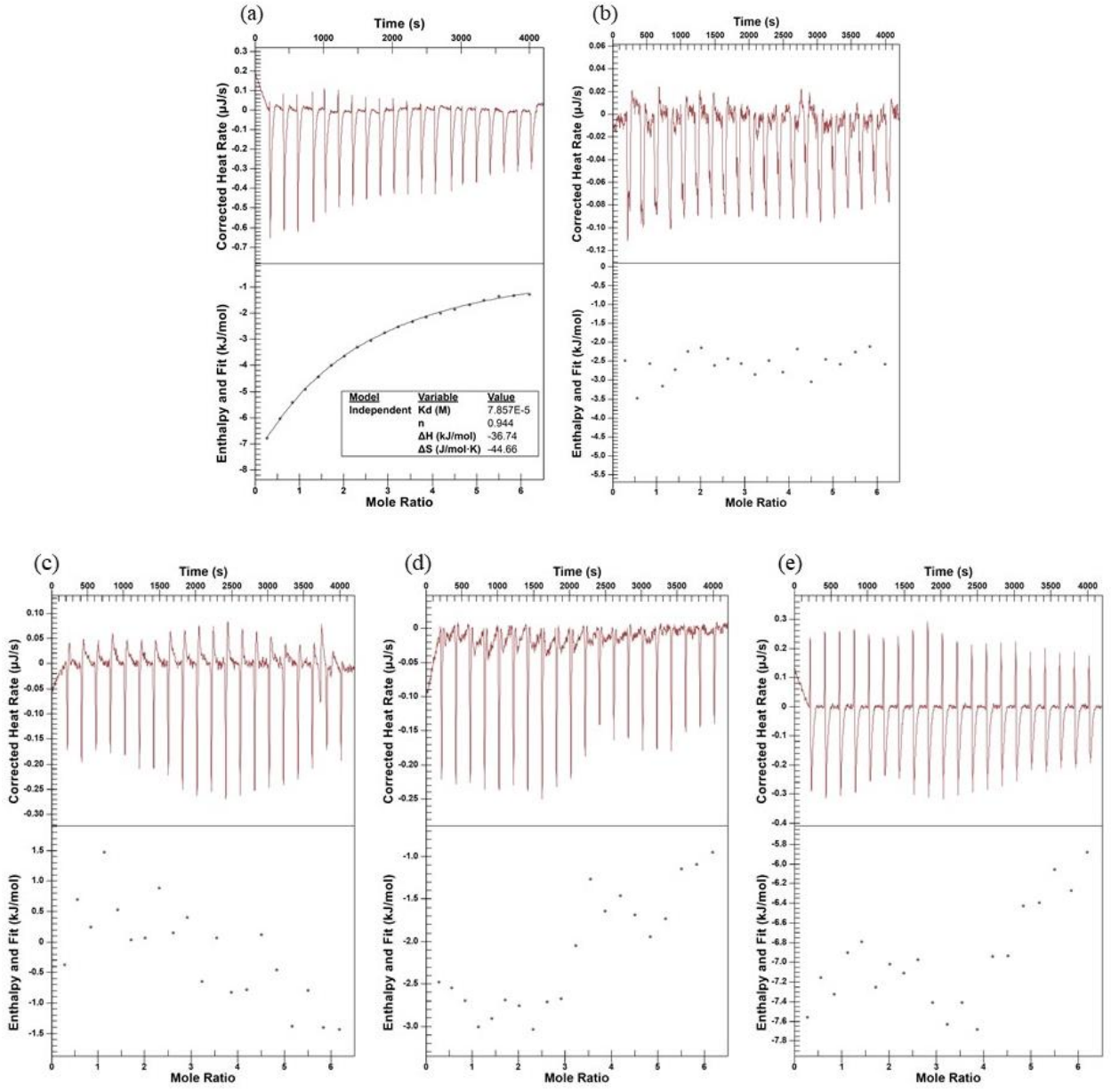


Figure 3.22. ITC studies with (a) **2** and (b) **3** titrated into d(CTG)16 and (c) **2**, (d) **3**, and (e) **2**+**3** titrated into a random DNA duplex. [Cell (nucleic acid)] = 20 µM, [Syringe (ligand)] = 500 µM, Buffer = 300 mM NaCl, 20 mM MOPS.

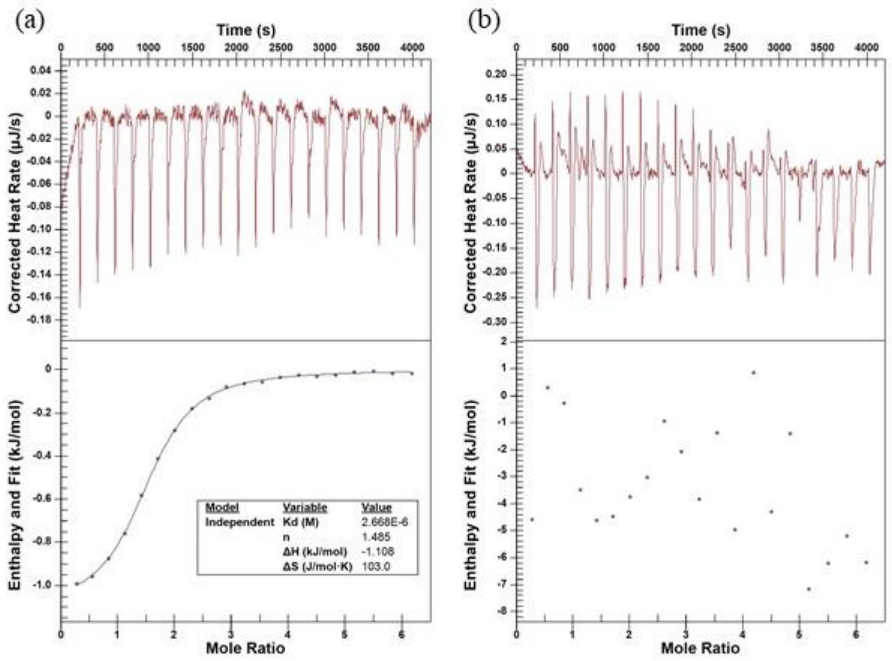


Figure 3.23. ITC study with **2**+**3** titrated into (a) r(CUG)$_{16}$ and (b) a random ssRNA. [Cell (nucleic acid)] = 20 µM, [Syringe (ligand)] = 500 µM, Buffer = 300 mM NaCl, 20 mM MOPS.

Ligand	Sequence	K_d (µM)	n	ΔH (kJ/mol)	ΔS (J/mol·K)
2	d(CTG)$_{16}$	81.57 ± 0.88	1.63 ± 0.24	-20.03 ± 7.02	11.11 ± 23.51
	DNA R.D	n.b.			
3	d(CTG)$_{16}$	n.b.			
	DNA R.D	n.b.			
2+3	d(CTG)$_{16}$	5.58 ± 0.54	2.37 ± 0.16	-4.01 ± 0.70	88.1 ± 2.87
	DNA R.D	n.b.			
	r(CUG)$_{16}$	4.30 ± 0.49	1.53 ± 0.03	-1.62 ± 0.24	97.9 ± 1.84
	RNA R.D	n.b.			

Table 3.1. Thermodynamic Values from ITC experiments using independent binding model. All values are reported as an average over 3 independent experiments with error reported as standard error. n.b. indicated no binding observed under these conditions.

3.5.5. in vitro Transcription Inhibition Assays

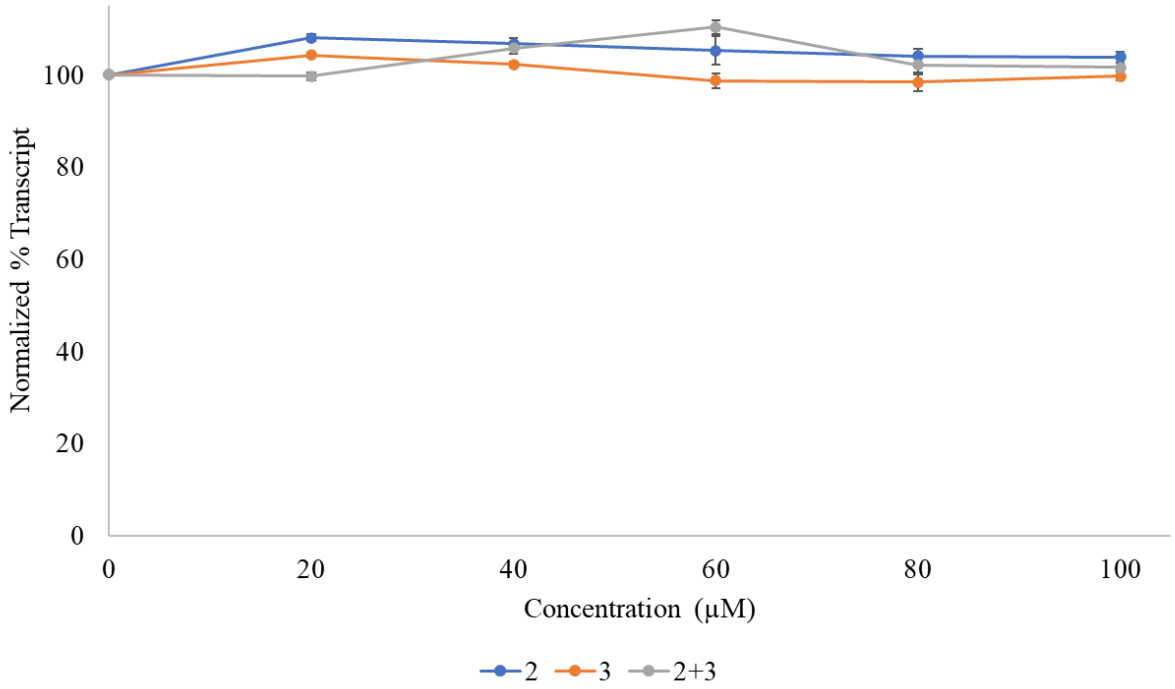


Figure 3.24. T7 *in vitro* transcription inhibition for compounds **2**, **3**, and **2**+**3** with Random DNA Duplex. Results are the average of at least 3 independent replicates.

3.5.6. SRB Cytotoxicity Assay

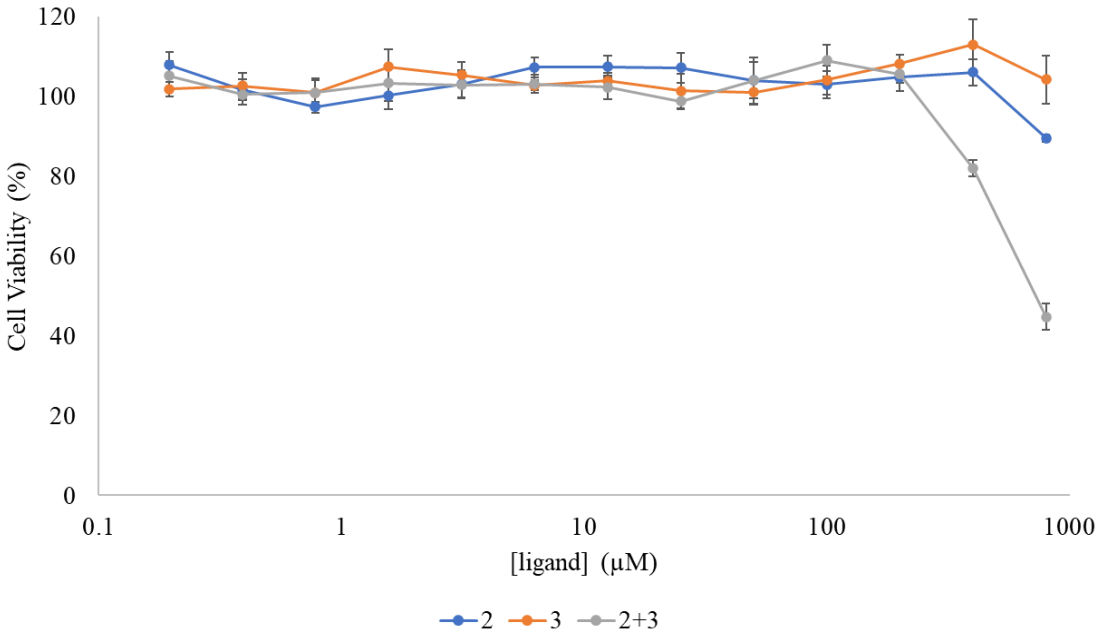


Figure 3.25. SRB Cytotoxicity Assay for compounds **2** and **3**. Results reported as the average of at least 5 independent replicates.

3.5.7. Insulin Receptor (IR) Mis-splicing Rescue Assay

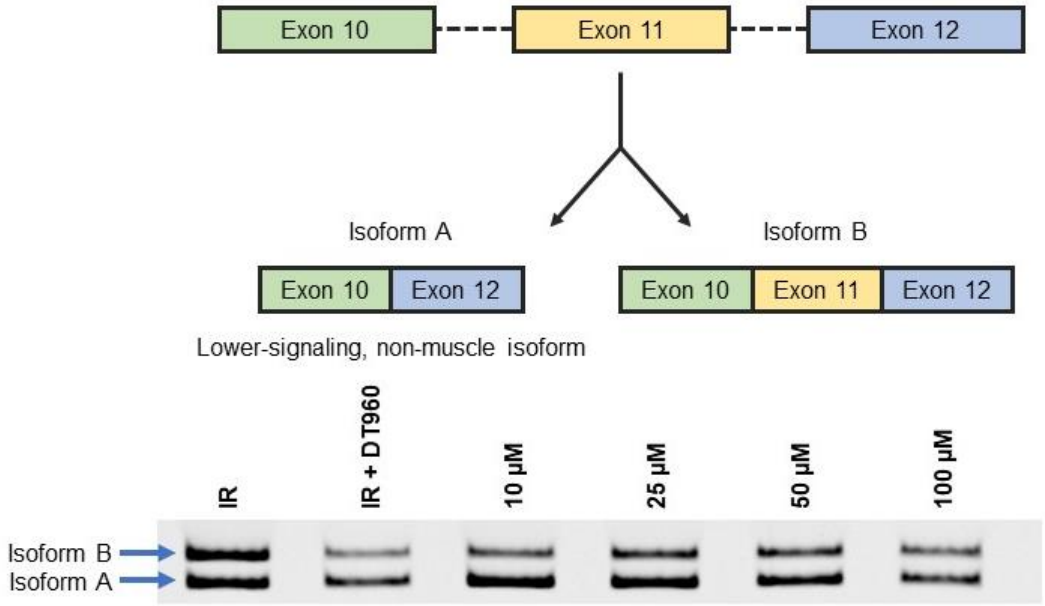


Figure 3.26. Gel image for IR mis-splicing assay reported in Figure 3.8. Splicing was rescued in a dose-dependent manner by **2**+**3**. 10 µM produced a 36% splicing rescue, 25 µM 40% rescue, 50 µM 63% recue, and 100 µM 65% rescue. Results are average of at least 3 independent replicates.

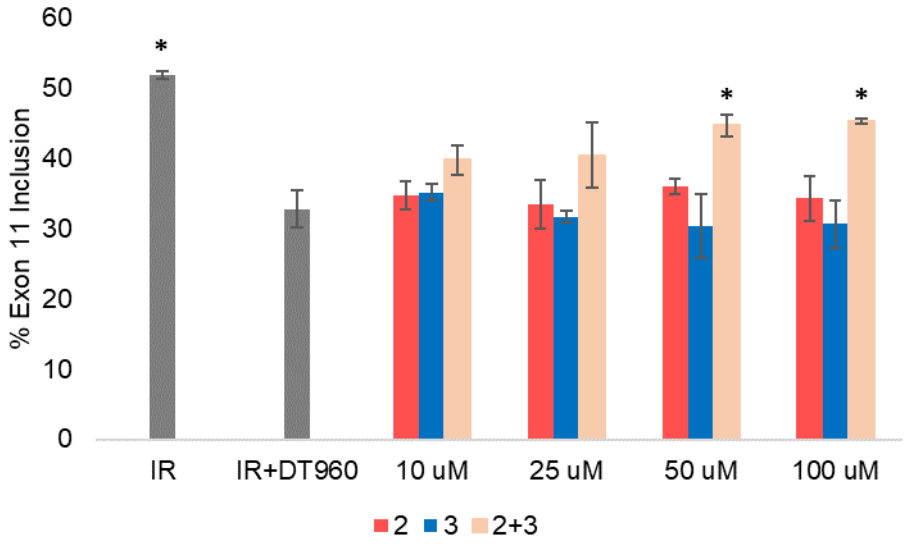


Figure 3.27. IR mis-splicing assay comparing monomers **2** and **3** to mixture of **2**+**3**. Splicing was rescued in a dose-dependent manner by **2**+**3**. * indicates $p < 0.05$ compared to the negative control.

3.6. REFERENCES

(1) Pennisi, E. ENCODE Project Writes Eulogy for Junk DNA. *Science* **2012**, *337* (6099), 1159–1161.

(2) Ross, S. J.; Revenko, A. S.; Hanson, L. L.; Ellston, R.; Staniszewska, A.; Whalley, N.; Pandey, S. K.; Revill, M.; Rooney, C.; Buckett, L. K.; Klein, S. K.; Hudson, K.; Monia, B. P.; Zinda, M.; Blakey, D. C.; Lyne, P. D.; Macleod, A. R. Targeting KRAS-Dependent Tumors with AZD4785, a High-Affinity Therapeutic Antisense Oligonucleotide Inhibitor of KRAS. *Sci. Transl. Med.* **2017**, *9*, 1–13.

(3) Moreno, A. M.; Fu, X.; Zhu, J.; Katrekar, D.; Shih, Y. R. V.; Marlett, J.; Cabotaje, J.; Tat, J.; Naughton, J.; Lisowski, L.; Varghese, S.; Zhang, K.; Mali, P. In Situ Gene Therapy via AAV-CRISPR-Cas9-Mediated Targeted Gene Regulation. *Mol. Ther.* **2018**, *26* (7), 1818–1827.

(4) Chi, X.; Gatti, P.; Papoian, T. Safety of Antisense Oligonucleotide and SiRNA-Based

Therapeutics. *Drug Discov. Today* **2017**, *22* (5), 823–833.

(5) Bahrami, A.; Aledavood, A.; Anvari, K.; Hassanian, S. M.; Maftouh, M.; Yaghobzade, A.; Salarzaee, O.; ShahidSales, S.; Avan, A. The Prognostic and Therapeutic Application of MicroRNAs in Breast Cancer: Tissue and Circulating MicroRNAs. *J. Cell. Physiol.* **2018**, *233* (2), 774–786.

(6) Donlic, A.; Morgan, B. S.; Xu, J. L.; Liu, A.; Roble, C.; Hargrove, A. E. Discovery of Small Molecule Ligands for MALAT1 by Tuning an RNA-Binding Scaffold. *Angew. Chemie - Int. Ed.* **2018**, *57* (40), 13242–13247.

(7) Liu, S.; Yang, Y.; Li, W.; Tian, X.; Cui, H.; Zhang, Q. Identification of Small-Molecule Ligands That Bind to MiR-21 as Potential Therapeutics for Endometriosis by Screening ZINC Database and in-Vitro Assays. *Gene* **2018**, *662* (April), 46–53.

(8) Matthes, F.; Massari, S.; Bochicchio, A.; Schorpp, K.; Schilling, J.; Weber, S.; Offermann, N.; Desantis, J.; Wanker, E.; Carloni, P.; Hadian, K.; Tabarrini, O.; Rossetti, G.; Krauss, S. Reducing Mutant Huntingtin Protein Expression in Living Cells by a Newly Identified RNA CAG Binder. *ACS Chem. Neurosci.* **2018**, *9* (6), 1399–1408.

(9) Gao, Z.; Cooper, T. A. Antisense Oligonucleotides: Rising Stars in Eliminating RNA Toxicity in Myotonic Dystrophy. *Hum. Gene Ther.* **2013**, *24* (5), 499–507.

(10) Rinaldi, C.; A Wood, M. J. Antisense Oligonucleotides: The next Frontier for Treatment of Neurological Disorders. *Nat. Publ. Gr.* **2017**.

(11) Provenzano, C.; Cappella, M.; Valaperta, R.; Cardani, R.; Meola, G.; Martelli, F.; Cardinali, B.; Falcone, G. CRISPR/Cas9-Mediated Deletion of CTG Expansions Recovers Normal

Phenotype in Myogenic Cells Derived from Myotonic Dystrophy 1 Patients. *Mol. Ther. - Nucleic Acids* **2017**, *9*, 337–348.

(12) Cerro-Herreros, E.; Sabater-Arcis, M.; Fernandez-Costa, J. M.; Moreno, N.; Perez-Alonso, M.; Llamusi, B.; Artero, R. MiR-23b and MiR-218 Silencing Increase Muscleblind-like Expression and Alleviate Myotonic Dystrophy Phenotypes in Mammalian Models. *Nat. Commun.* **2018**, *9* (1), 2482.

(13) Jenquin, J. R.; Coonrod, L. A.; Silverglate, Q. A.; Pellitier, N. A.; Hale, M. A.; Xia, G.; Nakamori, M.; Berglund, J. A. Furamidine Rescues Myotonic Dystrophy Type I Associated Mis-Splicing through Multiple Mechanisms. *ACS Chem. Biol.* **2018**, *13* (9), 2708–2718.

(14) Hsieh, W. C.; Bahal, R.; Thadke, S. A.; Bhatt, K.; Sobczak, K.; Thornton, C.; Ly, D. H. Design of a "Mini" Nucleic Acid Probe for Cooperative Binding of an RNA-Repeated Transcript Associated with Myotonic Dystrophy Type 1. *Biochemistry* **2018**, *57* (6), 907–911.

(15) Konieczny, P.; Selma-Soriano, E.; Rapisarda, A. S.; Fernandez-Costa, J. M.; Perez-Alonso, M.; Artero, R. Myotonic Dystrophy: Candidate Small Molecule Therapeutics. *Drug Discov. Today* **2017**, *22* (11), 1740–1748.

(16) Mirkin, S. M. Expandable DNA Repeats and Human Disease. *Nature*. 2007.

(17) Pearson, C. E. Slipped-Strand DNAs Formed by Long (CAG)(CTG) Repeats: Slipped-out Repeats and Slip-out Junctions. *Nucleic Acids Res.* **2002**, *30* (20), 4534–4547.

(18) Brook, D. J. et al. Molecular Basis of Myotonic Dystrophy: Expansionof a Trinucleotide (CTG) Repeat at the 3' End of a Transcript Encoding a Protein Kinase Family Member.

Cell **1992**, *68*, 799–808.

(19) Lin, X.; Miller, J. W.; Mankodi, A.; Kanadia, R. N.; Yuan, Y.; Moxley, R. T.; Swanson, M. S.; Thornton, C. A. Failure of MBNL1-Dependent Post-Natal Splicing Transitions in Myotonic Dystrophy. *Hum. Mol. Genet.* **2006**, *15* (13), 2087–2097.

(20) Charizanis, K. et al. Muscleblind-like 2-Mediated Alternative Splicing in the Developing Brain and Dysregulation in Myotonic Dystrophy. *Neuron* **2012**, *75* (3), 437–450.

(21) Zu, T. et al. Non-ATG-Initiated Translation Directed by Microsatellite Expansions. *Proc. Natl. Acad. Sci.* **2011**, *108* (1), 260–265.

(22) Krishnamurthy, V. M.; Estroff, L. A.; Whitesides, G. M. Multivalency in Ligand Design. In *Fragment-based Approaches in Drug Discovery*; 2006; Vol. 34, pp 11–53.

(23) Childs-Disney, J. L.; Yildirim, I.; Park, H.; Lohman, J. R.; Guan, L.; Tran, T.; Sarkar, P.; Schatz, G. C.; Disney, M. D. Structure of the Myotonic Dystrophy Type 2 RNA and Designed Small Molecules That Reduce Toxicity. *ACS Chem. Biol.* **2014**, *9* (2), 538–550.

(24) Bai, Y.; Nguyen, L.; Song, Z.; Peng, S.; Lee, J.; Zheng, N.; Kapoor, I.; Hagler, L. D.; Cai, K.; Cheng, J.; Chan, H. Y. E.; Zimmerman, S. C. Integrating Display and Delivery Functionality with a Cell Penetrating Peptide Mimic as a Scaffold for Intracellular Multivalent Multitargeting. *J. Am. Chem. Soc.* **2016**, *138* (30), 9498–9507.

(25) McMurray, C. T. Mechanisms of Trinucleotide Repeat Instability during Human Development. *Nature Reviews Genetics*. NIH Public Access November 2010, pp 786–799.

(26) Wong, C. H.; Nguyen, L.; Peh, J.; Luu, L. M.; Sanchez, J. S.; Richardson, S. L.; Tuccinardi, T.; Tsoi, H.; Chan, W. Y.; Chan, H. Y. E.; Baranger, A. M.; Hergenrother, P. J.;

Zimmerman, S. C. Targeting Toxic RNAs That Cause Myotonic Dystrophy Type 1 (DM1) with a Bisamidinium Inhibitor. *J. Am. Chem. Soc.* **2014**, *136* (17), 6355–6361.

(27) Luu, L. M.; Nguyen, L.; Peng, S.; Lee, J. Y.; Lee, H. Y.; Wong, C. H.; Hergenrother, P. J.; Chan, H. Y. E.; Zimmerman, S. C. A Potent Inhibitor of Protein Sequestration by Expanded Triplet (CUG) Repeats That Shows Phenotypic Improvements in a Drosophila Model of Myotonic Dystrophy. *ChemMedChem* **2016**, *11* (13), 1428–1435.

(28) Lee, J.; Bai, Y.; Chembazhi, U. V; Peng, S.; Yum, K.; Luu, L. M.; Hagler, L. D.; Serrano, J. F.; Chan, H. Y. E.; Kalsotra, A.; Zimmerman, S. C. Intrinsically Cell-Penetrating Multivalent and Multitargeting Ligands for Myotonic Dystrophy Type 1. *Proc. Natl. Acad. Sci.* **2019**, *116* (18), 8709–8714.

(29) Kiliszek, A.; Kierzek, R.; Krzyzosiak, W. J.; Rypniewski, W. Structural Insights into CUG Repeats Containing the "Stretched U-U Wobble": Implications for Myotonic Dystrophy. *Nucleic Acids Res.* **2009**, *37* (12), 4149–4156.

(30) Mao, J.; Desantis, C.; Bong, D. Small Molecule Recognition Triggers Secondary and Tertiary Interactions in DNA Folding and Hammerhead Ribozyme Catalysis. *J. Am. Chem. Soc.* **2017**, *139* (29), 9815–9818.

(31) NeuBase. Corporate Presentation January 2019 1. **2019**, No. January, 1–35.

(32) Yamada, T.; Miki, S.; Ni, L.; Nakatani, K. CGG Repeat DNA Assisted Dimerization of CGG/CGG Binding Molecule through Intermolecular Disulfide Formation. *Chem. Commun.* **2018**, *54*, 13072–13075.

(33) Ly, D. H.; Bahal, R.; Manna, A.; Bhunia, D.; Ra-Pireddy, S.; Cong, Y. Template-Directed

ΓPNA Synthesis and ΓPNA Targeting Compounds, 2014.

(34) Savkur, R. S.; Philips, A. V.; Cooper, T. A. Aberrant Regulation of Insulin Receptor Alternative Splicing Is Associated with Insulin Resistance in Myotonic Dystrophy. *Nat. Genet.* **2001**, *29* (1), 40–47.

(35) Jog, S. P.; Paul, S.; Dansithong, W.; Tring, S.; Comai, L.; Reddy, S. RNA Splicing Is Responsive to MBNL1 Dose. *PLoS One* **2012**, *7* (11), 1–8.

(36) Mills, J. E. J.; Dean, P. M. Three-Dimensional Hydrogen-Bond Geometry and Probability Information from a Crystal Survey. *J. Comput. Aided. Mol. Des.* **1996**, *10* (6), 607–622.

(37) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera - A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25* (13), 1605–1612.

(38) Long Island University. *Qualitative Tests for Carbonyls; Unknown Carbonyl*; 2015.

(39) Nguyen, L.; Luu, L. M.; Peng, S.; Serrano, J. F.; Chan, H. Y. E.; Zimmerman, S. C. Rationally Designed Small Molecules That Target Both the DNA and RNA Causing Myotonic Dystrophy Type 1. *J. Am. Chem. Soc.* **2015**, *137* (44), 14180–14189.

(40) Jahromi, A. H.; Nguyen, L.; Fu, Y.; Miller, K. A.; Baranger, A. M.; Zimmerman, S. C. A Novel CUGexp·MBNL1 Inhibitor with Therapeutic Potential for Myotonic Dystrophy Type 1. *ACS Chem. Biol.* **2013**, *8* (5), 1037–1043.

(41) Vichai, V.; Kirtikara, K. Sulforhodamine B Colorimetric Assay for Cytotoxicity Screening. *Nat. Protoc.* **2006**, *1* (3), 1112–1116.

CHAPTER 4: DYNAMIC COVALENT SCREEN FOR TARGETING AGENTS

4.1 INTRODUCTION

Myotonic Dystrophy Type 1 (DM1) and Huntington's Disease (HD) are both trinucleotide repeat expansion diseases (TREDs) that are caused by an extended d(CTG·CAG) repeating sequence in the DNA that leads to formation of toxic RNA and/or proteins.[1,2] Thus, these diseases provide ideal models to study multivalent nucleic acid targeting strategies. DM1 is characterized by an RNA gain-of-function in which the toxic r(CUG)^{exp} sequesters muscleblind-like proteins[3,4] and causes splicing defects, leading to symptoms including myotonia, insulin resistance, and cardiac defects.[5–8] HD is a neurodegenerative disease that leads to striatal and cortical degradation, causing defects in motor and cognitive functions as well as psychiatric disturbance, seizures, cardiac arrythmias, and muscle atrophy.[9] Many of the symptoms of HD overlap with DM1 and are caused by the same mechanisms: bidirectional transcription of d(CTG·CAG)^{exp} to form r(CUG)^{exp} and r(CAG)^{exp}, r(CUG)^{exp} sequestration of splicing proteins such as MBNL1, and translation of r(CUG)^{exp} and r(CAG)^{exp} to form toxic homopolymeric peptide sequences such as polyglutamine.[1] Although it is caused by the same repeat insertion as DM1, it is different in that the mutation in HD occurs in a coding region, specifically within the huntingtin (HTT) gene on chromosome 4.[1,10] In HD, the repeat expansion leads to formation of mutant huntingtin protein with an expanded polyglutamine region that causes formation of nuclear protein aggregates.[1]

Neither DM1 nor HD has a known cure. Treatment strategies for DM1 include antisense oligonucleotides (ASOs), CRISPR/Cas-9 mediated deletion, and small molecule therapeutics.[11–16] In HD, reduction of HTT protein aggregates through various therapeutic strategies has been shown to improve cognitive and motor performance in HD mouse models.[9,10,17–19] Here, we focus on the

potential of using small molecules to inhibit the formation of the toxic RNA and thus reduce the formation of toxic proteins, including the mutant HTT protein and homopolymeric peptides. This strategy would be advantageous in treating both DM1 and HD; transcription inhibition would result in decreased expression of (i) toxic r(CUG)^{exp} that sequesters MBNL1, (ii) r(CAG)^{exp} that undergoes RAN translation, and (iii) mutant proteins with extended polyglutamine that cause nuclear aggregation (HD).

Our previous target-guided screening approach utilized proximity-induced, irreversible azide-alkyne click as the ligation method.[22,23] We sought to employ a reversible approach where monomeric compounds could assemble on-template to form larger dimeric or oligomeric products that could be degraded upon dissociation from the template or degradation of template, re-forming the monomeric units (Figure 4.1). We envisioned that this reversible linkage could by achieved with a hydrolysable imine bond. Thus, in this screen, we utilized amine-aldehyde condensation to achieve template-selected, reversible assembly on the nucleic-acid target.

Dynamic covalent chemistry has been successfully employed by Miller (disulfide chemistry)[24] and Hargrove (amine chemistry)[25] to screen for RNA-binding agents. Herein, we have developed a target-guided screen that focuses on identifying novel small molecules that bind to and assemble

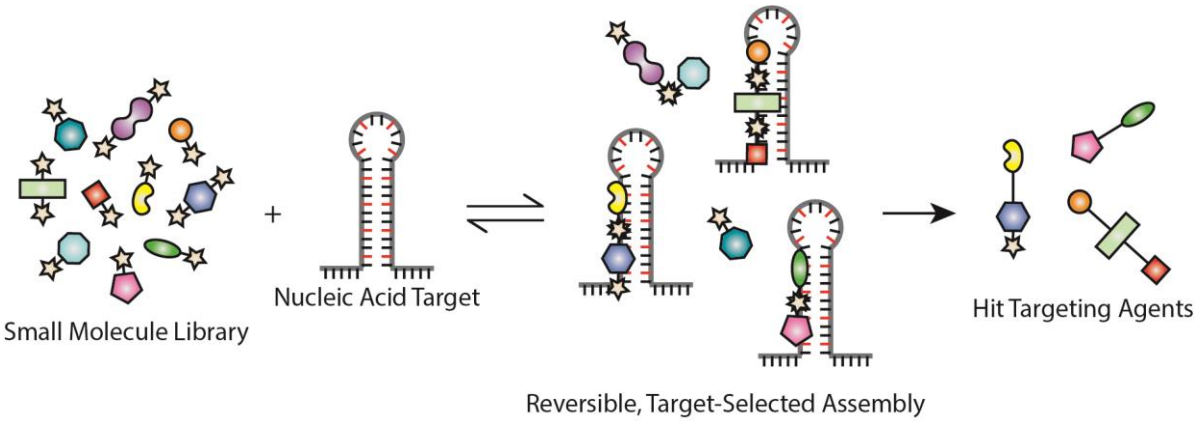


Figure 4.1. Dynamic Covalent Screen for Nucleic Acid-Targeting Agents. Aldehydes and amines (small molecule library) reversibly assemble on- and off-template, yielding template-selected hit combinations that bind to the nucleic acid target. The stars indicate reactive groups that can reversibly ligate to form multivalent targeting agents.

on DM1- and HD-relevant DNA targets. A particular interest for these diseases lies in targeting T-T and U-U mismatches that occur in the d(CTG)^{exp} and the r(CUG)^{exp}. Though we focus here on DM1 and HD, this simple, one-pot screening strategy performed at biologically relevant temperature and pH could be utilized for many other disease-relevant nucleic acid targets. Of note, identified hits did not need to be synthesized *ex vivo*, but were effective when treated in combination, suggesting potential for *in situ*, template-selected assembly.

4.2 RESULTS & DISCUSSION

4.2.1. Library Design

A library of aldehyde- and amine-functionalized compounds was designed based upon previously reported nucleic-acid targeting agents,[26–29] consisting primarily of general binders and thymine/uracil binders (Figure 4.2). The general nucleic acid binders include a bisamidinium unit, acridine, and neomycin. The bisamidinium unit utilized in **A1**, **A5**, **N1**, **N4**, **N11**, **N12**, **N20**, **N24**, and **N25** has been previously reported to bind in the major groove of A-form helices.[28] The acridine unit, utilized in **A7**, **A8**, **N9**, **N10**, **N23**, and **N26**, has been used as an intercalating agent.[30] Neomycin B, **N6**, is an RNA major-groove binder[31] and DNA triplex binder.[32] The nucleobase mimics that are included as potential T-T or U-U mismatch recognition units include melamine (**A1**, **A2**, **A5**, **A6**, **N1**, **N3**, **N8**, **N9**, **N11**, **N13**, **N14**, **N16**, **N20**, **N23**, and **N24**), amiloride (**N15**, **N21**, and **N22**), adenine (**N17**), and diaminopurine (**N18** and **N19**). Apart from adenine, each of these groups provides a donor-acceptor-donor hydrogen bonding interaction that complements the acceptor-donor-acceptor of the mismatched thymine or uracil bases. Adenine, as the natural base pair for T and U, provides a donor-acceptor match for the mismatched bases being targeted.

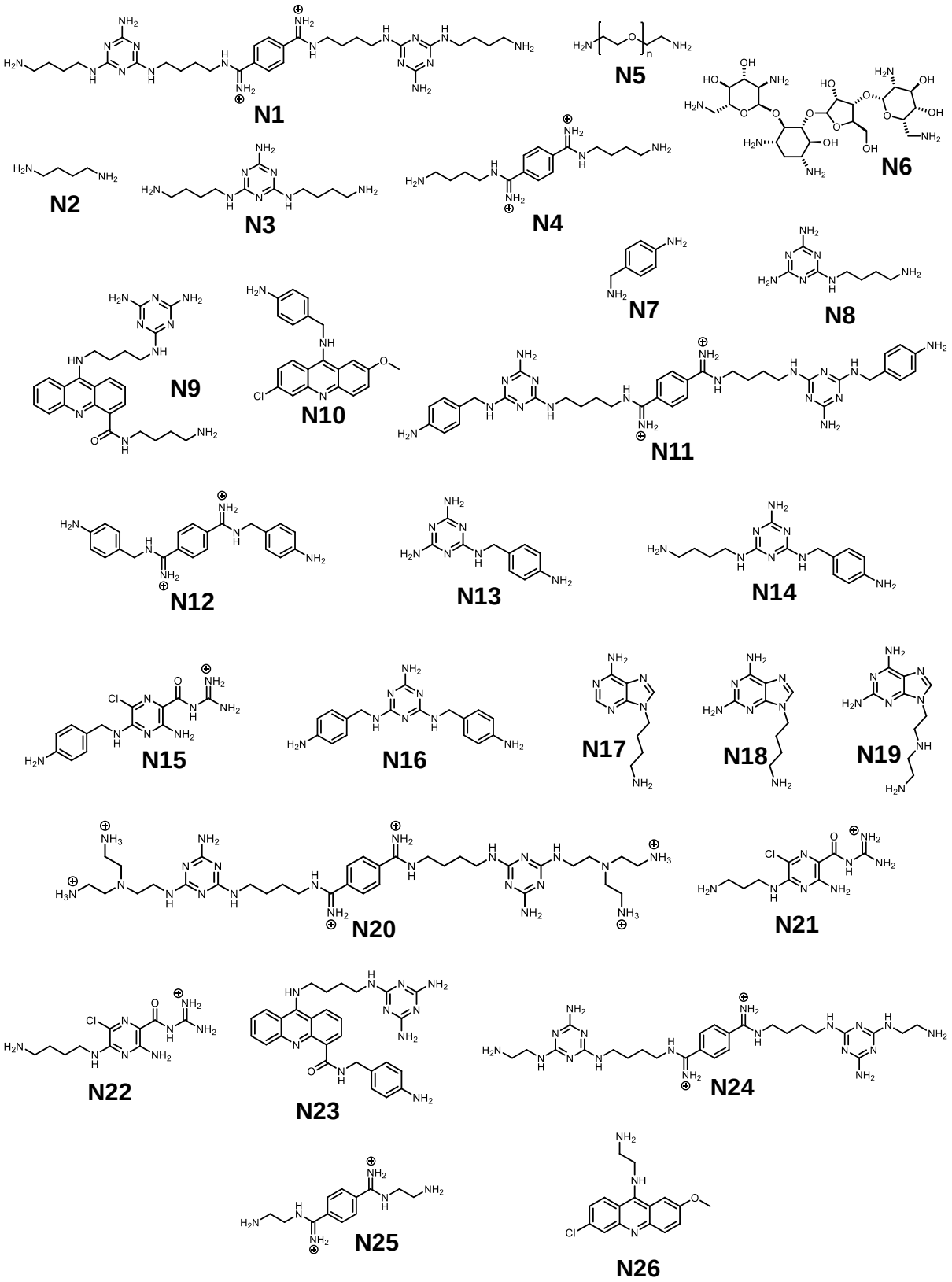


Figure 4.2. Dynamic Covalent Library. Amines **N1-N26** and aldehydes **A1-A8** were synthesized for reversible assembly on nucleic acids in the target-guided screen. (Cont. on next page)

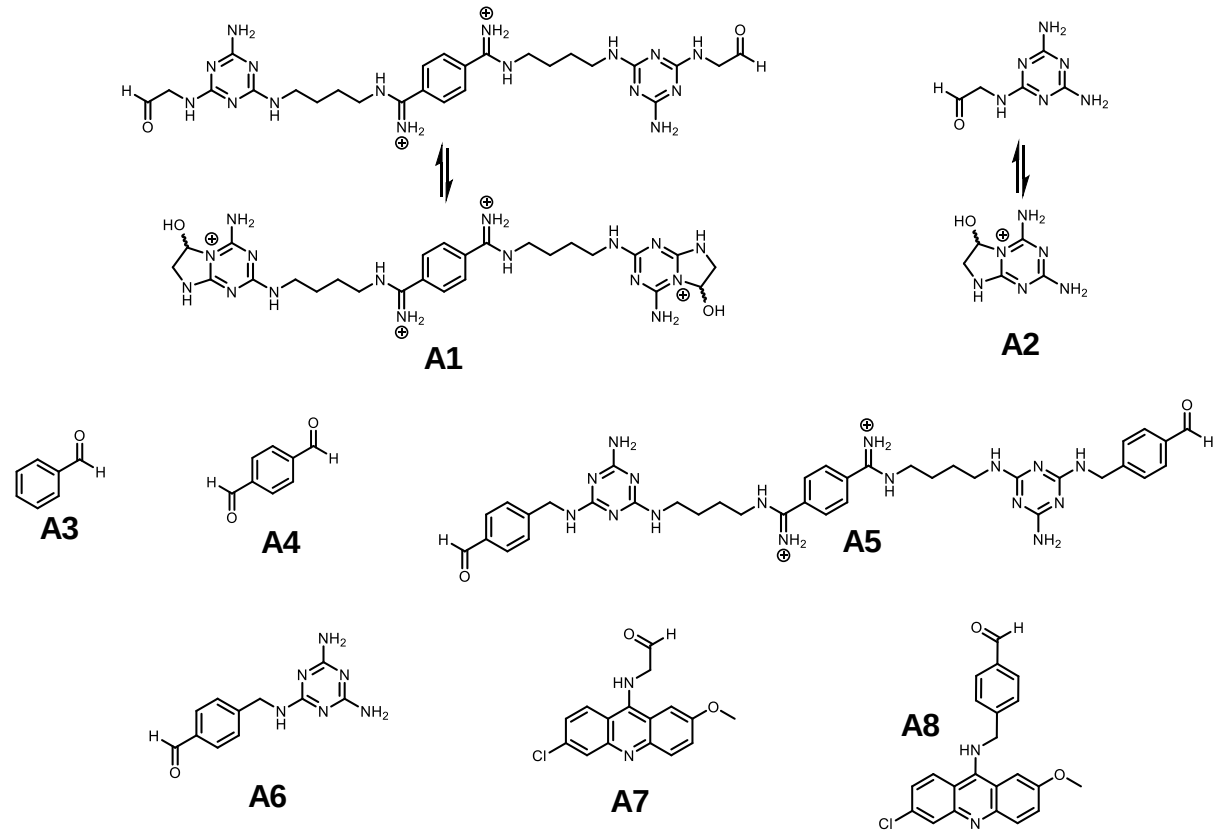


Figure 4.2. Dynamic Covalent Library. Amines **N1-N26** and aldehydes **A1-A8** were synthesized for reversible assembly on nucleic acids in the target-guided screen. (Cont. from previous page)

The library also includes several compounds with two reactive amines or aldehydes that are not known to bind the nucleic acid targets, but may act as short linkers (**A4**, **N2**, and **N7**) between compounds bound in adjacent sites that are not quite close enough to react. Compound **N5** may allow for the linkage of two compounds bound at distant sites on longer templates. Finally, the library includes compound **A3** that could form an imine, thus "capping" reactive end groups on the template and potentially allowing for hydrogen bonding with the phosphate backbone via the imine.

The library is made up of both larger, e.g., general binder linked to a targeting agent, and smaller, e.g., just a general binder, just a targeting agent, fragments that are functionalized with one, two, three, or four reactive aldehydes or amines to gain maximum diversity in the potential products formed. A mix of aromatic and aliphatic functionalities were utilized to obtain both rigid

and flexible linkers as well as varied reaction kinetics. These 8 aldehyde-containing compounds (**A1**-**A8**) and 26 amine-containing compounds (**N1**-**N26**) could potentially react to form as many as 2,173 dimeric and trimeric products.

4.2.2. Synthesis of Compounds.

Compounds **A3**, **A4**, **N2**, **N5**, **N6**, and **N7** were purchased from commercial sources, purified to achieve at least 95% purity, and utilized in the screen. All other compounds (**A1**, **A2**, **A5**-**A8**, **N1**, **N3**, **N4**, and **N8**-**N26**) were synthesized. See Appendix A, Section A.4.1 for synthetic methods.

4.2.3. Preliminary Study of Assembly

While this work was in progress, Hargrove reported a template-selected screen via dynamic covalent chemistry (DCC) with commercially available aromatic amines and a single amiloride-aldehyde.[33] In 2015, we began studying the *in situ* assembly of imines on DM1-relevant templates. The first compounds explored in this work are acridine-aromatic amine conjugate **N10** and pseudo-aldehyde **A1**. We initially performed assays with sodium cyanoborohydride used as the external reducing agent added at the end of the incubation period to allow for characterization of transient imine products as amines. A potential problem with this assay is that the addition of sodium cyanoborohydride may lead to a reductive amination that could be interpreted as product formation on or off template. Because sodium cyanoborohydride is generally accepted to be too gentle to reduce the aldehyde, the equilibrium between the imine that is formed on template and the aldehyde and amine starting materials may be depleted upon addition of the sodium cyanoborohydride, leading to the formation of more product (on or off template) by Le Chatelier's principle. To avoid this potential false positive readout, the reducing agent was later switched to sodium borohydride, a reagent that should reduce any imine bonds that form and any remaining

aldehyde starting material, effectively "freezing" the reaction mixture. This works under the assumption that the aldehyde is reduced faster than imines are formed.

There appears to be a new product peak in both the templated and the non-templated acridine-aniline + pseudo-aldehyde reactions (Figure 4.3). A spike-in of the acridine starting material was added to confirm that the peak was from a new species rather than a shift in the starting material (data not shown). Significant product formation was observed by HPLC after 1 d incubation at 37 °C. The HPLCs in Figure 4.3 were run after 2 d incubation at 37 °C. This new peak was observed at varied compound concentrations, including 10 µM compounds with 10 µM d(CTG)$_{16}$ template as well as 5 µM, 2.5 µM, 1 µM, and 0.5 µM compounds with 5 µM d(CTG)$_{16}$ template after 4 h incubation. An LC/MS study was performed to characterize this new peak. The mass of this new peak was found to be 259.07, consistent with the [M+5H]/5 peak of **N10+A1+N10** (calculated *m/z* = 261.12).

To investigate the possibility that the new peak observed was not just an aggregate formed

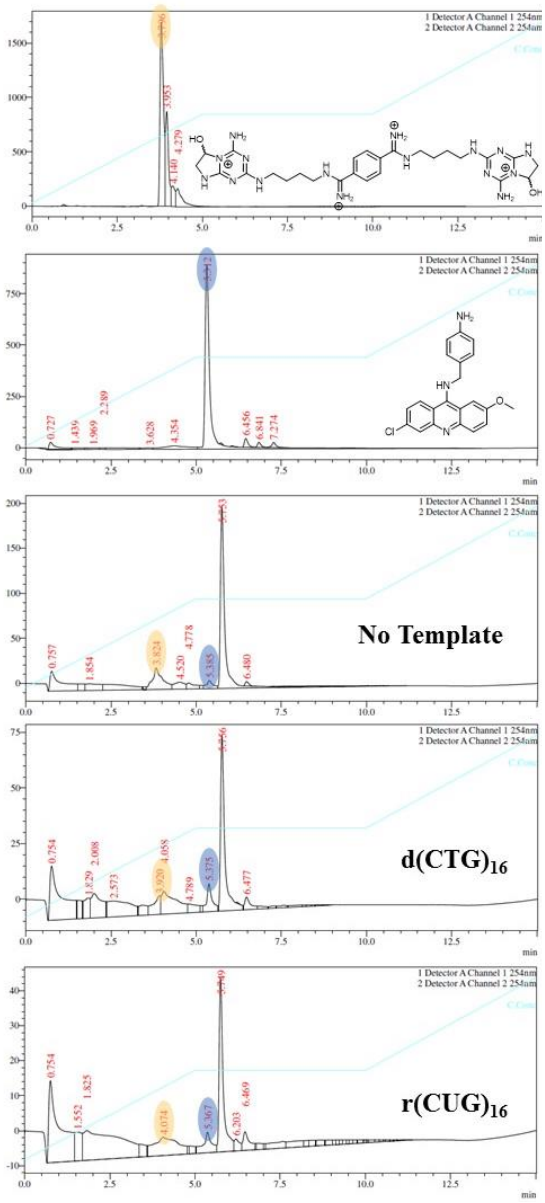


Figure 4.3. HPLC traces from condensation screen with pseudo-aldehyde **A1** (yellow) and aromatic amine **N10** (blue). Compounds were incubated in the presence or absence of template for two d and reduced with sodium cyanoborohydride. Compounds = 100 µM each, DNA = 10 µM, aqueous buffer = 2 mM Tris-HCl (pH 7), 2 mM KCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$. New peak appears around 5.75 min with water/acetonitrile gradient shown in blue trace.

off template by stacking of the acridine on the bisamidinium, but rather product formed on the template, the assay was run in the presence of 0.5% Triton-X 100 with 100 μM compounds and 10 μM $d(CTG)_{16}$. The Triton-X acts as a surfactant to break up aggregates, and thus should prevent any off template aggregates from forming. Consistent with the other assays, a new peak was observed after 4 h in the presence and absence of the template, shifted to about 6.4 min (data not shown), supporting that this new peak is a result of templated ligation and not aggregation. The slight shift observed in the retention time is likely due to the presence of detergent, as the monomer peaks were also shifted.

A similar HPLC assay was attempted with 100 μM concentration of aromatic amine **N11** and **A1**. Both the HPLC assay (100 μM **N11** and **A1**, 0% DMSO & 5% DMSO) and a MALDI assay (100 μM **N11** and **A1**, 5% DMSO) showed a peak consistent with **N11+A1** in the untemplated reaction following reduction after 4 h, but no product peaks in either $d(CTG)_{16}$ or $r(CUG)_{16}$ templated reactions. Similar results were obtained in PBS buffer. When a MALDI experiment was performed after incubation of $d(CTG)_{16}$ with **N11** but without the addition DNase after the incubation ($d(CTG)_{16}$ template was not degraded), no mass consistent with the DNA template was observed. Upon incubating **N11** and **A1** with $d(CTG)_{16}$ in the same buffer conditions above (0% DMSO) and reducing with sodium borohydride, a peak for the **N11+A1** product was observed by MALDI in the presence and absence of the template. Given that we were able to observe product with this MALDI-MS assay in the 2x2 screen and our previous successes with a library-based screening approach, we moved forward to screen the library of aldehydes and amines.

4.2.4. Library Screening.

With our library in hand, we set out to develop a one-pot screen in which we could quickly search for hit combinations that assemble specifically on the target nucleic acid. Thus, we

incubated the library with various nucleic acid templates ($d(CTG)_{16}$, $d(CAG)_{16}$, $r(CUG)_{16}$, $r(CUG)_{90}$, $r(CAG)_{90}$, random DNA duplex, or no template) in buffer at pH 7 for 24 h. Because aqueous hydrolysis of the imine makes the assembled products difficult to characterize, sodium borohydride was added after the incubation period to reduce the aldehydes and imines, thus "freezing" the equilibration of the library. MALDI-MS was utilized to screen for new peaks consistent with the mass of dimeric or trimeric products, including both homotrimers (e.g., **N22+A4+N22**) and heterotrimers (e.g., **A6+N3+A8**).

From the 34 monomers, there are a potential of 208 dimers, 286 homotrimers, and 1,579 heterotrimers, as not all arrangements of compounds can form trimers. The raw MALDI-MS data from three replicate assays was averaged and a heat map was constructed to represent the relative intensities of each peak relative to an internal standard, rutin. A representative heat map showing a few hit combinations is shown in Figure 4.4, with full heat maps shown in Figures 4.13 and 4.14. Based on the MALDI-MS data, 332 combinations comprised of 39 dimers and 293 trimers assembled in the assay.

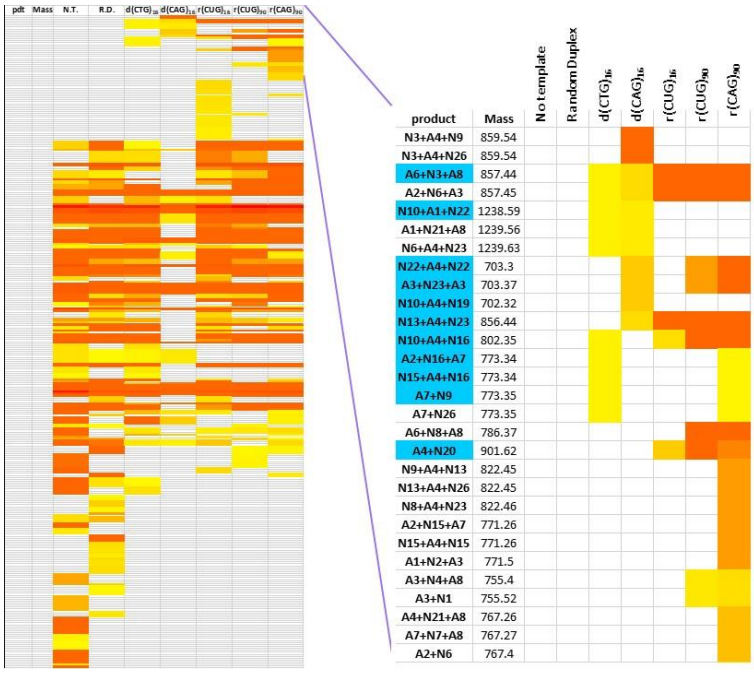


Figure 4.4. Representative Heat Map from MALDI-MS Screen. Aldehydes **A1-A8** and amines **N1-N26** were mixed with or without template. Raw MALDI-MS data was screened for new peaks consistent with dimers and trimers. The heat map represents the intensity as an average over 3 independent replicates. Coloring from red to yellow to white corresponds to normalized intensity, with red being the strongest intensity. Map is sorted from highest to lowest combined peak intensity on targets $d(CTG)_{16}$, $d(CAG)_{16}$, $r(CUG)_{16}$, $r(CUG)_{90}$, and $r(CAG)_{90}$. Blue highlight indicates further study in single point assay.

Of these 332 combinations, 88 (11 dimers, 77 trimers) assembled only on the random duplex and/or only in the buffer.

The 88 combinations that did not appear to specifically assemble include mostly conjugation of non-specific nucleic acid binders to non-binders, such as benzene or diaminobutane. For example, **A3+N25+A8** (benzene-bisamidinium-acridine), **N7+A4+N26** (benzene-benzene-acridine), and **A4+N25+A8** (benzene-bisamidinium-acridine) were all observed only on the random duplex. This is to be expected for the random duplex template, as each includes non-specific binders conjugated to non-binders. Other combinations that were hits in the random duplex tended to be conjugates that can stack or otherwise aggregate due to their hydrophobic nature. As an example, **A4+N9+A4** (benzene-acridine-melamine-benzene) was a hit only with the random duplex template.

The combinations that assembled only in the buffer reaction also tended to be made up of primarily hydrophobic fragments, including acridine, benzene, and/or multiple melamine units. It is interesting to note that most of the larger MW compounds that assembled only in the buffer reaction contain at least one bisamidinium unit. The amidinium groups could allow for increased water solubility of assembled products containing mainly hydrophobic fragments or the formation of a long linear oligomer instead of a large aggregate that would dissociate in MALDI-MS. A full table of the 88 combinations that assembled in the non-templated buffer reaction or the random duplex control is included in Table 4.3. These combinations were not further explored.

Although this assay allowed for detection of many potential hits, one possible challenge with MALDI-MS is that it may be not be able to detect the presence of a product that is bound tightly to the template. Thus, the strongest dimers and trimers may bind so tightly to the template that they cannot be characterized. The templates were not degraded in these assays because the

presence of the enzyme in the solution severely decreased the intensity of the MALDI-MS peaks in the low volume sample. Further, the enzyme added additional background peaks that could mask peaks from potential hits. A similar assay with the template degraded could yield different hits.

4.2.5. Single Point in vitro Transcription Inhibition Assay

Of the remaining 244 combinations, 24 dimers and 82 trimers were selected for analysis in a single point *in vitro* transcription assay (highlighted in blue in Figures 4.13 and 4.14) to choose the best combinations for further study. This assay was used to quickly test the numerous hits, as some masses correspond to more than one potential combination.

Most of the 106 combinations tested in the single point assay assembled on $d(CTG)_{16}$ in the MALDI-MS assay, with a few dimers and trimers tested that did not assemble on $d(CTG)_{16}$ tested as controls. Control dimers included those that assembled on at least one target RNA but not $d(CTG)_{16}$ (**A4+N20**, **A1+N21**, **A5+N14**), those that also assembled on random DNA and/or buffer (**A1+N17**, **A1+N21**, **A6+N23**, **A8+N23**, **A1+N15**, **A8+N11**), and those that assembled only in random DNA or buffer (**A5+N16**, **A1+N9**, and **A1+N26**). Trimers tested as controls included those that assembled on at least one target RNA but not $d(CTG)_{16}$ (**A3+N23+A3**, **N10+A4+N19**, and **N13+A4+N23**) and those that also assembled on random DNA and/or buffer (**N2+A4+N24**, **A7+N13+A7**, **N10+A4+N15**, **N1+A4+N11**, **N2+A5+N13**, and **N13+A1+N21**). Some combinations were tested that assembled on the $d(CAG)_{16}$ template but not the $d(CTG)_{16}$ template, including trimers **N22+A4+N22**, **A3+N23+A3**, **N10+A4+N19**, and **N13+A4+N23**. Importantly, trimers **N22+A4+N22**, **A3+N23+A3**, and **N10+A4+N19** have masses within 1 amu, and thus it is entirely possible that not all trimers identified from the MALDI-MS assay are template-selected products.

Of note, some combinations represent both dimers and trimers. For example, when **A5** and **N4** are mixed, a dimer or a trimer could form. This particular combination was a hit as both a dimer and a trimer. Some combinations could also represent two different trimers, such as the mixture of **A1** and **N7** that hit as both **A1+N7+A1** and **N7+A1+N7**.

Structure of the combinations was also considered when selecting hits for further testing in this single point assay. Thus, some combinations that appear to be hits from the heat map were not tested here. For example, dimer **A7+N26** was observed in the presence of $d(CTG)_{16}$ and $r(CAG)_{90}$, but the structure is two acridines linked together with no specific targeting agent. Because this di-acridine would likely bind non-specifically, it was deemed uninteresting and not studied further. Similarly, a peak consistent with trimer **A1+N21+A8** was observed in the MALDI-MS but is unlikely to form because **N21** has only one primary amine. Therefore, this matching MALDI-MS peak likely corresponds to a different trimer with the same mass, such as **N10+A1+N22**, a combination that was tested.

The capacity for the mixtures of compounds to inhibit the transcription of $d(CAG)_{90}$ to form $r(CUG)_{90}$ *in vitro*

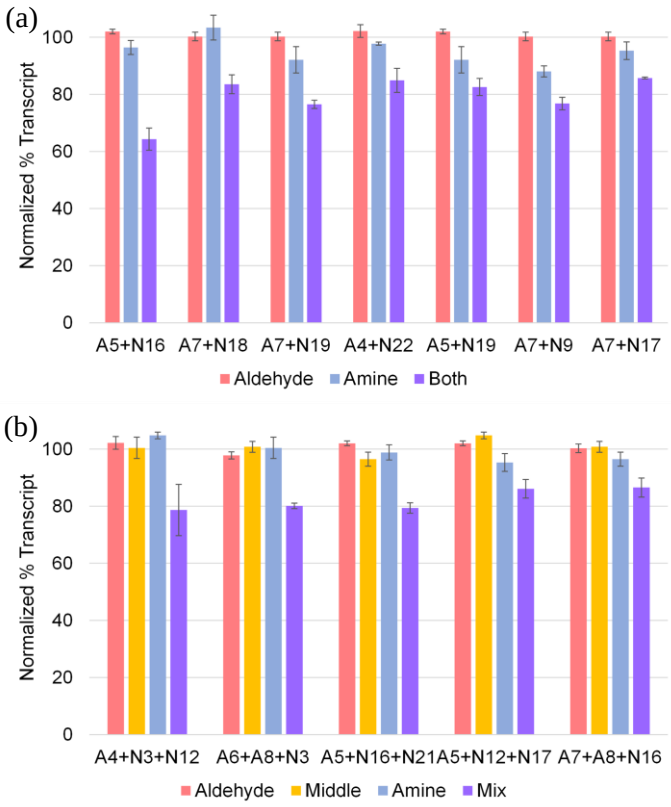


Figure 4.5. Representative Data from Single Point *in vitro* Transcription Inhibition Assay. Hit combinations (10 µM each) were incubated with a pSP72 plasmid containing $d(CTG\ CAG)_{90}$. The output $r(CUG)_{90}$ transcript was quantified to determine the normalized % transcript relative to the monomers for (a) dimers and homotrimers and (b) heterotrimers. Results are the average of at least 3 replicates, error reported as standard error.

was tested at a concentration of 10 μM each. The % transcription inhibition was compared to the monomeric units on their own to screen for compounds that cooperatively inhibited transcription, perhaps through assembly. A representative sample of the results is included in Figure 4.5, with full results shown in Figures 4.15 and 4.16. It is important to note that if a combination was selected from a homotrimer, for example **N22+A4+N22**, it was recorded as **A4+N22** and included in Figure 4.15 along with the dimeric hits even though the combination could form a dimer or a trimer.

Notably, none of the combinations tested as controls (listed above) inhibited transcription more strongly than the monomeric components except for **A5+N16** and **N1+A4+N11**. The combination **A5+N16** was observed only in the buffer, but strongly inhibited transcription. It could be that **A5+N16** was bound strongly enough to the template that its product peak was not observed in the MALDI-MS among the other product peaks, pointing to a potential limitation of this assay. Thus, we tested **A5+N16** in a 2x2 screen (see section 4.2.6). Similarly, combination **N1+A4+N11** assembled only off template and on r(CUG)$_{16}$. It is not known whether the compounds bind first and react on the template, react off-template and the ligated products are selectively amplified via protection from hydrolysis resulting from binding the template, or a mixture of both.

Interestingly, the combination **A4+N22**, selected from the trimeric hit **N22+A4+N22**, inhibited transcription even though it assembled on the d(CAG)$_{16}$ template and not the d(CTG)$_{16}$ template. This result suggests that binding either side of the expanded repeats may inhibit transcription by stalling the polymerase. The other three trimers that assembled on d(CAG)$_{16}$ but not d(CTG)$_{16}$ did not inhibit transcription more strongly than the monomeric components. Consistent with the hypothesis that the three trimers tested with the same mass may not actually all be hits, only **N22+A4+N22** and **N10+A4+N19** inhibited transcription and not **A3+N23+A3**. The other trimer

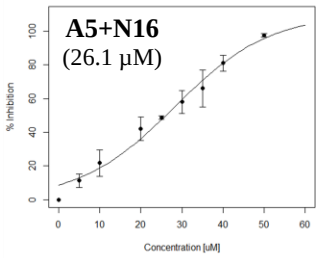
tested here that bound d(CAG)$_{16}$ but not d(CTG)$_{16}$, **N13+A4+N23**, inhibited transcription but not more strongly than its monomeric parts.

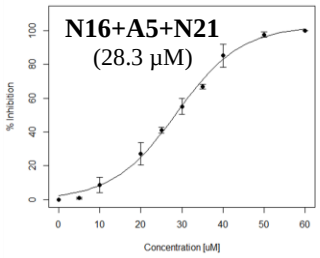
From this single point assay, combinations with the strongest inhibition and largest difference between the monomers and the combination were selected for further study.

4.2.6. IC$_{50}$ Determination

A total of 15 combinations (highlighted in green in Figures 4.15 and 4.16) were selected for further study of their ability to inhibit transcription, including **A4+N22**, **A5+N16**, **A5+N19**, **A6+N12+A7**, **A6+N3+A7**, **A6+N3+A8**, **A7+N16+A8**, **A7+N9**, **A7+N17**, **A7+N18**, **A7+N19**, **N12+A5+N17**, **N16+A5+N21**, **N3+A4+N10**, and **N3+A4+N12**. Each combination was tested at several concentrations, and half maximal inhibitory concentration (IC$_{50}$) values were calculated for 11 of these hits (Figures 4.6, 4.17). It is important to note that the other 4 compounds (**A4+22**, **A6+N3+A8**, **N3+A4+N10**, and **N3+A4+N12**) did inhibit transcription, but with IC$_{50}$ values greater than 60 µM. These IC$_{50}$ values were not calculated here were not more accurately determined here because of their comparatively weak binding.

Compound **A4+N22** is an amiloride (**N22**) conjugated to a benzyl imine (**A4**) and could also assemble to form **A4+N22+A4** or **N22+A4+N22**. This





Compound	Calc'd IC$_{50}$ (µM)
A5+N16	26.1 ± 0.6
N16+A5+N21	28.3 ± 0.3
N12+A5+N17	34.2 ± 0.6
A7+N9	36.2 ± 0.2
A7+N17	38.2 ± 0.4
A7+ N16+A8	39.8 ± 0.4
A7+N18	39.8 ± 0.3
A7+N19	41.6 ± 0.7

Compound	Calc'd IC$_{50}$ (µM)
A6+N12+A7	42.9 ± 0.2
A6+N3+A7	44.0 ± 0.4
A5+N19	53.7 ± 1.7
A4+N22	>60
A6+N3+A8	>60
N3+A4+N10	>60
N3+A4+N12	>60

Figure 4.6. IC$_{50}$ Determination for Hit Combinations. Hit combinations of aldehyde and amine were tested at multiple concentrations and IC$_{50}$ values were calculated for the inhibition of transcription to from r(CUG)$_{90}$. Results are the average of at least 3 replicates and error is standard error. See Figure 4.17 for other IC$_{50}$ curves.

combination was selected as a hit due to the assembly of trimer **N22+A4+N22** on the $d(CAG)_{16}$ template. The amiloride-benzyl di-imine-amiloride conjugate could bind $d(CAG)_{16}$, but it did not inhibit transcription as strongly as those ligands that bound $d(CTG)_{16}$. The reason for this observation could be threefold: (1) Trimer **N22+A4+N22** could bind more weakly to $d(CAG)_{16}$ than other combinations bind to $d(CTG)_{16}$, (2) The $d(CAG)_{16}$ hairpin, induced here with a GGG/CCC toehold, may not form as reliably as the d(CTG) hairpin during transcription, thus providing less opportunity for binding, and (3) Binding the (CAG) side of the expanded repeat could have less of an impact on transcription than binding to the (CTG) side. Although the (CAG) strand in the coding strand, expanded d(CAG) tends to be less structured than expanded d(CTG) regions that typically form hairpins. It is hypothesized that binding tightly to the d(CTG) hairpin can stall the polymerase and prevent bidirectional transcription. As only one (CAG) binder was tested here, further studies would be needed to make a stronger conclusion about the relationship between (CAG) binding and transcription inhibition and this relationship is beyond the scope of this work.

A few of the compounds with IC_{50} values greater than 60 μM have interesting structures. For example, **A6+N3+A8** is a melamine-melamine-acridine conjugate linked through a rigid aromatic unit in both connections. Similarly, **N3+A4+N10** is a melamine-benzyl imine-acridine conjugate with the first linker being the flexible hydrocarbon chain of **N3** and the second being a benzene ring. Compound **N3+A4+N12** has similar structure to **N3+A4+N10** but has a bisamidinium in place of the acridine. As noted previously, these compounds do inhibit transcription at a low level, but their IC_{50} values were not calculated here.

Of the 11 combinations for which IC_{50} values were calculated, the lowest was **A5+N16** at 26.1 ± 0.6 μM. This conjugate would be a melamine-bisamidinium-melamine-melamine conjugate,

with the dynamic covalent linkage being through two rigid benzene rings. As **A5** is a derivative of our previously reported ligands,[28,29] it is not surprising that the extension of the ligand to include another melamine enhances inhibition. Trimeric combination **N16+A5+N21** had the next lowest IC_{50} value of 28.3 ± 0.3 μM. This is interesting because the dimeric **A5+N16** has a very similar IC_{50} value, suggesting that addition of **N21** (an amiloride) does not further enhance the inhibition of transcription.

Combinations **A7+N17**, **A7+N18**, and **A7+N19** have similar structure and very close IC_{50} values of 38.2 ± 0.4, 39.8 ± 0.3, and 41.6 ± 0.7 μM, respectively. The structure of these compounds is an acridine linked to an adenine (**N17**) or diaminopurine (**N18**, **N19**). This result suggests that an acridine-adenine conjugate may be effective at binding DNA and inhibiting transcription to form $r(CUG)_{90}$. Further, the adenine and diaminopurine seem to be interchangeable, suggesting that the additional hydrogen bonding interaction from the NH_2 group on the 6-position of diaminopurine does not significantly enhance the interaction and may add steric bulk at the binding site.

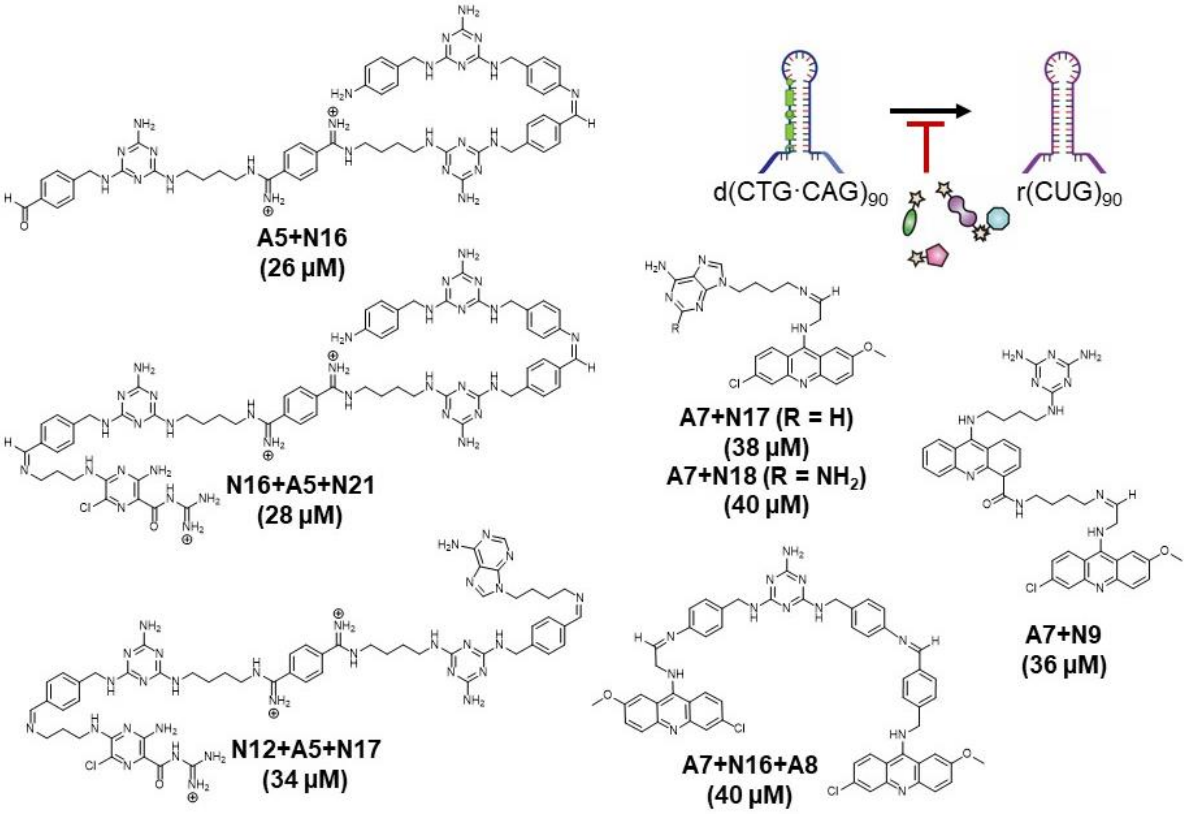
Several of the compounds that inhibited transcription in this concentration range include similar building blocks, including **A6+N3+A7** (melamine-melamine-acridine) with IC_{50} of 44.0 ± 0.4 μM, **A7+N16+A8** (acridine-melamine-acridine) with IC_{50} of 39.8 ± 0.4 μM, and **A7+N9** (acridine-acridine-melamine) with IC_{50} of 36.2 ± 0.2 μM. Our previously reported acridine-melamine conjugate[27] bound selectively to T-T and U-U mismatches, supporting that these building blocks can be used to form effective inhibitors of transcription.

The acridine building block was also found in hit **A6+N12+A7**, a melamine-bisamidinium-acridine conjugate with IC_{50} of 42.9 ± 0.2 μM. Other hits include bisamidinium-melamine-bisamidinium-melamine-adenine **N12+A5+N17** with IC_{50} of 34.2 ± 0.6 μM and melamine-

bisamidinium-melamine-diaminopurine **A5+N19** with IC_{50} of 53.7 ± 1.7 μM. The diverse scaffolds and fragments included in these hits, compared to the lack of diversity of hits obtained in our previously reported click screen,[23] suggests that the dynamic screen afforded more variety as the equilibration occurred in the presence of template. These compounds will be further explored for their potential to treat DM1 and HD.

4.2.7. Further Study of Hits

The combinations with the two lowest calculated IC_{50} values were selected for further study, specifically **A5+N16** and **N16+A5+N21**. To determine whether these combinations could assemble on the DNA template to form dimeric and/or trimer products, each combination was incubated with the template individually, as opposed to the one-pot screen from which the hits were initially selected. When combination **A5+N16** was incubated with the $d(CTG)_{16}$ template, a prominent peak was observed at 1081.729, consistent with the M+H peak for **A5+N16** (calc'd



1081.6081) along with a smaller peak at 1399.940 corresponding to the M+H peak for **N16+A5+N16** (calc'd 1399.7786) (Figure 4.18). Although MALDI-MS is not quantitative, based on peak intensity it seems that this combination favors formation of the dimeric product on the template. This is consistent with the screen data, in which the dimer **A5+N16** was observed as a hit, but no peak with mass of the trimer was observed. It is likely that a small peak for the trimer would have been buried in the additional peaks of the one-pot screen. Thus, it seems that the dimeric product is favored in this combination.

Interestingly, the MALDI-MS peaks observed when combination **N16+A5+N21** was incubated with $d(CTG)_{16}$ included 1081.734 (**A5+N16**) and a small peak at 1399.934 (**N16+A5+N16**), but not **N16+A5+N21** (calc'd 1349.7032) (Figure 4.19). The combination **N16+A5+N21** was initially observed in the screen as a weak hit in both the buffer reaction and the $d(CTG)_{16}$ templated reaction. Thus, it is possible that the dimer **A5+N16** is formed more readily and/or more favorably over trimer **N16+A5+N21**. As mentioned earlier, the similar IC_{50} values of these combinations suggested that the additional amiloride unit did not enhance the activity of this compound. Consistent with that observation, the MALDI-MS data suggests that the dimeric **A5+N16** product is formed preferentially over trimeric products and thus determines the IC_{50} value.

We also further studied these hit combinations via *in vitro* transcription inhibition assays. Notably, neither **A5+N16** or **N16+A5+N21** inhibited the formation of a random duplex template (Figure 4.20), supporting the potential for specific inhibition with these combinations. This is particularly important for combination **A5+N16** that was observed as a hit in the buffer reaction.

These two combinations were also compared to their individual monomers in an *in vitro* transcription inhibition assay (Figure 4.7). Monomers **A5**, **N16**, and **N21** all showed negligible transcription inhibition up to 60 μM, compared to the combinations **A5+N16** and **N16+A5+N21**

that showed full inhibition at 60 μM. This result demonstrates the cooperative effect of the hits found in the DCC screen. Although the transient nature of the linkage makes the product difficult to characterize, the cooperative effect of the combinations along with the MALDI-MS results supports the formation of dimeric and trimeric products on the template *in situ*.

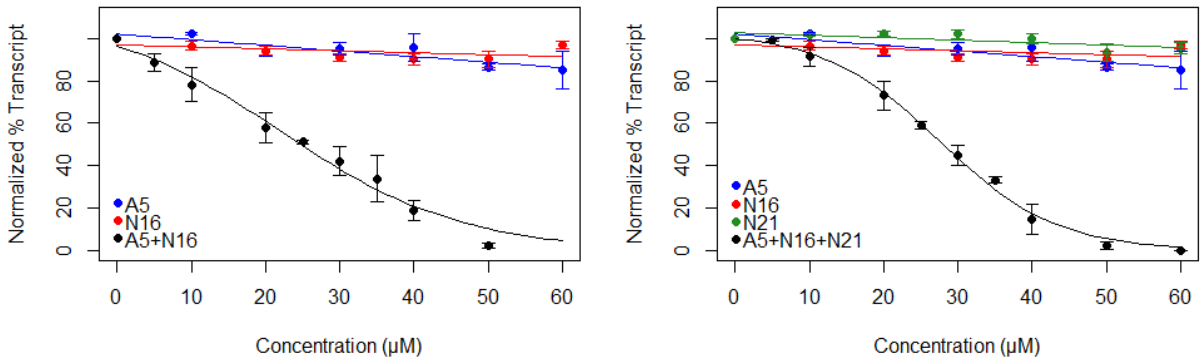


Figure 4.7. *in vitro* Transcription: Comparison of Monomers and Combinations. Hit combinations and monomers were tested at multiple concentrations for the inhibition of transcription to from r(CUG)$_{90}$. Results are the average of at least 3 replicates and error is reported as standard error.

4.3 CONCLUSIONS

From a diverse library consisting of general nucleic acid binders and nucleobase mimics for targeting T-T and U-U mismatches, a variety of hits were found that successfully inhibited transcription of d(CAG)$_{90}$ thereby preventing the formation of toxic r(CUG)$_{90}$. These hits provided new structural combinations to form multivalent targeting agents. With the goal of this screen being to identify new hit compounds for treatment of DM1 and HD, only the hits with the best cooperative inhibition were studied. Of note, weaker binding hits could also be detected from this screening method. This method could be utilized to identify hit compounds that can assemble reversibly *in situ* to form a multivalent targeting agent for DM1 and HD in addition to other diseases caused by nucleic acid repeats such as amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD).

4.4. METHODS

4.4.1. Synthetic Methods

Refer to the Appendix A, Section A.4.1 for synthesis of **A1**, **A2**, **A5-A8**, **N1**, **N3**, **N4**, and **N8-N26**.

4.4.2. Dynamic Covalent Library Screen

<u>*Nucleic Acid Sequences*</u>

$d(CTG)_{16}$ = 5’ – GGG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CCC – 3’ (purchased from IDT, MW = 16,554.6 g/mol)

$d(CAG)_{16}$ = 5’ – GGG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CCC – 3’ (purchased from IDT, MW = 16,698.8 g/mol)

$r(CUG)_{16}$ = 5’ – GGG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CCC – 3’ (purchased from IDT, MW = 17,194.1 g/mol)

Random DNA Duplex = 5'-GCC ATC ACG GAT CAC GTC-3' // 3’-CGG TAG TGC CTA GTG CAG-5’ (purchased from IDT, MW = 11,000.2 g/mol)

$r(CUG)_{90}$ = 5’ – G GGAGACCGGC AGAUCUGAUA UCAUCGAUGA AUUC CUG GAGCUC GGUACCCGG G – 3’(synthesized from pSP72 $(CTG{\cdot}CAG)_{90}$ plasmid, MW = 102,571.1 g/mol, calculated using IDT OligoAnalyzer Tool)

$r(CAG)_{90}$ = 5’ – GAACUCGAGC AGCUGAAGCU UGCAUGCCUG CAGGUCGACU CUAGAGGAUC CCCGGGUACC GAGCUC CAG CAG CAG CAG CAG CAG CAG CAG

CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG G – 3'(synthesized from pSP72 $(CTG{\cdot}CAG)_{90}$ plasmid, MW = 109,737.8 g/mol, calculated using OligoAnalyzer)

Condensation Screen MALDI-MS Assay

Compounds (100 µM each) and DNA or RNA templates (10 µM for $d(CTG)_{16}$, $d(CAG)_{16}$, and $r(CUG)_{16}$ and 500 nM for $r(CUG)_{90}$, and $r(CAG)_{90}$) were incubated in aqueous buffer consisting of 2 mM each of KCl, $MgCl_2$, $CaCl_2$, and Tris-HCl, pH 7 in a final volume of 5 µL. The solutions were annealed in buffer before addition of compounds on a thermocycler at 95 °C for 5 min and allowed to cool by 0.1 °C/s to 25 °C. The solutions were incubated at 37 °C for 24 h before reduction with 0.3 µL of 100 mM sodium borohydride. Following 6 h incubation at 37 °C (for reduction step), a 1 µL aliquot was removed for MALDI-MS (run in 1 µL DHB matrix). Raw MALDI-MS data was screened for masses matching within ±1 *m*/*z* of the calculated masses for all possible dimers and trimers using RStudio.

4.4.3. *in vitro* transcription of $d(CTG{\cdot}CAG)_{90}$ to form $r(CUG)_{90}$ and $r(CAG)_{90}$

See Appendix A, Section A.1.11

4.4.4. *in vitro* Transcription Inhbition Assays

Refer to Appendix A, Section A.1.8 for general *in vitro* transcription inhibition procedures.

4.4.5. 2x2 MALDI-MS Assay

Compounds (100 µM each) and 10 µM d(CTG)$_{16}$ were incubated in aqueous buffer consisting of 2 mM each of KCl, $MgCl_2$, $CaCl_2$, and Tris-HCl, pH 7 in a final volume of 50 µL. The solutions were annealed in buffer before addition of compounds on a thermocycler at 95 °C for 5 min and cooled at a rate of 0.1 °C/s to 25 °C. The solutions were incubated at 37 °C for 24 h before reduction with 2.5 µL of 100 mM sodium borohydride. Following 6 h incubation at 37 °C (for reduction step), a 1 µL aliquot was removed for MALDI-MS (run in 1 µL DHB matrix).

4.5. SUPPLEMENTAL FIGURES AND DATA

4.5.1. Preliminary reaction screening

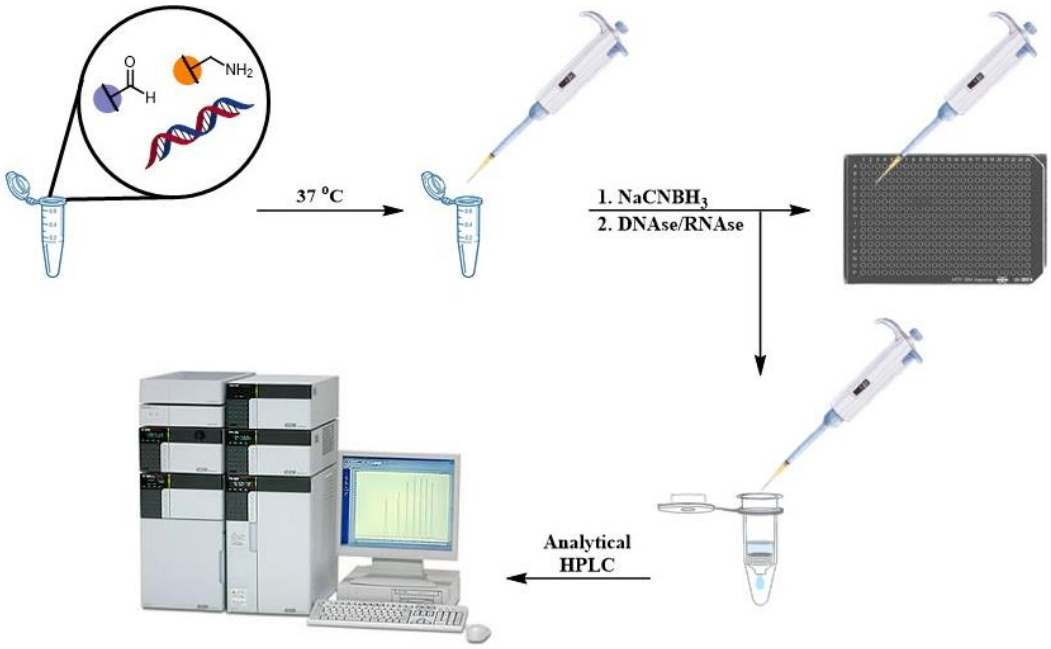


Figure 4.8. Condensation Screening Assay. Compounds are mixed with or without nucleic acid template and incubated at 37 °C. Aliquots are removed at several time points and analyzed by MALDI-MS. Samples with new peaks in the MALDI would be filtered through a silica spin column to remove excess $NaCNBH_3$ and analyzed by HPLC. Buffer conditions tested include 300 mM NaCl/20 mM MOPS (pH 7), 100 mM NaCl/10 mM phosphate buffer (pH 6.5), 2 mM Tris-HCl (pH 7.0)/2 mM KCl/2 mM $MgCl_2$/2 mM $CaCl_2$. Image credits: Image credits: https://www.starlabgroup.com, https://www.indiamart.com, http://www.ms-textbook.com, https://www.protocols.io, HPLC from Shimadzu.

Nucleic Acid Sequences: same as above in Section 4.4.2

Preliminary HPLC Study with A1, N10, and N11:

Compound solutions were incubated with freshly annealed $d(CTG)_{16}$ (master solution 2 mM) or $r(CUG)_{16}$ (master solution 2 mM). An aliquot of nucleic acid solution was annealed on a thermocycler at 95 °C for 5 min and cooled at a rate of 0.1 °C/s to 25 °C. To each reaction to contain nucleic acid template was added the proper volume of the freshly annealed DNA or RNA solution. A master mix of buffer solution was prepared and the proper aliquot of the solution was added to each Eppendorf tube. ACR-ANI was used as a 2 mM stock solution in DMSO Ald-CHL-Ald and Amine-CHL-Amine were used as 100 μM stock solutions in molecular biology grade water. Each reaction was set up in a final volume of 50 μL with molecular biology grade water. The Eppendorf tubes were incubated at 37 °C.

For MALDI analysis, 1 d samples: A 10 μL aliquot was removed from the reactions and 5 μL 100 μM $NaCNBH_3$ (aqueous) was added. The aliquots were incubated at 37 °C for 60 min or 16 h on a shaker and 1 μL DNase I (NEB) was added to each tube to degrade the template. The tubes were incubated at 37 °C for 20 min. MALDI-MS of each sample was measured (matrix: DHB) by mixing 1 μL matrix with 1 μL sample on the MALDI plate. The sample was run through a Zip Tip by SCS CORES Staff before spotting on MALDI plate. MALDI spectra were screened for new mass peaks. No significant new peaks were observed.

For HPLC analysis, 3 d samples: A 5 μL aliquot was removed from the reactions and 2.5 μL 100 μM $NaCNBH_3$ (aqueous) was added. The aliquots were incubated at 37 °C for ~16 h and filtered through an alumina column. Analytical HPLC was run with gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min. The dialdehyde peak persisted throughout the analysis.

For HPLC analysis, 7 d samples: A 10 μL aliquot was removed from the reactions and 1 μL 10 mM $NaCNBH_3$ (aqueous) was added. The aliquots were incubated at 37 °C for ~16 h and filtered through an 0.2 μm syringe filter. To each tube was added 20 μL DMSO, and tubes were vortexed to dissolve precipitates. Analytical HPLC was run with gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min. The peak from the dialdehyde starting material persisted, but also observed a new peak at 5.742 min. To check to be sure this was not the acridine starting material, a spike-in of 1 μL ACR-ANI (2 mM) was added to the HPLC samples and re-run. The peaks are indeed separate, this peak appears to be from a new product.

Master solutions $MgCl_2$ 100 mM (aq), $CaCl_2$ 100 mM (aq), KCl 100 mM (aq), Tris HCl pH 7 100 mM (aq), **A1** 1 mM (aq), **N11** 10 mM (aq), **N10** 100 mM (DMSO), $(CTG)_{16}$ 2 mM (aq), $(CUG)_{16}$ 2 mM (aq). HPLC samples were run in 5% DMSO to solubilize everything. A master solution containing proper aliquots of molecular biology grade water, tris HCl, KCl, $MgCl_2$, and $CaCl_2$ was mixed. The procedure was followed as reported above. A 25 μL aliquot of each reaction was removed after 45 min and after 1 d and 1 μL 10 mM NaCNBH3 was added. The aliquots were incubated at 37 °C overnight. To each aliquot was added 1 μL of either molecular biology grade water, DNase I, or RNase H, depending on the contents of the reaction, and the solutions were incubated at 37 °C for 15 min. To each aliquot was added 2.5 μL DMSO and 46.5 μL molecular biology grade water. 50 μL of solution was injected for analytical HPLC. The retention times shown in bold indicate the peaks with most area.

Sample	Peaks (retention time, min)
Buffer Only	0.530, 0.938 (all broad, weak)
DNA Only	0.380, 0.756, 2.005, 3.430, 4.103, 5.391, 5.771 (all weak)
RNA Only	0.397, 0.743, 4.080, 5.365, 5.763 (all weak)
N11 + **A1** (no template)	0.751, 3.795 (dialdehyde), 3.971, 4.110
N10 + **A1** (no template)	0.290, 0.747, 3.802, 3.943, 4.488 (weak), 5.364 (acridine), **5.741** (strongest), 6.459
N11 + **A1** + DNA	0.747, 1.996, 2.358, 3.470 (weak), 3.808 (dialdehyde), 3.976, 4.155 (all weak)
N10 + **A1** + DNA	0.428, 0.750, 2.020, 4.064, 5.376 (acridine), **5.759**, 6.471
N11 + **A1** + RNA	0.475, 0.751, 1.525 (weak), 3.790 (dialdehyde), 3.974, 4.147, 5.346
N11 + **A1** + RNA	0.748, 4.083 (weak), 5.368 (acridine), **5.752** (strongest), 6.471

Table 4.1. HPLC peaks observed in preliminary screening at 45 minutes. Previously observed retention times for starting materials in same gradient: Dialdehyde **A1** 3.796 min, Acridine-Aniline **N10** 5.312 min, Diamine **N11** 4.094 min. Note that the "Buffer Only" and "DNA/RNA Only" peaks are not reported in the subsequent reactions.

Sample	Peaks (retention time, min)
Buffer Only	0.756, 4.325, 5.364, 5.763 (all weak)
DNA Only	0.761, 2.029, 2.842, 3.029, 3.401, 3.921, 4.056, 4.804, 5.292, 5.768 (all weak)
RNA Only	0.758, 1.551, 3.415, 4.058, 7.574, 8.508 (all weak)
N11 + **A1** (no template)	0.753, 3.810, 3.958, 4.099 (strongest, diamine), 8.250 (v. weak)
N10 + **A1** (no template)	0.757, 1.854 (weak), 3.824, 4.520 (weak), 4.778 (weak), 5.385 (acridine), **5.753** (strongest), 6.480
N11 + **A1** + DNA	0.757, 2.047, 3.061, 4.133, 5.309, 5.781 (all weak)
N10 + **A1** + DNA	0.754, 1.829 (weak), 2.008, 2.573 (weak), 3.920, 4.048, 4.789 (weak), 5.375 (acridine), **5.756** (strongest), 6.477
N11 + **A1** + RNA	0.751, 1.551, 3.791 (dialdehyde), 4.017 (diamine), 4.126, 5.302, 5.762 (all weak)
N11 + **A1** + RNA	0.754, 1.552 (weak), 1.825 (weak), 4.074, 5.367 (acridine), **5.749** (strongest), 6.203, 6.469

Table 4.2. HPLC peaks observed in preliminary screening after 1 d. Previously observed retention times for starting materials in same gradient: Dialdehyde **A1** 3.796 min, Acridine-Aniline **N10** 5.312 min, Diamine **N11** 4.094 min. Note that the "Buffer Only" and "DNA/RNA Only" peaks are not reported in the subsequent reactions.

Preliminary Library Screen: Before the full library shown above was synthesized, a screen with **A1-A4**, **O1**, and **N1-N14** was performed to determine if the project was worth pursuing further. The conditions for this assay are as follows: 100 μM compounds (from 10 mM stock solutions), ±

10 μM $d(CTG)_{16}$ or $r(CUG)_{16}$ template in buffer of 2 mM Tris-HCl (pH 7), 2 mM KCl, 2 mM $MgCl_2$, and 2 mM $CaCl_2$. The compounds with or without template were incubated in buffer at 37 °C. Samples were removed after 4 h, 1 d, 3d, and 7 d, reduced with $NaBH_4$ (1 μL of 10 mM aqueous solution) and the template was degraded with 1 μL DNaseI or RNase H (1 μL water was added to the non-templated reactions.) The samples were subjected to MALDI-TOF-MS analysis in DHB matrix. The raw data was exported to an Excel file and an algorithm designed in R was used to search a list of potential masses (composed of each starting material as well as all dimeric and some trimeric products) for matches within the raw data (Figure 4.9). The only trimeric products that have not yet been included in the analysis are trimers composed of 3 different monomers. The output at each time point was recorded and is summarized in the heat maps shown in Figure 4.10.

Mass List Searched
Total Monomers = 34
Total Potential Dimers = 208
Total Potential Homotrimers = 286
Total Potential Heterotrimers = 1,579

Total list of masses analyzed = **2,173**
332 potential hits

```r
require(openxlsx)

# Load data
folder <- 'C:/Users/sbonson2/Box Sync/Documents/Sarah/Signals Notebook/Excel
setwd(folder)
raw <- read.xlsx('14023_massfinder_3.xlsx', sheet = 2) # Raw data from instr
mass <- read.xlsx('14023_massfinder_3.xlsx', sheet = 4) # Calculated product

# Set up to find peaks
rng <- 1 # peaks within +/- this value found
result <- data.frame(product = character(0), calcMass = numeric(0),
                     A = numeric(0), Intensity_A = numeric(0))

for (i in 1:length(mass$Mass)) {
  # Separate raw into run numbers
  rawA <- raw$`m/z`[raw$Run == 'A']
  intensAraw <- raw$Intens.[raw$Run == 'A']

  # Find peaks within range [m/z] of calculated mass
  valsA <- rawA[abs(rawA-mass$Mass[i]) <= rng]
  intensA <- intensAraw[abs(rawA-mass$Mass[i]) <= rng]


  # Add 0 if peak not returned
  valsA <- ifelse(length(valsA)==0, 0, valsA)
  intensA <- ifelse(length(intensA)==0, 0, intensA)


  # Build table
  app <- data.frame(product = mass$Product[i], calcMass = mass$Mass[i],
                    A = valsA, Intensity_A = intensA)

  result <- rbind(result,app)
}

#result[result == 0]  <- NA

write.table(result, file = '14023_3_peakfinder_intensities.csv',
            row.names = F, sep = ',', na = '')
```

Figure 4.9. Search algorithm for matching observed peaks with potential product peaks from the DCC screen.

Preliminary Nucleic Acid Templated Condensation Screen Procedures

Reactions were run in 0.5 mL Eppendorf tubes to a final volume of 50 μL in molecular biology grade water. A master solution of buffer (for final concentration of 2 mM each of Tris-HCl (pH 7), KCl, $MgCl_2$, $CaCl_2$) was made prior to each experiment and the proper aliquot of the solution was added to each Eppendorf tube. To reactions with template was added the proper amount of nucleic acid solution to achieve the desired final concentration. The tubes were annealed in buffer

before addition of compounds on a thermocycler at 95 °C for 5 min and cooled at a rate of 0.1 °C/s to 25 °C. Compounds were added after annealing as 10 mM stock solutions to achieve the desired final concentration of each compound. Compound solutions were incubated at 37 °C with freshly annealed template in buffer for 3 h. After 3 h, 2.5 µL of 10 mM $NaBH_4$ was added and the sample was incubated at 37 °C overnight. A 1 µL aliquot was removed and spotted for MALDI-MS with 1 µL DHB matrix (50 mg/mL in 70 water:30 MeCN, 0.1% TFA).

For MALDI analysis, a 10 µL aliquot was removed from the reactions and 1 µL 100 µM $NaCNBH_3$ or $NaBH_4$ (aqueous) was added. The aliquots were incubated at 37 °C overnight and 1 µL DNase I (NEB), 1 µL RNase H (NEB), or 1 µL molecular biology grade water was added to the DNA-templated, RNA-templated, and non-templated reactions, respectively. The tubes were

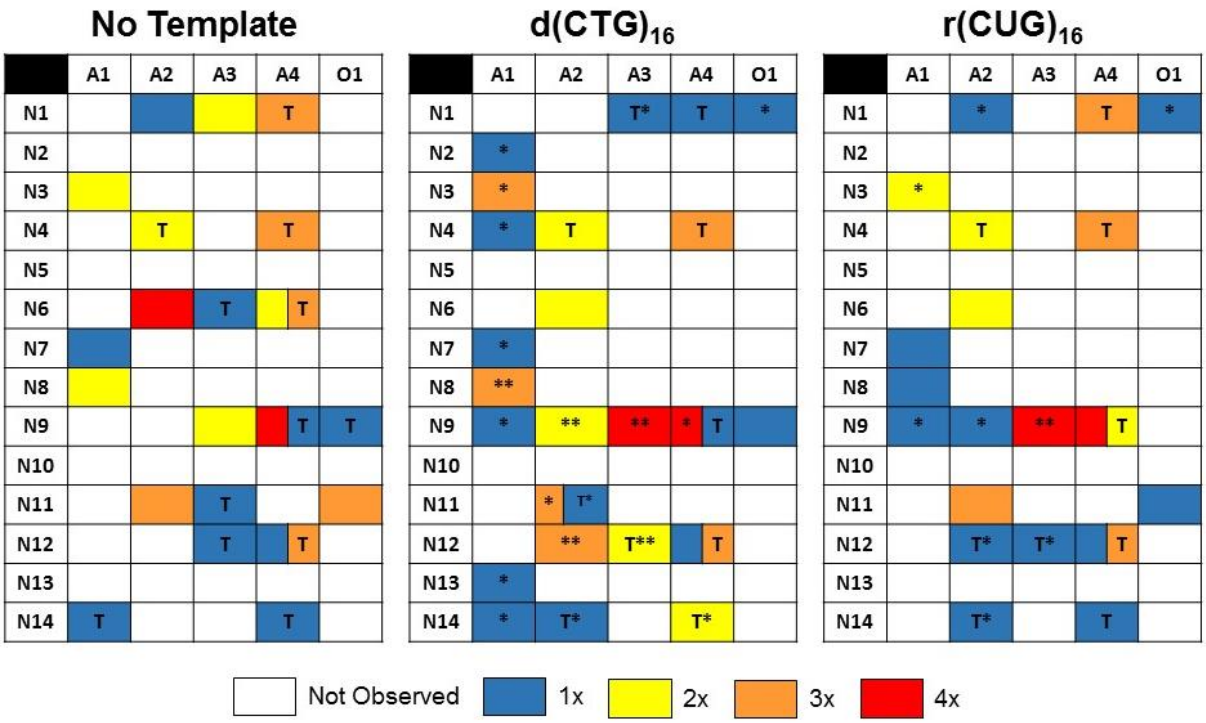


Figure 4.10. Initial DCC library screen heat maps. 10 µM $d(CTG)_{16}$ or $r(CUG)_{16}$ (H_2O for no template), 100 µM compounds, 2 mM Tris-HCl, 2 mM KCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$, + $NaBH_4$ after 4 h, 1 d, 3 d, 7d, 1 µL H_2O, DNaseI or RNaseH, analyzed by MALDI (DHB matrix) and screened for peaks matching product. *indicates an experiment in which the product peak was observed on template but not off template. T indicates trimer product observed. Boxes without a T correspond to dimeric products.

incubated at 37 °C for 20 min. MALDI-MS of each sample was measured (DHB matrix) by mixing 1 µL matrix with 1 µL sample on the MALDI plate. MALDIs were screened for new mass peaks.

For HPLC analysis, an aliquot (5-50 µL depending on compound concentration) was removed from the reactions and 1 µL 100 µM $NaCNBH_3$ or $NaBH_4$ (aqueous) was added. The aliquots were incubated at 37 °C overnight and 1 µL DNase I (NEB), 1 µL RNase H (NEB), or 1 µL molecular biology grade water was added to the DNA-templated, RNA-templated, and non-templated reactions, respectively, to degrade the template. Analytical HPLC was run with gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min. Injection volume 50 µL.

Master solutions: $MgCl_2$ 100 mM (aq), $CaCl_2$ 100 mM (aq), KCl 100 mM (aq), Tris HCl pH 7 100 mM (aq), $d(CTG)_{16}$ 2 mM (aq), $r(CUG)_{16}$ 2 mM (aq), compounds 10 mM (aq) except **N9** is 10 mM in DMSO. DMSO was added to some experiments (5% or 10%) as noted above.

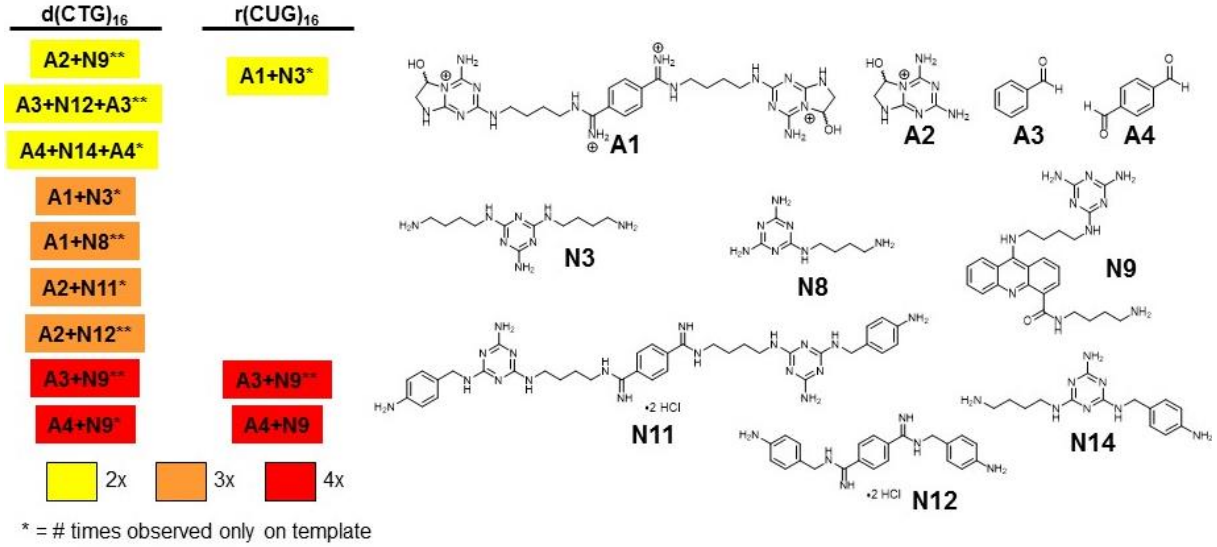


Figure 4.11. Initial DCC library screen hits. Color corresponds to the number of time points the hit was observed (out of 4 possible) and the number of stars after the combination indicates the number of time points the hit was observed only on the template and not off the template.

in vitro Transcription Inhibition: 2x2 DCC Screen

To test for a potential synergistic effect upon assembly of the DCC ligands used in the initial 2x2 HPLC and MALDI screens described above (**A1**, **N9**, and **N11**), the d(CTG·CAG)$_{90}$ pSP72 plasmid was used with both T7 and SP6 polymerases to study the ability to inhibit the formation of r(CUG)^{exp} and r(CAG)^{exp} with each monomer and a mixture of the two compounds. The formation of r(CAG)$_{90}$ from this plasmid can be performed using SP6 RNA polymerase, whereas the formation of r(CUG)$_{90}$ utilizes the T7 RNA polymerase and proceeds in the reverse direction. Of note, this screen was performed with the initial hit compounds from the 2x2 HPLC and MALDI screens and has not yet been performed with the hits from the library screen. Although no synergistic effect was observed for these compounds (Figure 4.5.5), other combinations could be tested.

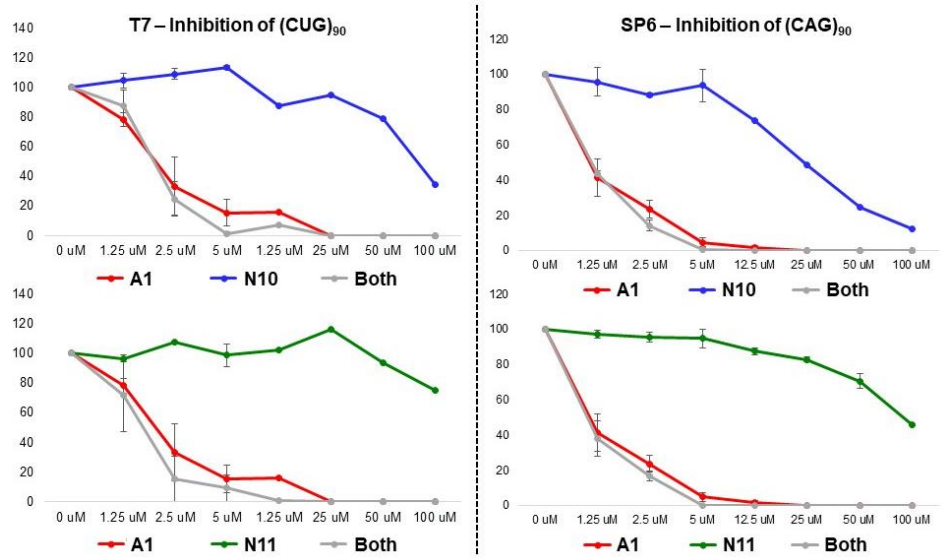


Figure 4.12. Bidirectional *in vitro* transcription inhibition study with compounds from initial hit from the 2x2 DCC screen. It appears that there is no synergistic effect between the acridine and bisamidinium ligands shown here on the formation of r(CUG)$_{90}$ or r(CAG)$_{90}$ from the d(CTG·CAG)$_{90}$ template.

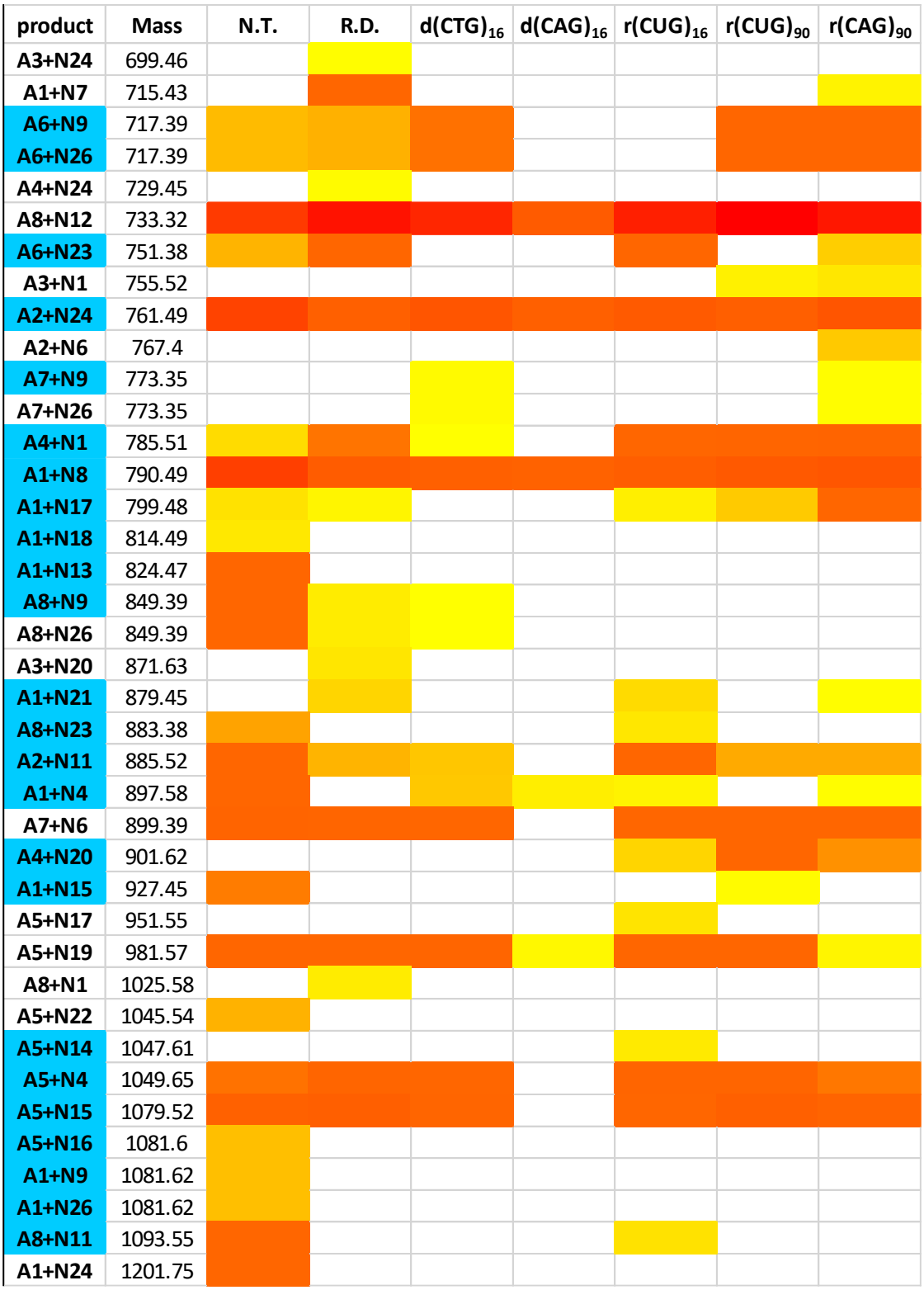


Figure 4.13. Full heat map including dimeric hits from MALDI-MS Screen. Raw MALDI-MS data was screened for new peaks consistent with dimers. The heat map represents the intensity as an average over 3 independent replicates. Coloring from red to yellow to white corresponds to normalized intensity, with red being the strongest intensity. Map is sorted by calculated mass from lowest to highest. Blue highlight indicates further study in single point assay.

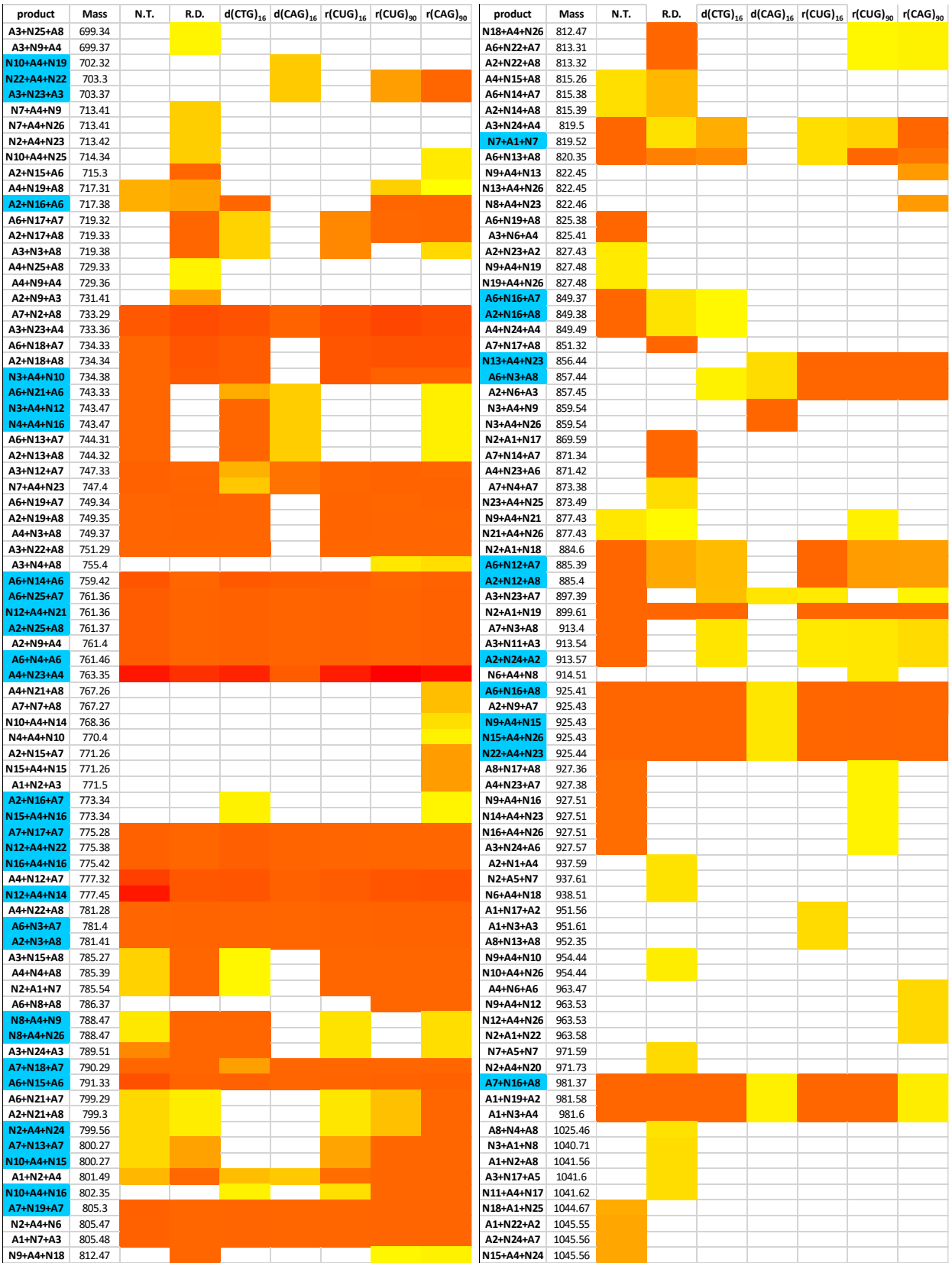
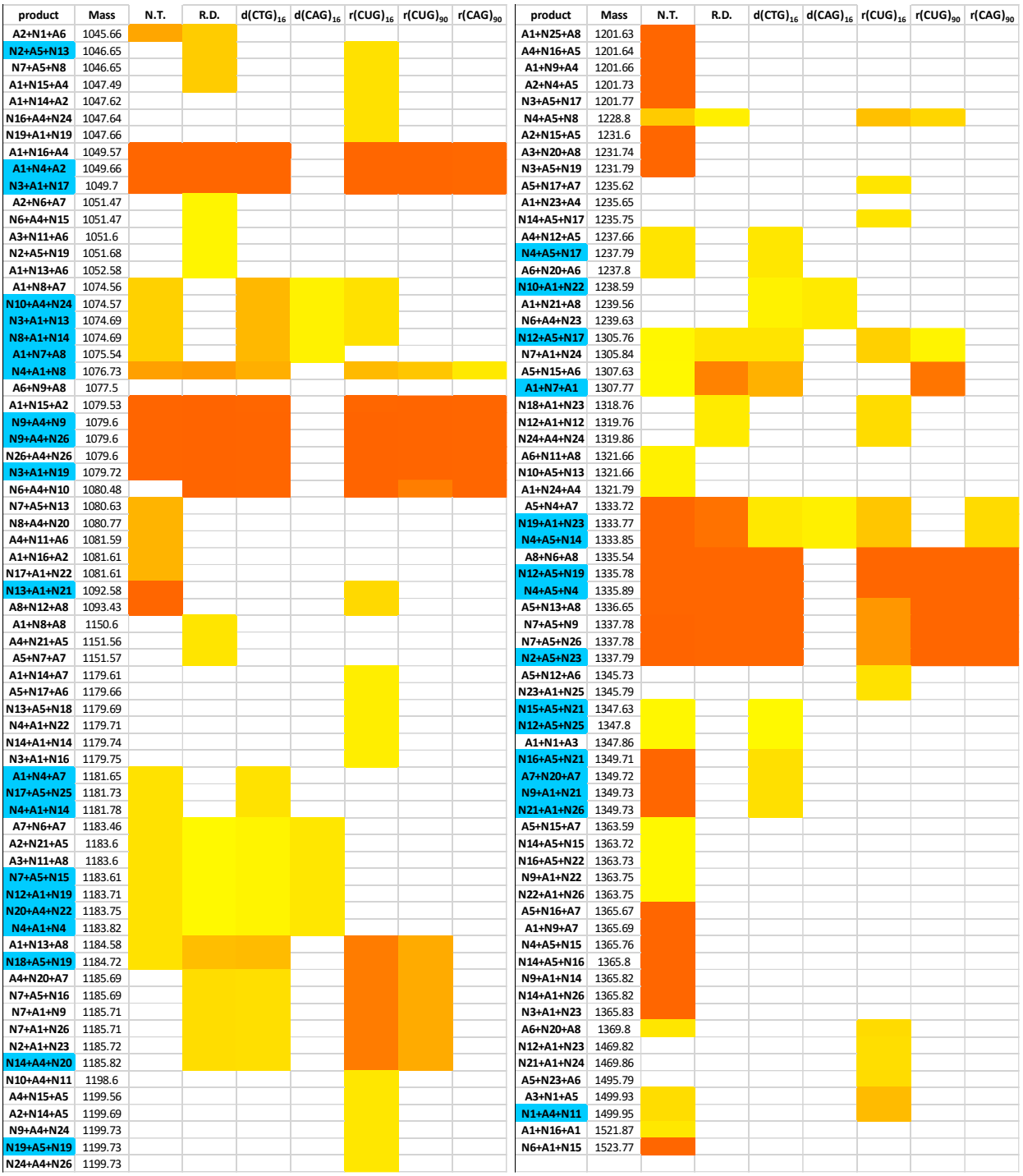


Figure 4.14. Full heat map including trimeric hits from MALDI-MS Screen. See full caption on next page.



product	Mass	N.T.	R.D.	d(CTG)$_{16}$	d(CAG)$_{16}$	r(CUG)$_{16}$	r(CUG)$_{90}$	r(CAG)$_{90}$
A2+N1+A6	1045.66							
N2+A5+N13	1046.65							
N7+A5+N8	1046.65							
A1+N15+A4	1047.49							
A1+N14+A2	1047.62							
N16+A4+N24	1047.64							
N19+A1+N19	1047.66							
A1+N16+A4	1049.57							
A1+N4+A2	1049.66							
N3+A1+N17	1049.7							
A2+N6+A7	1051.47							
N6+A4+N15	1051.47							
A3+N11+A6	1051.6							
N2+A5+N19	1051.68							
A1+N13+A6	1052.58							
A1+N8+A7	1074.56							
N10+A4+N24	1074.57							
N3+A1+N13	1074.69							
N8+A1+N14	1074.69							
A1+N7+A8	1075.54							
N4+A1+N8	1076.73							
A6+N9+A8	1077.5							
A1+N15+A2	1079.53							
N9+A4+N9	1079.6							
N9+A4+N26	1079.6							
N26+A4+N26	1079.6							
N3+A1+N19	1079.72							
N6+A4+N10	1080.48							
N7+A5+N13	1080.63							
N8+A4+N20	1080.77							
A4+N11+A6	1081.59							
A1+N16+A2	1081.61							
N17+A1+N22	1081.61							
N13+A1+N21	1092.58							
A8+N12+A8	1093.43							
A1+N8+A8	1150.6							
A4+N21+A5	1151.56							
A5+N7+A7	1151.57							
A1+N14+A7	1179.61							
A5+N17+A6	1179.66							
N13+A5+N18	1179.69							
N4+A1+N22	1179.71							
N14+A1+N14	1179.74							
N3+A1+N16	1179.75							
A1+N4+A7	1181.65							
N17+A5+N25	1181.73							
N4+A1+N14	1181.78							
A7+N6+A7	1183.46							
A2+N21+A5	1183.6							
A3+N11+A8	1183.6							
N7+A5+N15	1183.61							
N12+A1+N19	1183.71							
N20+A4+N22	1183.75							
N4+A1+N4	1183.82							
A1+N13+A8	1184.58							
N18+A5+N19	1184.72							
A4+N20+A7	1185.69							
N7+A5+N16	1185.69							
N7+A1+N9	1185.71							
N7+A1+N26	1185.71							
N2+A1+N23	1185.72							
N14+A4+N20	1185.82							
N10+A4+N11	1198.6							
A4+N15+A5	1199.56							
A2+N14+A5	1199.69							
N9+A4+N24	1199.73							
N19+A5+N19	1199.73							
N24+A4+N26	1199.73							

product	Mass	N.T.	R.D.	d(CTG)$_{16}$	d(CAG)$_{16}$	r(CUG)$_{16}$	r(CUG)$_{90}$	r(CAG)$_{90}$
A1+N25+A8	1201.63							
A4+N16+A5	1201.64							
A1+N9+A4	1201.66							
A2+N4+A5	1201.73							
N3+A5+N17	1201.77							
N4+A5+N8	1228.8							
A2+N15+A5	1231.6							
A3+N20+A8	1231.74							
N3+A5+N19	1231.79							
A5+N17+A7	1235.62							
A1+N23+A4	1235.65							
N14+A5+N17	1235.75							
A4+N12+A5	1237.66							
N4+A5+N17	1237.79							
A6+N20+A6	1237.8							
N10+A1+N22	1238.59							
A1+N21+A8	1239.56							
N6+A4+N23	1239.63							
N12+A5+N17	1305.76							
N7+A1+N24	1305.84							
A5+N15+A6	1307.63							
A1+N7+A1	1307.77							
N18+A1+N23	1318.76							
N12+A1+N12	1319.76							
N24+A4+N24	1319.86							
A6+N11+A8	1321.66							
N10+A5+N13	1321.66							
A1+N24+A4	1321.79							
A5+N4+A7	1333.72							
N19+A1+N23	1333.77							
N4+A5+N14	1333.85							
A8+N6+A8	1335.54							
N12+A5+N19	1335.78							
N4+A5+N4	1335.89							
A5+N13+A8	1336.65							
N7+A5+N9	1337.78							
N7+A5+N26	1337.78							
N2+A5+N23	1337.79							
A5+N12+A6	1345.73							
N23+A1+N25	1345.79							
N15+A5+N21	1347.63							
N12+A5+N25	1347.8							
A1+N1+A3	1347.86							
N16+A5+N21	1349.71							
A7+N20+A7	1349.72							
N9+A1+N21	1349.73							
N21+A1+N26	1349.73							
A5+N15+A7	1363.59							
N14+A5+N15	1363.72							
N16+A5+N22	1363.73							
N9+A1+N22	1363.75							
N22+A1+N26	1363.75							
A5+N16+A7	1365.67							
A1+N9+A7	1365.69							
N4+A5+N15	1365.76							
N14+A5+N16	1365.8							
N9+A1+N14	1365.82							
N14+A1+N26	1365.82							
N3+A1+N23	1365.83							
A6+N20+A8	1369.8							
N12+A1+N23	1469.82							
N21+A1+N24	1469.86							
A5+N23+A6	1495.79							
A3+N1+A5	1499.93							
N1+A4+N11	1499.95							
A1+N16+A1	1521.87							
N6+A1+N15	1523.77							

Figure 4.14 (cont). Full heat map including trimeric hits from MALDI-MS Screen. Raw MALDI-MS data was screened for new peaks consistent with homotrimers and heterotrimers. The heat map represents the intensity as an average over 3 independent replicates. Coloring from red to yellow to white corresponds to normalized intensity, with red being the strongest intensity. Map is sorted by calculated mass from lowest to highest. Blue highlight indicates further study in single point assay.

product	Mass	N.T.	R.D.	Structure	General Binders	Base Mimics	Non-Binders
A7+N17+A8	851.32			acridine - adenine - acridine	2	1	
N2+A1+N17	869.59			diaminobutane - melamine-bisamidinium-melamine - adenine	2	3	
A7+N14+A7	871.34			acridine - melamine - acridine	2	1	
A4+N23+A6	871.42			benzene - acridine-melamine - melamine	1	2	1
A2+N9+A3	731.41			melamine - acridine-melamine - benzene	1	2	1
A4+N15+A8	815.26			benzene - amiloride - acridine	1	1	1
A6+N14+A7	815.38			melamine - melamine - acridine	1	2	
A2+N14+A8	815.39			melamine - melamine - acridine	1	2	
A2+N1+A6	1045.66			melamine - melamine-bisamidinium-melamine - melamine	1	4	
N7+A4+N9	713.41			benzene - benzene - acridine-melamine	1	1	2
N7+A4+N26	713.41			benzene - benzene - acridine	1		2
N2+A4+N23	713.42			diaminobutane - benzene - acridine-melamine	1	1	2
N7+A5+N7	971.59			benzene - melamine-bisamidinium-melamine - benzene	1	2	2
N2+A4+N20	971.73			diaminobutane - benzene - melamine-bisamindinium-melamine	1	2	2
N3+A1+N8	1040.71			melamine - melamine-bisamidinium-melamine - melamine	1	4	
A1+N2+A8	1041.56			melamine-bisamidinium-melamine - diaminobutane - acridine	2	2	1
A3+N17+A5	1041.6			benzene - adenine - melamine-bisamidinium-melamine	1	3	1
N11+A4+N17	1041.62			melamine-bisamidinium-melamine - benzene - adenine	2	2	1
A3+N20	871.63			benzene - melamine-bisamidinium-melamine	1	2	1
A7+N4+A7	873.38			acridine - bisamidinium - acridine	3		
N23+A4+N25	873.49			acridine-melamine - benzene - bisamidinium	2	1	1
A2+N1+A4	937.59			melamine - melamine-bisamidinium-melamine - benzene	1	3	1
N2+A5+N7	937.61			diaminobutane - melamine-bisamidinium-melamine - acridine	2	2	1
N6+A4+N18	938.51			neomycin B - benzene - diaminopurine	1	1	1
A8+N4+A8	1025.46			acridine - bisamidinium - acridine	3		
A8+N1	1025.58			acridine - melamine-bisamidinium-melamine	2	2	
A1+N8+A8	1150.6			melamine-bisamidinium-melamine - melamine - acridine	2	3	
A4+N21+A5	1151.56			benzene - amiloride - melamine-bisamidinium-melamine	1	3	1
A5+N7+A7	1151.57			melamine-bisamidinium-melamine - benzene - acridine	2	2	1
N9+A4+N10	954.44			acridine-melamine - benzene - acridine	2	1	1
N10+A4+N26	954.44			acridine - benzene - acridine	2		1
A4+N25+A8	729.33			benzene - bisamidinium - acridine	2		1
A4+N9+A4	729.36			benzene - acridine-melamine - benzene	1	1	2
A4+N24	729.45			benzene - melamine-bisamidinium-melamine	1	2	1
A3+N25+A8	699.34			benzene - bisamidinium - acridine	2		1
A3+N9+A4	699.37			benzene - acridine-melamine - benzene	1	1	2
A3+N24	699.46			benzene - melamine-bisamidinium-melamine	1	2	1
A2+N6+A7	1051.47			melamine - neomycin B - acridine	2	1	
N6+A4+N15	1051.47			neomycin B - benzene - amiloride	1	1	1
A3+N11+A6	1051.6			benzene - melamine-bisamidinium-melamine - melamine	1	3	1
N2+A5+N19	1051.68			diaminobutane - melamine-bisamidinium-melamine - diaminopurine	1	3	1
A1+N13+A6	1052.58			melamine-bisamidinium-melamine - melamine - melamine	1	4	
A2+N15+A5	1231.6			melamine - amiloride - melamine-bisamidinium-melamine	1	4	
A3+N20+A8	1231.74			benzene - melamine-bisamidinium-melamine - acridine	2	2	1
N3+A5+N19	1231.79			melamine - melamine-bisamidinium-melamine - diaminopurine	1	4	
A1+N25+A8	1201.63			melamine-bisamidinium-melamine - bisamidinum - acridine	3	2	
A4+N16+A5	1201.64			benzene - melamine - melamine-bisamidinium-melamine	1	3	1
A1+N9+A4	1201.66			melamine-bisamidinium-melamine - acridine-melamine - benzene	2	3	1
A2+N4+A5	1201.73			melamine - bisamidinium - melamine-bisamidinium-melamine	2	3	
A1+N24	1201.75			melamine-bisamidinium-melamine - melamine-bisamidinium-melamine	2	4	
N3+A5+N17	1201.77			melamine - melamine-bisamidinium-melamine - adenine	1	4	
A5+N16+A7	1365.67			melamine-bisamidinium-melamine - melamine - acridine	2	3	
A1+N9+A7	1365.69			melamine-bisamidinium-melamine - acridine-melamine - acridine	3	3	
N4+A5+N15	1365.76			bisamindium - melamine-bisamidinium-melamine - amiloride	2	3	
N14+A5+N16	1365.8			melamine - melamine-bisamidinium-melamine - melamine	1	4	
N9+A1+N14	1365.82			acridine-melamine - melamine-bisamidinium-melamine - melamine	2	4	
N14+A1+N26	1365.82			melamine - melamine-bisamidinium-melamine - acridine	2	3	
N3+A1+N23	1365.83			melamine - melamine-bisamidinium-melamine - amiloride	1	4	
N6+A1+N15	1523.77			neomycin B - melamine-bisamidinium-melamine - amiloride	2	3	
A1+N13	824.47			melamine-bisamidinium-melamine - melamine	1	3	
A6+N19+A8	825.38			melamine - diaminopurine - melamine		3	
A3+N6+A4	825.41			benzene - neomycin B - benzene	1		2
A5+N22	1045.54			melamine-bisamidinium-melamine - amiloride	1	3	
A1+N22+A2	1045.55			melamine-bisamidinium-melamine - amiloride - acridine	2	3	
A2+N24+A7	1045.56			melamine - melamine-bisamidinium-melamine - acridine	2	3	
N15+A4+N24	1045.56			amiloride - benzene - melamine-bisamidinium-melamine	1	3	1
N18+A1+N25	1044.67			diaminopurine - melamine-bisamidinium-melamine - bisamidinium	2	3	
N7+A5+N13	1080.63			benzene - melamine-bisamidinium-melamine - melamine	1	3	1
N8+A4+N20	1080.77			melamine - benzene - melamine-bisamidinium-melamine	1	3	1
A4+N11+A6	1081.59			benzene - melamine-bisamidinium-melamine - melamine	1	3	1
A5+N16	1081.6			melamine-bisamidinium-melamine - melamine	1	3	
A1+N16+A2	1081.61			melamine-bisamidinium-melamine - melamine - amiloride	1	4	
N17+A1+N22	1081.61			adenine - melamine-bisamidinium-melamine - amiloride	2	3	
A1+N9	1081.62			melamine-bisamidinium-melamine - acridine-melamine	2	3	
A1+N26	1081.62			melamine-bisamidinium-melamine - acridine	2	2	
A1+N18	814.49			melamine-bisamidinium-melamine - diaminopurine	1	3	
A1+N16+A1	1521.87			melamine-bisamidinium-melamine - melamine - melamine-bisamidinium-melamine	2	5	
A2+N23+A2	827.43			melamine - acridine-melamine - melamine	1	3	
N9+A4+N19	827.48			acridine-melamine - benzene - diaminopurine	1	2	1
N19+A4+N26	827.48			diaminopurine - benzene - acridine	1	1	1
A6+N11+A8	1321.66			melamine - melamine-bisamidinium-melamine - acridine	2	3	
N10+A5+N13	1321.66			acridine - melamine-bisamidinium-melamine - melamine	2	3	
A1+N24+A4	1321.79			melamine-bisamidinium-melamine - melamine-bisamidinium-melamine - benzene	2	4	1
A5+N15+A7	1363.59			melamine-bisamidinium-melamine - amiloride - acridine	2	3	
N14+A5+N15	1363.72			melamine - melamine-bisamidinium-melamine - amiloride	1	4	
N16+A5+N22	1363.73			melamine - melamine-bisamidinium-melamine - amiloride	1	4	
N9+A1+N22	1363.75			acridine-melamine - melamine-bisamidinium-melamine - amiloride	2	4	
N22+A1+N26	1363.75			amiloride - melamine-bisamidinium-melamine - acridine	2	3	

Table 4.3. Identity of compounds that assembled in the buffer and/or on the random duplex template.

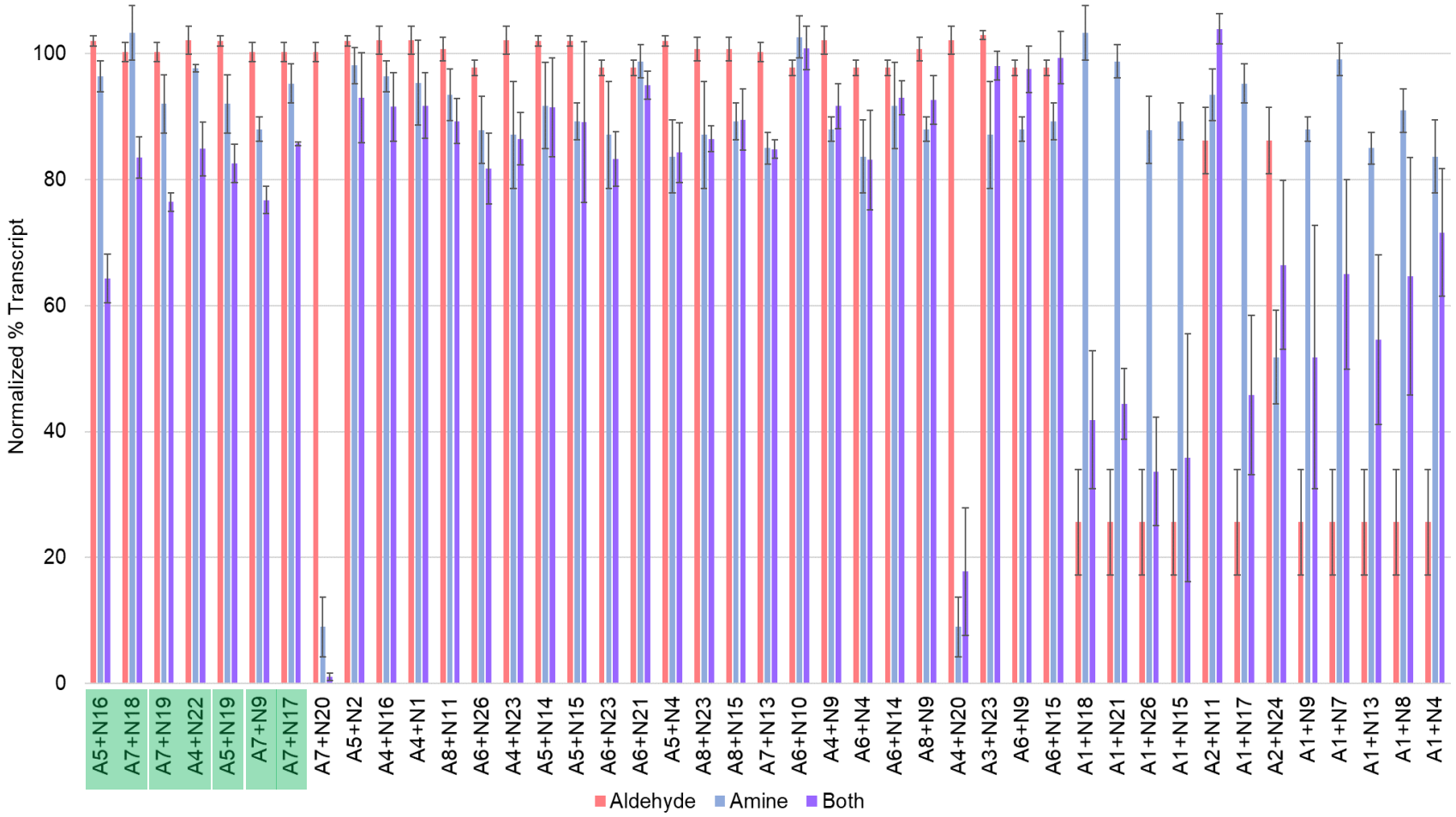


Figure 4.15. Full Results for Single Point *in vitro* Transcription Inhibition with Dimers and Homotrimers. Hit combinations (10 μM each) were incubated with a pSP72 plasmid containing d(CTG CAG)$_{90}$. The output r(CUG)$_{90}$ transcript was quantified to determine the normalized % transcript relative to the monomers. Aldehyde monomer is pink, amine monomer is blue, and the mixture is shown in purple. Data is sorted by difference between the monomers and the mixture (largest difference to smallest) and combinations selected for further study are highlighted in

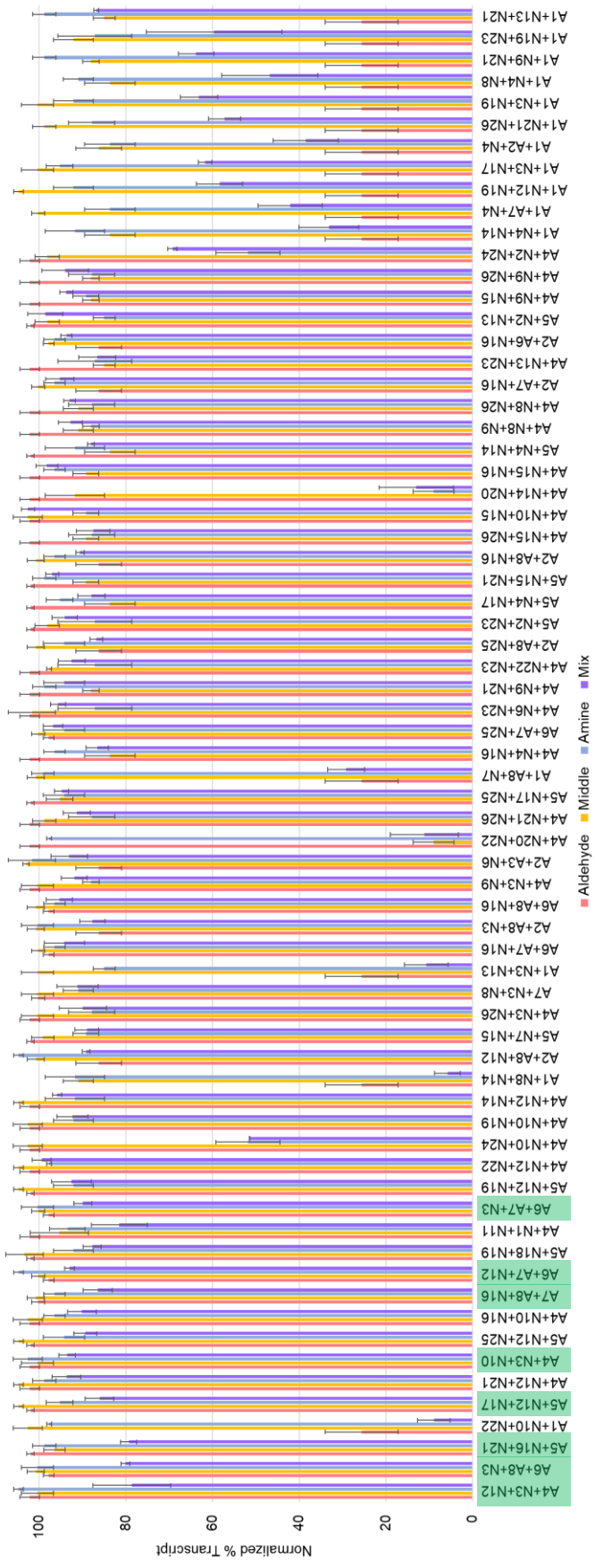


Figure 4.16. Full Results for Single Point *in vitro* Transcription Inhibition with Heterotrimers. Hit combinations (10 μM each) were incubated with a pSP72 plasmid containing d(CTG CAG)$_{90}$. The output r(CUG)$_{90}$ transcript was quantified to determine the normalized % transcript relative to the monomers. Aldehyde monomer is pink, amine monomer is blue, and "middle" monomer (could be aldehyde or amine) is yellow. The data from the mixture is shown in purple. Data is sorted by difference between the monomers and the mixture (largest difference to smallest) and combinations selected for further study are highlighted in green.

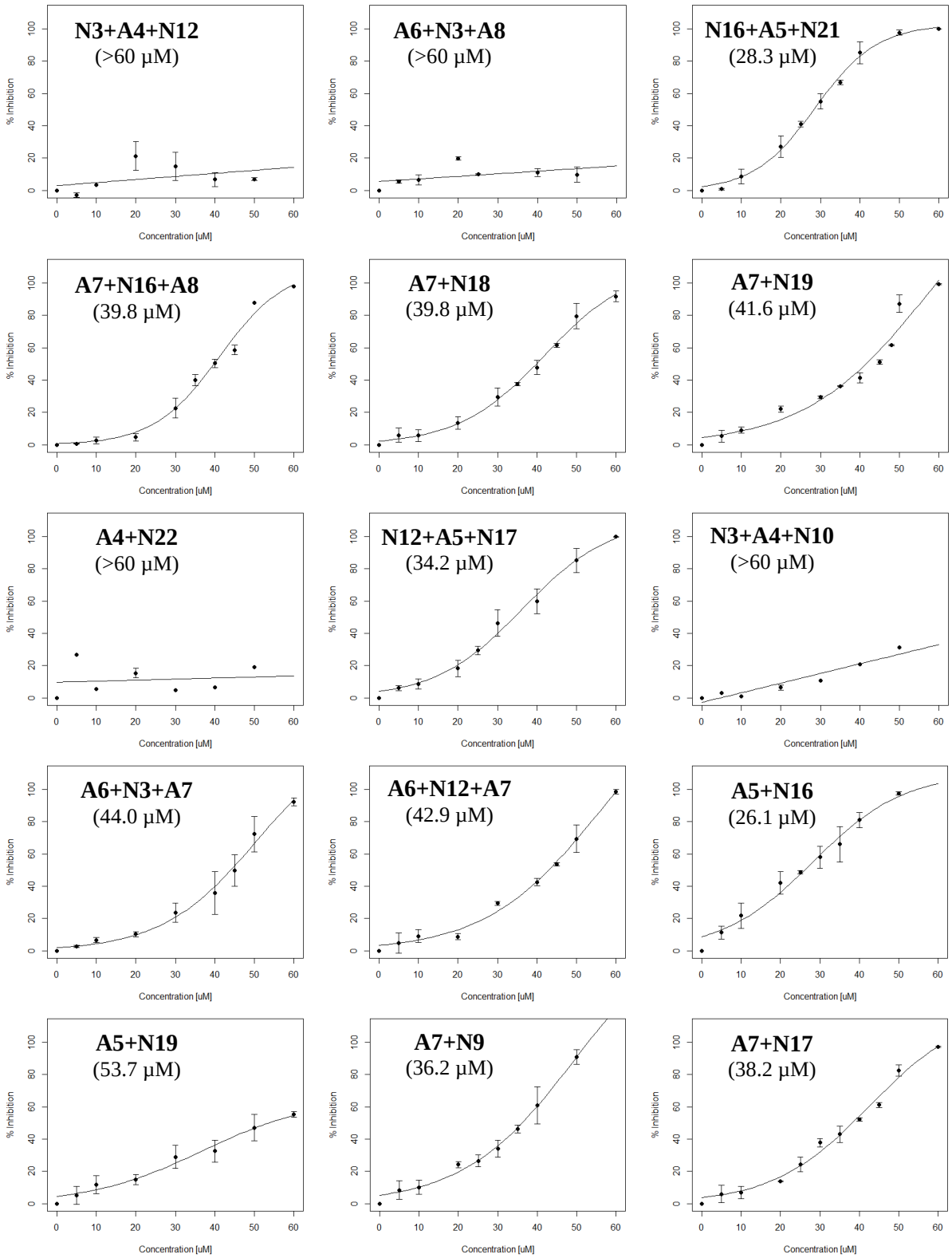


Figure 4.17. IC_{50} curves for hit combinations obtained from MALDI-MS screen and single-point *in vitro* transcription inhibition assay. Calculated IC_{50} values are shown in parentheses under the combination.

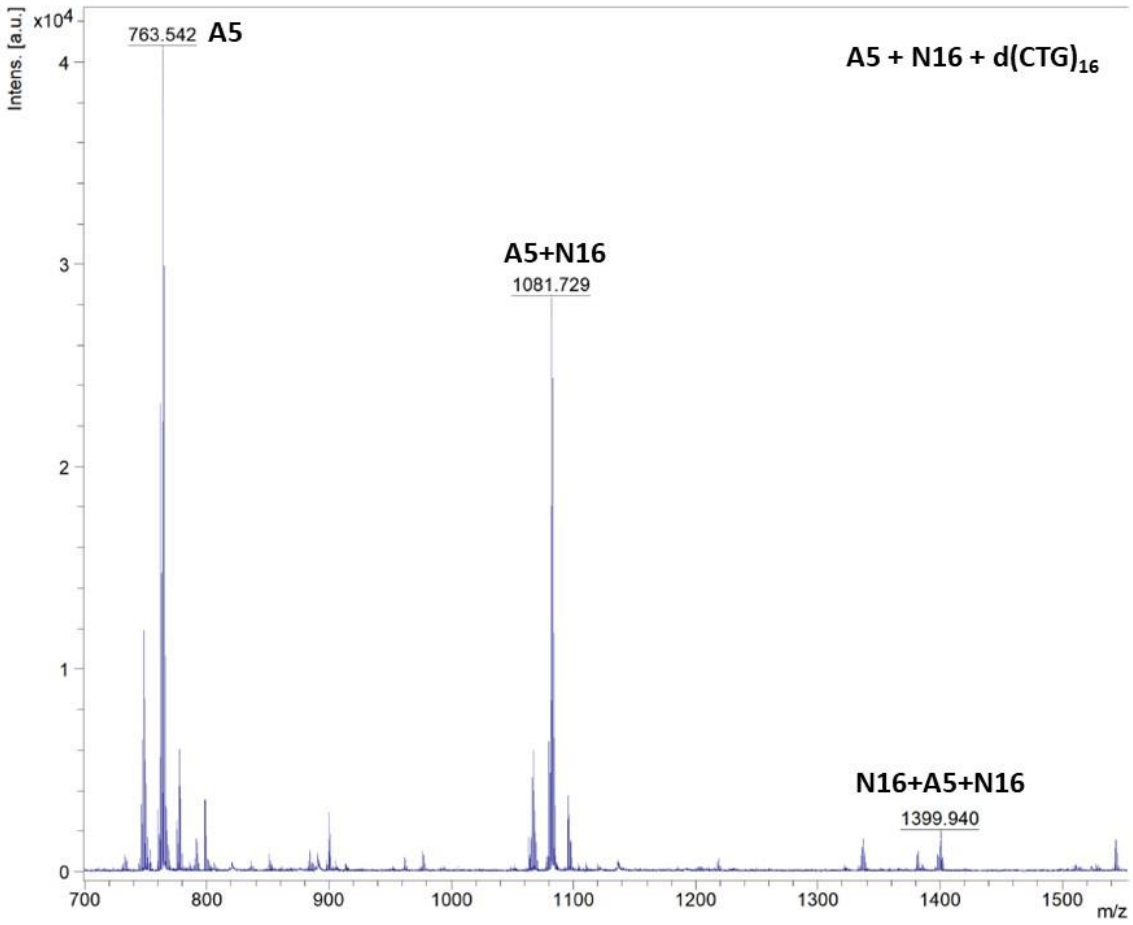


Figure 4.18. MALDI-MS for incubation of combination A5+N16 with $d(CTG)_{16}$ template. Compounds (100 µM each) and $d(CTG)_{16}$ (10 µM). Aqueous buffer: 2 mM each of KCl, $MgCl_2$, $CaCl_2$, and Tris-HCl, pH 7. The solutions were incubated at 37 °C for 24 h before reduction with 2.5 µL of 100 mM sodium borohydride. Following 6 h incubation at 37 °C (for reduction step), a 1 µL aliquot was removed for MALDI-MS (run in 1 µL DHB matrix).

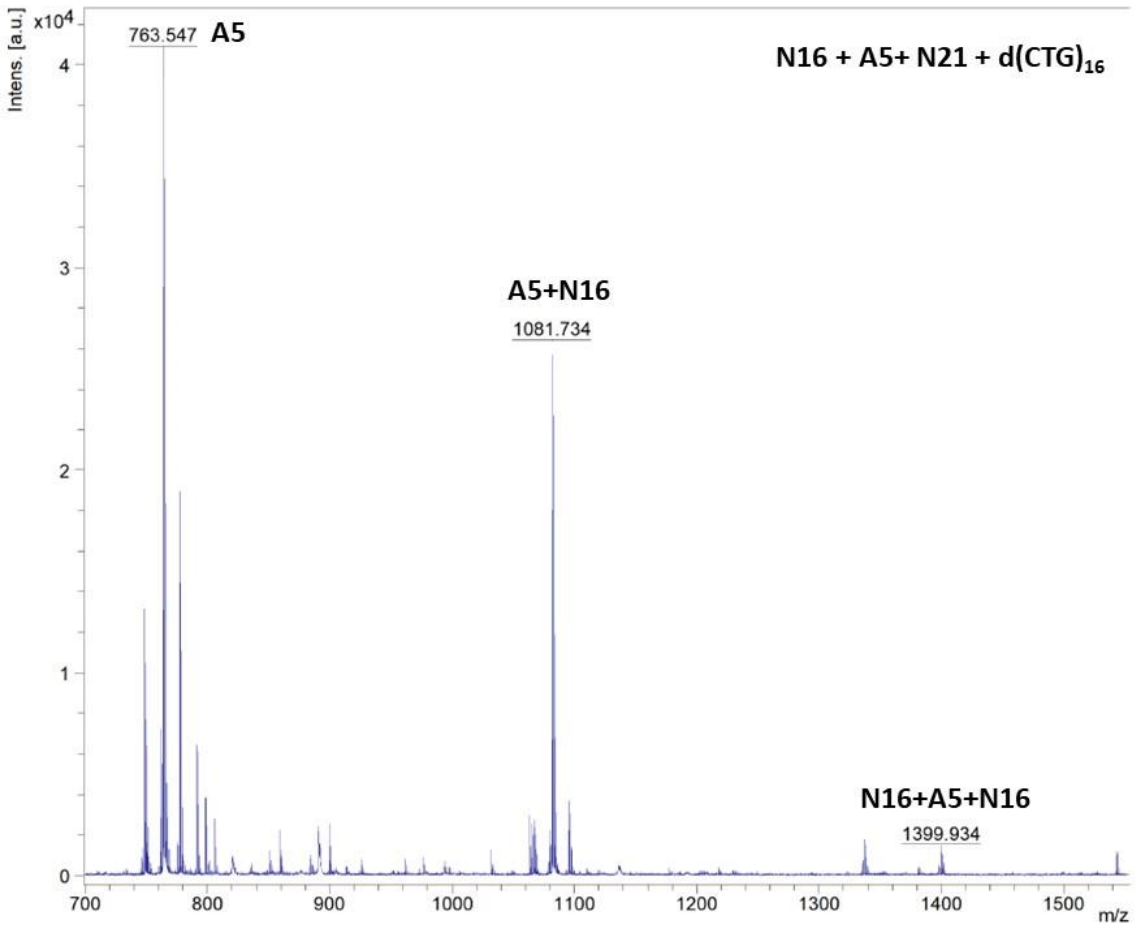


Figure 4.19. MALDI-MS for incubation of combination A5+N16 with $d(CTG)_{16}$ template. Compounds (100 µM each) and $d(CTG)_{16}$ (10 µM). Aqueous buffer: 2 mM each of KCl, $MgCl_2$, $CaCl_2$, and Tris-HCl, pH 7. The solutions were incubated at 37 °C for 24 h before reduction with 2.5 µL of 100 mM sodium borohydride. Following 6 h incubation at 37 °C (for reduction step), a 1 µL aliquot was removed for MALDI-MS (run in 1 µL DHB matrix).

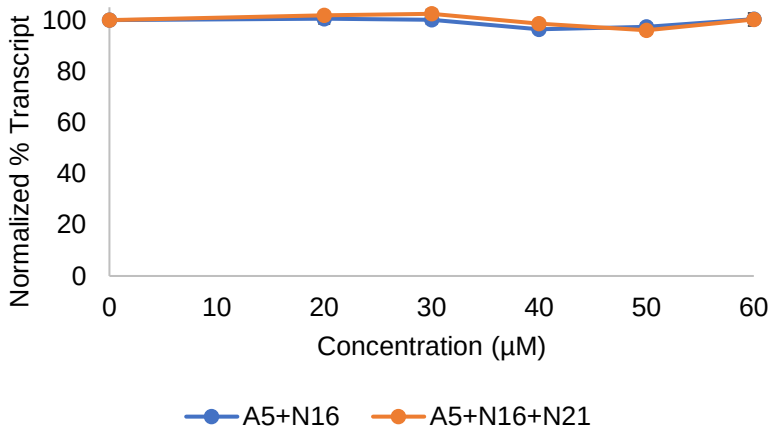


Figure 4.20. *in vitro* transcription of random duplex template with hit combinations A5+N16 and N16+A5+N21. No inhibition was observed. Results are reported as the average of three independent experiments and error is reported as standard error of the mean.f1

4.6. REFERENCES

(1) Mirkin, S. M. Expandable DNA Repeats and Human Disease. *Nature* **2007**, *447* (7147), 932–940.

(2) Khristich, A. N.; Mirkin, S. M. On the Wrong DNA Track: Molecular Mechanisms of Repeat-Mediated Genome Instability. *J. Biol. Chem.* **2020**, *295* (13), 4134–4170.

(3) Lin, X.; Miller, J. W.; Mankodi, A.; Kanadia, R. N.; Yuan, Y.; Moxley, R. T.; Swanson, M. S.; Thornton, C. A. Failure of MBNL1-Dependent Post-Natal Splicing Transitions in Myotonic Dystrophy. *Hum. Mol. Genet.* **2006**, *15* (13), 2087–2097.

(4) Charizanis, K. et al. Muscleblind-like 2-Mediated Alternative Splicing in the Developing Brain and Dysregulation in Myotonic Dystrophy. *Neuron* **2012**, *75* (3), 437–450.

(5) Miller, J. W.; Urbinati, C. R.; Teng-umnuay, P.; Stenberg, M. G.; Byrne, B. J.; Thornton, C. A.; Swanson, M. S. Recruitment of Human Muscleblind Proteins to (CUG)n Expansions Associated with Myotonic Dystrophy. *EMBO J.* **2000**, *19* (17), 4439–4448.

(6) Savkur, R. S.; Philips, A. V.; Cooper, T. A. Aberrant Regulation of Insulin Receptor Alternative Splicing Is Associated with Insulin Resistance in Myotonic Dystrophy. *Nat. Genet.* **2001**, *29* (1), 40–47.

(7) Philips, A. V; Timchenko, L. T.; Cooper, T. A. Disruption of Splicing Regulated by a CUG-Binding Protein in Myotonic Dystrophy. *Science* **1998**, *280*, 737–741.

(8) Konieczny, P.; Stepniak-Konieczna, E.; Sobczak, K. MBNL Proteins and Their Target RNAs, Interaction and Splicing Regulation. *Nucleic Acids Res.* **2014**, *42* (17), 10873–10887.

(9) Simmons, D. A.; Belichenko, N. P.; Yang, T.; Condon, C.; Monbureau, M.; Shamloo, M.; Jing, D.; Massa, S. M.; Longo, F. M. A Small Molecule TrkB Ligand Reduces Motor Impairment and Neuropathology in R6/2 and BACHD Mouse Models of Huntington's Disease. *J. Neurosci.* **2013**, *33* (48), 18712–18727.

(10) Budworth, H.; McMurray, C. T. Bidirectional Transcription of Trinucleotide Repeats: Roles for Excision Repair. *DNA Repair (Amst).* **2013**, *12* (8), 672–684.

(11) Gao, Z.; Cooper, T. A. Antisense Oligonucleotides: Rising Stars in Eliminating RNA Toxicity in Myotonic Dystrophy. *Hum. Gene Ther.* **2013**, *24* (5), 499–507.

(12) Rinaldi, C.; A Wood, M. J. Antisense Oligonucleotides: The next Frontier for Treatment of Neurological Disorders. *Nat. Publ. Gr.* **2017**.

(13) Provenzano, C.; Cappella, M.; Valaperta, R.; Cardani, R.; Meola, G.; Martelli, F.; Cardinali, B.; Falcone, G. CRISPR/Cas9-Mediated Deletion of CTG Expansions Recovers Normal Phenotype in Myogenic Cells Derived from Myotonic Dystrophy 1 Patients. *Mol. Ther. - Nucleic Acids* **2017**, *9*, 337–348.

(14) Jenquin, J. R.; Coonrod, L. A.; Silverglate, Q. A.; Pellitier, N. A.; Hale, M. A.; Xia, G.; Nakamori, M.; Berglund, J. A. Furamidine Rescues Myotonic Dystrophy Type I Associated Mis-Splicing through Multiple Mechanisms. *ACS Chem. Biol.* **2018**, *13* (9), 2708–2718.

(15) Hsieh, W. C.; Bahal, R.; Thadke, S. A.; Bhatt, K.; Sobczak, K.; Thornton, C.; Ly, D. H. Design of a "Mini" Nucleic Acid Probe for Cooperative Binding of an RNA-Repeated Transcript Associated with Myotonic Dystrophy Type 1. *Biochemistry* **2018**, *57* (6), 907–911.

(16) Konieczny, P.; Selma-Soriano, E.; Rapisarda, A. S.; Fernandez-Costa, J. M.; Perez-Alonso, M.; Artero, R. Myotonic Dystrophy: Candidate Small Molecule Therapeutics. *Drug Discov. Today* **2017**, *22* (11), 1740–1748.

(17) Hagihara, M.; Nakatani, K. Inhibition of DNA Replication by a d(CAG) Repeat Binding Ligand. *Nucleic Acids Symp. Ser.* **2006**, *50*, 147–148.

(18) Nakatani, K.; Hagihara, S.; Goto, Y.; Kobori, A.; Hagihara, M.; Hayashi, G.; Kyo, M.; Nomura, M.; Mishima, M.; Kojima, C. Small-Molecule Ligand Induces Nucleotide Flipping in (CAG)n Trinucleotide Repeats. *Nat. Chem. Biol.* **2005**, *1* (1), 39–43.

(19) Nakamori, M. et al. A Slipped-CAG DNA-Binding Small Molecule Induces Trinucleotide-Repeat Contractions in Vivo. *Nat. Genet.* **2020**, *52* (2), 146–159.

(20) Rideout, D. Self-Assembling Cytotoxins. *Science* **1986**, *233* (4763), 561–563.

(21) Rideout, D.; Calogeropoulou, T.; Jaworski, J.; McCarthy, M. Synergism through Direct Covalent Bonding between Agents: A Strategy for Rational Design of Chemotherapeutic Combinations. *Biopolymers* **1990**, *29* (1), 247–262.

(22) Hagler, L. D.; Krueger, S. B.; Luu, L. M.; Lanzendorf, A. N.; Mitchell, N. L.; Vergara, J. I.; Curet, L. D.; Zimmerman, S. C. Versatile Target-Guided Screen for Discovering Bidirectional Transcription Inhibitors of a Trinucleotide Repeat Disease. *ACS Med. Chem. Lett.* **2021**, *12* (6), 935–940.

(23) Hagler, L. D.; Luu, L. M.; Tonelli, M.; Lee, J.; Hayes, S. M.; Bonson, S. E.; Vergara, J. I.; Butcher, S. E.; Zimmerman, S. C. Expanded DNA and RNA Trinucleotide Repeats in Myotonic Dystrophy Type 1 Select Their Own Multitarget, Sequence-Selective Inhibitors.

Biochemistry **2020**, *59* (37), 3463–3472.

(24) Gareiss, P. C.; Sobczak, K.; Mcnaughton, B. R.; Palde, P. B.; Thornton, C. A.; Miller, B. L. Dynamic Combinatorial Selection of Molecules Capable of Inhibiting the (CUG) Repeat RNA-MBNL1 Interaction In Vitro: Discovery of Lead Compounds Targeting Myotonic Dystrophy (DM1). *J. Am. Chem. Soc.* **2008**, *130*, 16254–16261.

(25) Umuhire Juru, A.; Cai, Z.; Jan, A.; Hargrove, A. E. Template-Guided Selection of RNA Ligands Using Imine-Based Dynamic Combinatorial Chemistry. *Chem. Commun.* **2020**, *56* (24), 3555–3558.

(26) Galindo, M. A.; Amantia, D.; Martinez, A. M.; Clegg, W.; Harrington, R. W.; Martinez, V. M.; Houlton, A. Probing Metal-Ion Purine Interactions at DNA Minor-Groove Sites. *Inorg. Chem.* **2009**, *48* (21), 10295–10303.

(27) Arambula, J. F.; Ramisetty, S. R.; Baranger, A. M.; Zimmerman, S. C. A Simple Ligand That Selectively Targets CUG Trinucleotide Repeats and Inhibits MBNL Protein Binding. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (38), 16068–16073.

(28) Wong, C. H.; Nguyen, L.; Peh, J.; Luu, L. M.; Sanchez, J. S.; Richardson, S. L.; Tuccinardi, T.; Tsoi, H.; Chan, W. Y.; Chan, H. Y. E.; Baranger, A. M.; Hergenrother, P. J.; Zimmerman, S. C. Targeting Toxic RNAs That Cause Myotonic Dystrophy Type 1 (DM1) with a Bisamidinium Inhibitor. *J. Am. Chem. Soc.* **2014**, *136* (17), 6355–6361.

(29) Nguyen, L.; Luu, L. M.; Peng, S.; Serrano, J. F.; Chan, H. Y. E.; Zimmerman, S. C. Rationally Designed Small Molecules That Target Both the DNA and RNA Causing Myotonic Dystrophy Type 1. *J. Am. Chem. Soc.* **2015**, *137* (44), 14180–14189.

(30) Constant, J. F.; Laugaa, P.; Roques, B. P.; Lhomme, J. Heterodimeric Molecules Including Nucleic Acid Bases and 9-Aminoacridine. Spectroscopic Studies, Conformations, and Interactions with DNA. *Biochemistry* **1988**, *27*, 3997–4003.

(31) Michael, K.; Tor, Y. Designing Novel RNA Binders. *Chem. - A Eur. J.* **1998**, *4* (11), 2091–2098.

(32) Arya, D. P.; Micovic, L.; Charles, I.; Coffee, R. L.; Willis, B.; Xue, L. Neomycin Binding to Watson-Hoogsteen (W-H) DNA Triplex Groove: A Model. *J. Am. Chem. Soc.* **2003**, *125* (13), 3733–3744.

(33) Umuhire Juru, A.; Cai, Z.; Jan, A.; Hargrove, A. E. Template-Guided Selection of RNA Ligands Using Imine-Based Dynamic Combinatorial Chemistry. *Chem. Commun.* **2020**, *56* (24), 3555–3558.

APPENDIX A: MATERIALS, METHODS, AND SUPPLEMENTAL FIGURES

A.1 GENERAL METHODS

A.1.1. General Synthetic Methods

All reagents were purchased from Sigma-Aldrich, Fisch-Scientific, Oakwood Chemical, Acros Organics, Cambridge Chemical Technologies, Chem-Impex International, AK Scientific, or TCI America and used without further purification unless otherwise noted. Solvents were obtained from the Solvent Purification System and stored over activated 4Å molecular sieves unless otherwise noted. Ammonium hydroxide solution is aqueous, 28% by weight, referred to as "NH_4OH" in this work. Reactions were stirred with a magnetic stir bar and reactions run at elevated temperature were heated in an oil bath unless otherwise noted. Molecular-Biology grade water was used in all biological experiments including isothermal titration calorimetry and cytotoxicity assays. DNA templates were purchased from Integrated DNA Technologies unless otherwise noted. Mass spectra were obtained by the Mass Spectrometry Laboratory, School of Chemical Sciences, University of Illinois with ESI on a Waters Micromass Q-Tof spectrometer and field desorption (FD) on a Waters 70-VSE spectrometer. MALDI experiments were performed using Bruker UltrafleXtreme or Autoflex Speed in the University of Illinois Mass Spectrometry Laboratory. DHB matrix was prepared as 5 mg/mL 2,5-dihydroxybenzoic acid in 70 H_2O:30 MeCN, 0.1% TFA. Nuclear Magnetic Resonance Spectra were recorded on Varian 500 MHz, Carver-Bruker 500 MHz CryoProbe, or Bruker Prodigy (600 MHz) instruments in the NMR Laboratory, School of Chemical Sciences, University of Illinois at 22±3 °C unless otherwise noted. Spectra were processed using MestReNova (v12.0). Chemical shifts are reported in parts per million (ppm) and coupling constants (J) are reported in Hertz (Hz). ^{1}H NMR chemical shifts were

referenced to residual solvent peaks at 7.26 ppm for chloroform-*d* ($CDCl_3$), 2.50 ppm in DMSO-d_6, or 4.79 in d_2O. ^{13}C NMR shifts were referenced to the solvent peak at 77.16 ppm for $CDCl_3$ and 40.45 ppm for DMSO-d_6. Analytical thin-layer chromatography (TLC) was performed on 0.2 mm silica 60 coated on glass plates with F254 indicator. Flash column chromatography was performed on 40-63 μm silica gel (SiO_2) unless otherwise noted. Solvent mixtures used in chromatography are reported as a volume ratio (v:v). Analytical HPLC was performed on Shimadzu Prominence System, C-18 column, length 5 cm, diameter 4.6 mm, particle size 3 μm, flow rate 1,500 μL/min, detector wavelength 254 nm. Preparative HPLC was performed on Agilent Technologies 1260 Infinity II, C-18 250x50.0 mm column.

A.1.2. General Procedures: HeLa Cell Growth and Maintenance

HeLa cells (from cryogenic storage, 1 mL media/10% DMSO) were grown at 37 °C in T75 flasks in Dulbecco's Modified Eagle Medium media prepared in the Cell Media Facility, University of Illinois, with 10% Fetal Bovine Serum added. Frozen cells were thawed and transferred to a T75 flask. Cell media was changed after 1 d to remove DMSO. Otherwise, cell media was changed at least every 2 d. When cells reached 100% confluence, the cells were washed with 1X PBS (37 °C) and about 10 mL 37 °C trypsin (0.05% (1x) with EDTA) was added. After incubation at 37 °C for about 2 min, the sides of the flask were gently tapped to detach cells and contents of the T75 was transferred to a conical tube containing 3 volumes of cell media (relative to trypsin). The cells were counted, centrifuged at 500 G for 5 min, and the supernatant was pipetted off. The cells were diluted to 50,000 cells/mL. Harvested cells were used immediately in assays or transferred to a new T75 flask for further growth.

Freezing HeLa Cells. Some cells were frozen to thaw and use later. Cells were grown in a T175 culture flask in DMEM until 100% confluent. The media was aspirated, about 10 mL 37 °C trypsin (0.05% (1x) with EDTA) was added to the flask, and the cells were incubated at 37 °C for about 2 min. The sides of the flask were tapped to detach cells. About 20 mL DMEM was added to deactivate the trypsin. The cells were vortexed and to foam was aspirated. The cells were centrifuged at 500 G for 5 min, the media/trypsin solution was aspirated, and the cells were resuspended in 9 mL DMEM + 1 mL DMSO. The cell suspension was carefully transferred to cryotubes (1 mL each), sealed with parafilm, stored at -80 °C overnight, and transferred to liquid nitrogen cryogenic storage for long-term preservation.

A.1.3. General Procedure: Sulforhodamine B Cytotoxicity Assay[1]

HeLa cells (from cryogenic storage) were grown at 37 °C in T75 flasks in Dulbecco's Modified Eagle Medium media prepared in the Cell Media Facility, University of Illinois, with 10% Fetal Bovine Serum added. Cell media was changed at least every 2 d. When cells reached 100% confluence, the cells were washed with PBS (37 °C) and about 10 mL 37 °C trypsin (0.05% (1x) with EDTA) was added. After incubation at 37 °C for exactly 2 min, the contents of the T75 was transferred to a conical tube containing 3 volumes of cell media (relative to trypsin). The cells were counted, centrifuged at 500 G for 5 min, and the supernatant was pipetted off. The cells were diluted to 50,000 cells/mL. To split cells, 5 mL of cells (concentration 50,000 cells/mL) and 15 mL media was added to a new T75 flask. Harvested cells were used immediately in cytotoxicity assays. A 96-well plate was used according to the diagram below. The outer edge wells (white) were filled with 200 µL PBS. The dead controls (blacks) were filled with 200 µL cell media. The live controls (green) were filled with 100 µL cell media and 100 µL cells. The gradient-colored

blue boxes signify treatment with compound, with [1] being the highest, and [10] being the lowest. The compound master solutions were made from a serial dilution, and each treatment well was filled with 2 µL compound solution, 98 µL cell media, and 100 µL cells. The plates were incubated at 37 °C for 72 h, checking each day to monitor progress. After 72 h, 100 µL 10% w/v aqueous trichloroacetic acid at 4 °C was added and the plates were incubated for 1 h at 4 °C. The plates were washed with DI water, dried overnight, and 100 µL 0.05% w/v aqueous sulforhodamine B was added to each well. The plates were placed on the shaker for 15 min, rinsed quickly with 1% v/v aqueous acetic acid and then DI water. After the plates were dry, 200 µL of 10 mM tris base (pH 10.5) was added to each well. The optical density was measured at 510 nm on a SpectraMax Multi-Mode Microplate Reader with SoftMaxPro software, settings: 1 wavelength, 510 nm, plate check, auto mix. Excel was used to plot average results over at least 5 replicates with error reported as standard error.

A.1.4 General Procedure: Transformation of E. coli

High efficiency, competent *E. coli* cells, type C2984I in SOC buffer (stored at -80 °C) were placed on ice to thaw until crystals disappeared. To a new 0.5 mL Eppendorf tube was added 50 µL *E. coli* cells and the tube was placed on ice. To the tube was added 2 µL of desired plasmid (e.g., $(CTG{\cdot}CAG)_{74}$, $(CTG{\cdot}CAG)_{90}$, DT0, DT960, insulin resistance (IR) minigene, ERT, pEGFP-Q23, or pEGFP-Q74) (300-900 ng/µL). Both $(CTG{\cdot}CAG)_{90}$ and $(CTG{\cdot}CAG)_{74}$ were obtained from Maurice Swanson's lab and have the (CTG·CAG) repeat sequence inserted between BamHI and EcoRI sites. The DT960 plasmid was obtained from the lab of Thomas Cooper (Baylor College of Medicine) and the IR minigene plasmid was obtained from the lab of Nicholas Webster (University of California, San Diego). The tube was flicked to mix, incubated on ice (0 °C) for 30 min, heat

shocked for exactly 30 s, and cooled on ice for 5 min. To a new 1 mL tube was added 950 µL SOC buffer and the *E. coli*/plasmid mixture. The samples were incubated on a shaker at 37 °C for 1 h, pipetted onto agar/ampicillin or agar/kanamycin plates (depending on the antibiotic resistance gene in the plasmid, pEGFP plasmids contain kanamycin and all others listed above conatin ampicillin resistance gene), and spread. The plates were incubated at 37 °C with agar side down for 15 min, agar side up for 14 h, and stored at -4 °C until ready to use single colonies. In a new 50 mL centrifuge tube was mixed 10 mL LB media and 10 µL ampicillin or kanamycin B (50 mg/mL). A single colony was isolated from the plates of transformed *E. coli* and added to the 50 mL centrifuge tube. The caps were loosely secured, and the tubes were incubated on a shaker at 37 °C for about 16 h to obtain a tan, cloudy solution. The caps were tightened, the tubes were centrifuged at 5,000 rpm for 20 min, the supernatant was removed, and the cell pellet was used to extract plasmids or stored at -80 °C until plasmids were extracted.

A.1.5. General Procedure: Plasmid Extraction

QIAprep Spin Mini Prep Kit (Qiagen) was used to extract the plasmid from the cells, following manufacturer protocols with double the volume. The pelleted bacteria cells were resuspended in 500 µL buffer P1 and transferred to a 2 mL Eppendorf tube. To the tube was added 250 µL buffer P2 (lysis buffer) and the tube was inverted to mix until the solution became colorless/translucent, keeping the lysis process under 5 min for each tube. To stop lysis, 700 µL buffer N3 was added to the tube and it was inverted several times to mix. The tube was centrifuged at 13,000 rpm for 10 min, the supernatant (750 µL at a time) was transferred to a QIA prep 2.0 spin column, and the column was centrifuged at 13,000 rpm for 60 s (repeated until all supernatant was passed through the column). The liquid flow-through was discarded and the column was washed with 500 µL

buffer PB, centrifuged for 60 s at 13,000 rpm, and the flow-through was discarded. To dry the column, it was centrifuged at 13,000 rpm for 60 s and the flow-through was discarded. The column was placed in a new 1.5 mL Eppendorf tube and the DNA was eluted by adding 50 µL buffer EB (10 mM TrisCl, pH 8.5), incubating at room temperature for 60 s, and centrifuging at 13,000 rpm for 60 s. The concentration was measured with a Thermo Scientific nanodrop spectrophotometer using Nucleic Acid, DNA setting and blanking with EB buffer. Final concentrations ranged from about 100 ng/µL – 900 ng/µL, depending on the plasmid, bacteria storage time, and isolation efficiency.

A.1.6. General Procedure: Restriction Enzyme Digest

Restriction enzymes (BamHI-HF, HindIII-HF, SalI-HF, EcoRI-Hf, NsiI) were purchased from New England Biolabs (NEB) and the manufacturer protocol was followed. Each reaction tube contained 50 µL solution: 1 µg DNA, 5 µL 10X cut smart buffer (NEB), 1 µL of the respective restriction enzyme, diluted to 50 µL with nuclease-free water. The proper volume of nuclease-free water (dependent on plasmid concentration) was added to a new 0.5 µL Eppendorf tube, 5 µL cut smart buffer was added, and 1 µL restriction enzyme was added to the tube. The samples were incubated at 37 °C for 15 min, 2 µL shrimp alkaline phosphatase (rSAP, from NEB) was added to prevent re-cyclization of the plasmid, and the tubes were incubated at 37 °C for 30 min. The cut plasmids were purified with QIAquick PCR Purification Kit (Qiagen), according to manufacturer protocol. To a new 1.5 mL Eppendorf tube was added 500 µL buffer PB and 2 x 50 µL digestion reactions from above (or 250 µL buffer PB with 1 x 50 µL reaction.) The tubes were inverted to mix, the contents were transferred to a QIAquick spin column, and the columns were centrifuged at 13,000 rpm for 60 s. The flow-through was discarded, the column was washed with 750 µL

buffer PE, and centrifuged at 13,000 rpm for 60 s. The flow-through was discarded and the column was dried by centrifuging at 13,000 rpm for 60 s. The column was transferred to a new 1.5 mL Eppendorf tube and 30 µL buffer EB was added to center of the column and incubated for 60 s. The column was centrifuged at 13,000 rpm for 60 s to elute the DNA. The concentration was measured with a Thermo Scientific nanodrop spectrophotometer using Nucleic Acid, DNA setting and blanking with EB buffer. Final concentrations (in 50 µL) ranged from about 10-176 ng/µL depending on the plasmid and isolation efficiency. An analytical agarose gel was run to confirm successful cleavage.

A.1.7. General Procedure: Analytical Agarose Gel.

The gel was cast by mixing 0.5 g agarose, 5 mL 10X TBE buffer, 45 mL DI water in a 50 mL beaker, microwaving 15 s at a time until all agarose dissolved, , adding 1 µL ethidium bromide solution (10 mg/mL), slowly pouring into the mold, removing bubbles, sliding the comb into the liquid, and allowing to set for about 30 min. The buffer solution for running the gel was made with 50 mL 10X TBE Buffer, 450 mL DI water, and 5 µL ethidium bromide. Samples were prepared by mixing an appropriate amount of analyte (dependent on concentration) with 6X loading dye (purple). The gel was loaded with 2 µL of 1 kb DNA ladder (NEB) and 2 µL of each sample, run at 150 V for 30 min or until the ladder had migrated about two-thirds of the way down the gel, and imaged on a Bio-Rad Gel Doc System. A representative gel from BamHI cleavage is provided in Figure A.1.

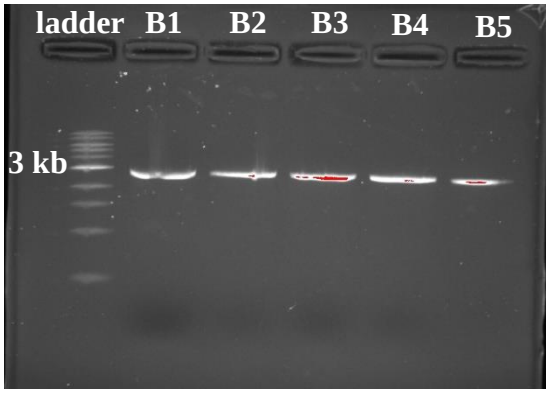


Figure A.1. Sample analytical agarose gel of product from BamHI-HF cleavage reaction. Lane 1 contains NEB 1 kb DNA ladder (labeled at 3 kb). The plasmid is about 3 kb. The circular (uncut or re-cyclized plasmid) should appear at higher molecular weight (and sometimes with 2 bands due to supercoiled and open circular forms) due to structural difference from the linear 3 kb restriction-enzyme cleaved plasmid. B1-B5 in lanes 1-5 indicate replicate reactions of the same Bam-HI cleavage reaction described above.

A.1.8. in vitro Transcription Inhibition Assays

For plasmid growth and amplification procedures, see General Procedures A.1.4, A.1.5, and A.1.6 above.

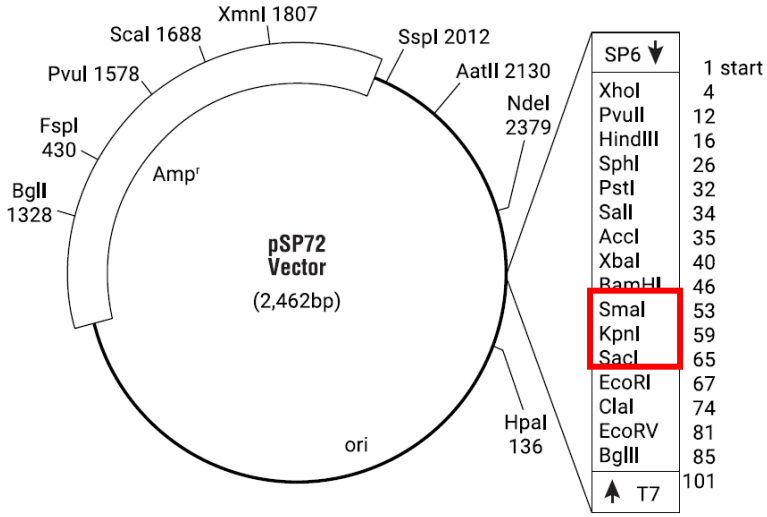


Figure A.2. Plasmid pSP72. The red box indicates where the (CTG·CAG) repeat sequence was cloned into the plasmid.

***in vitro* Transcription Inhibition Assay with T7 RNA Polymerase:**[2] Using linearized BamHI-cleaved $(CTG{\cdot}CAG)_{90}$ pSP72 plasmid (15 ng/µL) or PCR amplified pGEM random duplex

plasmid (15 ng/μL) and T7 RNA polymerase from NEB, following manufacturer's protocols for *in vitro* transcription. Because BamH1 restriction enzyme digest used with T7, the $(CAG)_{90}$ strand will be transcribed to form $(CUG)_{90}$. Each reaction tube contained 10 μL solution: 0.015 μg (1 μL) BamHI-digested plasmid solution, 1 μL 10X polymerase reaction buffer, 4 μL 10 mM NTPs (1 μL each of 10 mM CTP, 10 mM GTP, 10 mM ATP, 10 mM UTP), a proper volume of 100 μM compound solution, and diluted to a final volume of 10 μL with nuclease-free water. To a new 0.5 μL Eppendorf tube was added proper amounts of nuclease-free water, reaction buffer, NTPs, template DNA, and compound. The Eppendorf tubes were flicked to mix and incubated at 37 °C for 3.5 h or 24 h (for 24 h incubation experiments.) To each tube was added 1 μL T7 RNA polymerase and the tubes were incubated for an additional 1.5 h. The reaction was stopped by adding 6.5 μL 8 M urea solution (aqueous) followed by 1 μL RNA loading dye (95% Formamide, 0.02% SDS, 0.02 % Bromophenol Blue, 0.01% Xylene Cyanol, 1 mM EDTA). The tubes were heated at 95 °C for 5 min and immediately cooled to 4 °C. A 10% urea PAGE gel was run at 250 V (80 mA) for 30 min and post-stained with EtBr (10 μL, 10 mg/mL solution in ~15 mL DI water, incubated on shaker at RT for 15 min.) The gel was imaged on a BioRad Gel Doc XR and RNA was quantified based on fluorescence intensity. Bands were quantified with ImageJ by drawing a box of same size around each band, plotting the lanes, drawing a straight line under each curve, and integrating with the wizard tool in ImageJ. Normalized % transcript = integration of treatment / integration of control x 100%. Percent inhibition = 1 – normalized % transcript. Data was plotted in RStudio and fitted to a logistic curve using the SSlogis model. The output logistic equation was used to calculate IC_{50} values and error is reported from standard error of the coefficients in the fit equation. Curves that did not fit the logistic model were plotted with a linear model for visual comparison.

***in vitro* Transcription Inhibition Assay with SP6 RNA Polymerase:**[3] Using linearized EcoRI-cleaved $(CTG{\cdot}CAG)_{90}$ pSP72 plasmid (15 ng/μL) or PCR amplified pGEM random duplex plasmid described above (15 ng/μL) and SP6 RNA polymerase from NEB, following manufacturer's protocols for *in vitro* transcription. Because EcoRI restriction enzyme digest plasmid was used with SP6, the $(CTG)_{90}$ strand will be transcribed to form $(CAG)_{90}$. Each reaction tube contained 10 μL solution: 0.015 μg (1 μL) EcoRI-digested plasmid solution, 1 μL 10X polymerase reaction buffer, 4 μL 10 mM NTPs (1 μL each of 10 mM CTP, 10 mM GTP, 10 mM ATP, 10 mM UTP), a proper volume of compound solution to achieve desired final concentration, and diluted to a final volume of 10 μL with nuclease-free water. To a new 0.5 μL Eppendorf tube was added proper amounts of nuclease-free water, reaction buffer, NTPs, template DNA, and compound. The Eppendorf tubes were flicked to mix and incubated at 37 °C for 3.5 h. To each tube was added 1 μL SP6 RNA polymerase and the tubes were incubated for an additional 1.5 h. The reaction was stopped by adding 6.5 μL 8 M urea solution (aqueous) followed by 1 μL RNA loading dye (95% Formamide, 0.02% SDS, 0.02 % Bromophenol Blue, 0.01% Xylene Cyanol, 1 mM EDTA). The tubes were heated at 95 °C for 5 min and immediately cooled to 4 °C. A 10% urea PAGE gel was run at 250 V (80 mA) for 30 min and post-stained with EtBr (10 μL, 10 mg/mL solution in ~15 mL DI water, incubated on shaker at RT for 15 min.) The gel was imaged on a BioRad Gel Doc XR and RNA was quantified based on fluorescence intensity. Bands were quantified with ImageJ by drawing a box of same size around each band, plotting the lanes, drawing a straight line under each curve, and integrating with the wizard tool in ImageJ. Normalized % transcript = integration of treatment / integration of control x 100%. Note: Random plasmid pTRI-Xef (Invitrogen) was used in some experiments, but the experiment did not work

well (the transcription inhibition tracked directly with the pSP72 plasmid with no observed difference in the random control.)

A.1.9. General Procedure: Insulin Receptor (IR) Mis-Splicing Rescue Assay[4]

HeLa cells (from cryogenic storage) were grown at 37 °C in T75 flasks in Dulbecco's Modified Eagle Medium media prepared in the Cell Media Facility, University of Illinois, with 10% Fetal Bovine Serum added. Cell media was changed at least every 2 d. When cells reached 100% confluence, the cells were washed with PBS (37 °C) and about 10 mL 37 °C trypsin (0.05% (1x) with EDTA) was added. After incubation at 37 °C for 2 min, the T75 flask was gently tapped to remove any cells that remained stuck on the flask. The contents of the T75 was transferred to a conical tube containing 3 volumes of cell media (relative to trypsin). The cells were counted, centrifuged at 500 G for 5 min, and the supernatant was aspirated off. The cells were diluted to 50,000 cells/mL. Harvested cells were used immediately in mis-splicing assays. A 6-well plate was used for this assay. To each well was added 1 mL of the 50,000 cells/mL solution and the cells were incubated at 37 °C for about 12 h until the cells reached 70-80% confluence.

Cells were transfected with IR minigene plasmid and 5 wells were transfected with DT960 plasmid (see A.1.4-A.1.5 for *E. coli* amplification of plasmid). The transfection protocol is as follows. In a 1.5 mL Eppendorf tube was mixed 600 μL Gibco OPTI-MEM media and 22.5 μL lipofectamine 2000 (Invitrogen), a 1:3 plasmid:lipofectamine ratio in each of 6 wells. The solution was vortexed, spun down, and incubated exactly 5 min at room temperature. In a separate 1.5 mL Eppendorf tube was mixed 1.25 μg IR minigene plasmid and 100 μL OPTI-MEM media for well 1. In a third Eppendorf tube was mixed 2.5 μg IR minigene plasmid, 3.75 μg DT960 plasmid, and 500 μL OPTI-MEM media. After 5 min incubation of lipofectamine mixture, 103 μL was added

to the IR plasmid solution and 515 μL was added to the IR + DT960 plasmid solution. The resulting solution was incubated at room temperature for 20 min. During the incubation, cells were removed from the incubator and washed with 1X PBS. To each well was added 2 mL OPTI-MEM and the plate was returned to the 37 °C incubator. After 20 min incubation of transfection solutions, 200 μL of the IR solution was added to well 1 and 200 μL of the IR + DT960 solution was added to wells 2-5. The cells were incubated at 37 °C for 4 h.

The transfection media was replaced with DMEM + FBS cell media and treated with ligand. The plates were incubated at 37 °C for 72 h, after which the media was aspirated and the wells were washed with 1X PBS. To each well was added 750 μL warm trypsin and the plate was incubated at 37 °C for 2 min. The plate was tapped to release remaining cells and 2 mL DMEM + FBS cell media was added to each well. The detached cells were transferred to 15 mL falcon tubes (1 for each well), the wells were rinsed with 1 mL media, and the tubes were centrifuged at 500 G for 5 min. The supernatant was aspirated off and the falcon tubes were placed on ice.

Total RNA extraction was performed with E.Z.N.A. Total RNA Kit I. To each tube was added 350 μL TRK lysis buffer, the cells were pipetted up and down until the cells were dissolved into clear, homogeneous solution which was loaded onto a QIAshredder column. The columns were centrifuged at 13,300 rpm for 2 min to remove cell debris. To the filtrate was added 350 μL 70% ethanol, pipetted up and down, the solution was loaded onto a HiBind RNA mini column, and centrifuged at 10,000 xG for 1 min. The filtrate was discarded, and the columns were washed twice with 500 μL RNA wash buffer II and centrifuged at 10,000 xG for 1 min. The columns were centrifuged at 13,300 rpm for 2 min to dry and the columns were transferred to 1.5 mL Eppendorf tubes. The RNA was eluted by adding 70 μL DEPC water (included in EZNA kit), incubating for 1 min, and centrifuging at 10,000 xG for 1 min. The concentration was measured with a Thermo

Scientific nanodrop spectrophotometer using Nucleic Acid, RNA setting and blanking with DEPC water from kit.

The samples were prepared for reverse transcription with BioRad iScript cDNA Synthesis kit. In a 0.5 mL Eppendorf tube was mixed 4 μL 5X Transcription Buffer, 1 μg RNA, nuclease-free water to volume of 19 μL, and 1 μL RT Enzyme from kit (20 μL final volume). The tubes were incubated at 25 °C for 4 min, 46 °C for 20 min, and 95 °C for 1 min, then cooled to 4 °C. The reaction mix was purified with QIAquick PCR purification kit, following manufacturer protocol. To each tube was added 5 volumes (100 μL) Buffer PB to the PCR reaction. The sample was applied to a QIAquick column and centrifuged at 13,300 rpm for 1 min. The flow-through was discarded and the column was washed with 750 μL Buffer PE and centrifuged at 13,300 rpm for 1 min. The flow-through was discarded and the column was spun at 13,300 rpm for 1 min to dry. The DNA was eluted by adding 50 μL Buffer EB, incubating for 1 min, and centrifuging at 13,300 rpm for 1 min. The concentration was measured with a Thermo Scientific nanodrop spectrophotometer using "Nucleic Acid," "DNA" setting and blanking with Buffer EB.

PCR amplification of cDNA was run with forward primer: 5'-GTA CCA GCT TGA ATG CTG CTC CTG-3' and reverse primer 5'-CTC GAG CGT GGG CAC GCT-3'. A 10 μM master mix of primers was made and mixed well. To each of 6 0.5 mL Eppendorf tubes was added 2 μL of the primer mix, 70 ng cDNA, nuclease-free water to a volume of 25 μL, and 25 μL of Go Taq Green Master Mix (Promega). The tubes were incubated at 95 °C for 2 min followed by 35 PCR cycles (95 °C for 30 s, 65 °C for 45 s, 72 °C for 30 s), and 72 °C for 5 min. The tubes were cooled to 0 °C and a non-denaturing PAGE gel was run to analyze samples. A 1 mm gel was set by mixing 1 mL 10X TBE, 9 mL DI water, 2.5 mL acrylamide (37.5:1), 100 μL 30% ammonium persulfate solution, and 4 μL tetramethylethylenediamine (TEMED). The gels were pre-run at 250 V for 30

min, and 20 μL of each sample was loaded onto the gel and it was run at 250 V for 20 min. The gel was stained with 5 μL ethidium bromide (10 mg/mL) in 10 mL DI water, incubated on a shaker for 20 min, and imaged on a Bio-Rad Gel Doc System. Bands were quantified with ImageJ by drawing a box of same size around each band, plotting the lanes, drawing a straight line under each curve, and integrating with the wizard tool in ImageJ. Microsoft Excel was used to make figures using the average of at least 3 independent experiments with error reported standard error.

A.1.10. General Procedure: Construction and PCR Amplification of Random pGEM Plasmid

Construction of Random pGEM Plasmid: The plasmid was created using Promega's pGEM-T Easy Vectors Kit and following the manufacturer's protocol. Kit components (2X Rapid Ligation Buffer, pGEM-T Easy Vector, and Control Insert DNA) were thawed, vortexed, and centrifuged down before use. In a 0.5 mL Eppendorf tube was mixed 5 μL of 2X Rapid Ligation Buffer, 1 μL of pGEM-T Easy Vector, and 2 μL of Control Insert DNA. Added 1 μL of T4 DNA Ligase to the tube (kept in the freezer until ready to use.) The resulting solution was incubated at 4 °C for 24 h before transforming into *E. coli*.

Promega A363A Control Insert Sequence (used from Promega pGEM-T Easy Vector Kit as delivered):
CTTGATTGACAAGGATGGATGGCTACATTCTGGAGACATAGCTTACTGGGACGAAG
ACGAACACTTCTTCATAGTTGACCGCTTGAAGTCTTTAATTAAATACAAAGGATATC
AGGTGGCCCCCGCTGAATTGGAATCGATATTGTTACAACACCCCAACATCTAATTAG
CGGAGTGGCAGGTCTTCCCGACGATGACGCCGGTGAACTTCCCGCCGCCGTTGTTGT
TTTGGAGCACGGAAAGACGATGACGGAAAAAGAGATCGTGGATTACGTGGCCAGTC
AAGTAACAACCGCGAAAAAGTTGCGCGGAGGAGTTGTGTTTGTGGACGAAGTACCG
AAAGGTCTTACCGGAAAACTCGACGCAAGAAAAATCAGAGAGATCCTCATAAAGGC
CAAGAAGGGCGGAAAGTCCAAATTGTAAAATGTAACTGTATTCAGCGATGACGAAA
TTCTTAGCTATTGTAATCCTCCGAGGCCTCGAGGTCGACGAATTCGAGCTCGGCCGA
CTTGGCCAATTCGCCCTATAGTGAGTCGTATTAA

PCR Amplification: Using pGEM normal duplex plasmid solution prepared from T-Easy kit, mixed 10 ng of plasmid, 2.5 µL each of forward (T7) and reverse (SP6) primers (from 10 µM final volume, from 100 µM solution) and the appropriate aliquot of molecular biology grade water for a final volume of 25 µL. Vortexed and centrifuged before adding 25 µL of Thermo Scientific 2X Phusion Master Mix w/ HF Buffer (Ref. F-531) that contains polymerases as well as NTPs. The PCR solution was incubated on a thermocycler using the following program: 98 °C 30 sec then 30 cycles of 98 °C 30 sec, 56 °C 40 sec, 72 °C 60 sec, and 72 °C 5 min followed by cooling to 4 °C. After PCR reaction, purified with QIAGEN QIAquick PCR Purification Kit. Combined 5x50 µL reactions (total 250 µL) in one 2 mL Eppendorf tube and added 5 volumes (1250 µL) of Buffer PB. The tube was inverted tube several times to mix and the contents were transferred to a spin column 500 µL at a time. The spin columns were spun down on a microcentrifuge at the 13,300 rpm for one min and the flow through was discarded. This process was continued until all the solution was washed through the spin column. To wash the column, added 750 µL Buffer PE and centrifuged for one min at 13,300 rpm. The flow through was discarded and the column was centrifuged for one min at 13,300 rpm to dry. The flow through was discarded and the spin column was transferred to a new 1.5 mL Eppendorf tube. To the center of the spin column was added 50 µL Buffer EB (elution buffer) and the column was allowed to sit for 1 min followed by centrifugation for 1 min at 13,300 rpm. The concentration was tested using a NanoDrop and the PCR amplification yielded pGEM random duplex plasmid ranging from concentration 174 to 222 ng/µL in 50 µL total volume. The resulting plasmid solution was used within 1 week of synthesis and freeze-thaw cycles were minimized.

Primers: F-T7 = 5'-TAA TAC GAC TCA CTA TAG GG-3', R-SP6 = 5'-ATT TAG GTG ACA CTA TAG AAT AC-3'

A.1.11. in vitro Transcription of d(CTG·CAG)n Template to form RNA Templates

***in vitro* transcription of $(CTG·CAG)_n$ template.** The T7 or SP6 promoter was used to transcribe $d(CTG·CAG)_{16}$ template to form r(GGG $(CUG)_{16}$ CCC). DNA Strands: "A" = 5' – ACG CAC GCT GTA ATA CGA CTC ACT ATA GGG $(CTG)_{16}$ CCC – 3'; "B" = 5' – GGG $(CAG)_{16}$ CCC TAT AGT GAG TCG TAT TAC AGC GTG CGT – 3'. T7 was also used to transcribe $d(CTG·CAG)_{74}$ and $d(CTG·CAG)_{90}$ (both plasmids obtained from Maurice Swanson's group and cleaved at BamHI to linearize as described above) to make $r(CUG)_{74}$ and $r(CUG)_{90}$ or $r(CAG)_{90}$. cleaved at BamHI for T7 transcription or EcoRI for SP6 transcription to linearize as described above) to make $r(CUG)_{90}$ and $r(CAG)_{90}$, respectively. T7 binds at the promoter sequence 5' – TAATACGACTCACTATAG – 3', starts transcription at the underlined G, and transcribes using the opposite strand as a template from 5' to 3' to form $r(CUG)_{16}$. The manufacturer protocols for T7 and SP6 polymerases (NEB) were followed. In a 20 µL total volume was mixed nuclease-free water, reaction buffer (from NEB), and DNA template (about 1 µg DNA). RNase inhibitor (1 µL) was added followed by 2 µL T7 or SP6 RNA polymerase, and the tube was inverted to mix and incubated on a shaker at 37 ^{o}C for 3 h (for longer $r(CUG)_{74}$ and $r(CUG)_{90}$ synthesis).

For $r(CUG)_{16}$ synthesis, a master mix was prepared (of the appropriate amounts for the number of replicate reactions, each 20 µL total volume) of nuclease-free water, reaction buffer (from NEB), and both DNA template strands (about 1 ug DNA). The Eppendorf tube containing this master mix was heated at 95 °C for 5 min and then allowed to cool slowly to room temperature to anneal the DNA (annealing performed for single stranded $(CTG)_{16}$ template strands only). The appropriate aliquot of the master mixture was added to a 0.5 mL Eppendorf tube and 4 µL of NTP mixture was added (consisting of 1 µL of each from 100 mM stock solutions from NEB). RNase

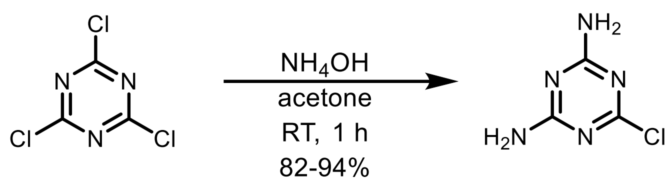
inhibitor (1 µL) was added to the tube, 2 µL T7 polymerase was added, and the tube was inverted to mix and incubated on a shaker at 37 °C for 18-20 h (for shorter r$(CUG)_{16}$ synthesis only).

To remove the DNA template, 2 µL DNase (NEB) was added, and the tubes were flicked to mix. After incubation at 37 °C for 15 min, all replicate reactions were combined into a 1.5 mL Eppendorf tube, and the product was purified by either phenol:chloroform extraction or using NEB Monarch RNA Cleanup Kit following the manufacturer protocols. For phenol:chloroform extraction, 1 volume of cold phenol:chloroform:isoamyl alcohol 25:24:1, saturated with 10 mM Tris, pH 8.0, 1 mM EDTA (4 °C) was added to the tube, inverted to mix, incubated on ice for 10 min, and inverted again. The resulting clear solution was centrifuged at 13,000 rpm for 2 min, the top layer was removed and placed in a new 1.5 mL Eppendorf tube. To the top layer was added 1 volume chloroform and the tube was mixed to invert, microfuged at 13,000 rpm for 1 min, and the top layer was placed in a new 1 mL Eppendorf tube. To the top layers was added 0.10 volume 3 M sodium acetate (aqueous), pH 5.2, and 2.5 volumes cold 100% ethanol (-20 °C). The suspension was mixed by inverting and incubated at -20 °C overnight. The tube was centrifuged at 13,000 rpm for 30 min at 4 °C and the supernatant ethanol solution was removed. The pellet was washed by adding 500 µL of 75% ethanol in nuclease-free water (-20 °C) and the tube was centrifuged at 13,000 rpm for 10 min. The supernatant ethanol layer was removed, and the pellet was washed again with 500 µL 75% ethanol in nuclease-free water, centrifuged at 13,000 rpm for 10 min, and the supernatant was removed. The pellet was dried via Thermo Scientific Savant ISS110 SpeedVac and dissolved in 50 µL RNA storage buffer. The concentration was measured using a Thermo Scientific nanodrop spectrophotometer set to Nucleic Acids, RNA. For Monarch Spin Columns, manufacturer instructions were followed. The purify of samples was assessed via analytical polyacrylamide gel electrophoresis (PAGE).

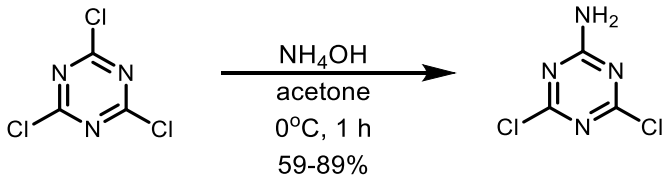
r(CUG)$_{90}$ = 5’ – G GGAGACCGGC AGAUCUGAUA UCAUCGAUGA AUUC CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG CUG GAGCUC GGUACCCGG G – 3’(synthesized from pSP72 (CTG·CAG)$_{90}$ plasmid, MW = 102,571.1 g/mol, calculated using IDT OligoAnalyzer Tool)

r(CAG)$_{90}$ = 5’ – GAACUCGAGC AGCUGAAGCU UGCAUGCCUG CAGGUCGACU CUAGAGGAUC CCCGGGUACC GAGCUC CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG G – 3’(synthesized from pSP72 (CTG·CAG)$_{90}$ plasmid, MW = 109,737.8 g/mol, calculated using OligoAnalyzer)

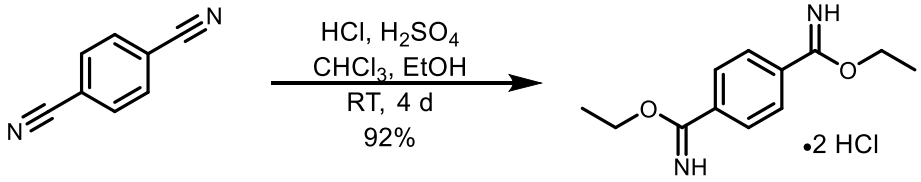
A.1.12. General Synthetic Methods



6-Chloro-1,3,5-triazine-2,4-diamine.[5] To a 500 mL round-bottom flask was added 25.3 g (137 mmol) cyanuric chloride and 50 mL acetone. The mixture was stirred at room temperature to make a white suspension and a mixture of 50 mL water and 50 mL 28% w/w aqueous ammonium hydroxide solution (400 mmol) was added dropwise via an addition funnel, giving a white precipitate. TLC (9 DCM:1 MeOH, Rf 0.5) was used to monitor formation of product. The solid was filtered using a Buchner funnel, washed with DI water, and dried on a lyophilizer to afford 18.7 g (94%) of the title compound as a white solid: ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.13 (bd, J = 39.9, 4H); ^{13}C NMR (500 MHz, d$_6$-DMSO) δ 169.11, 167.54; LR-ESI-MS (m/z) calcd for [M+H$^+$] 146.02; found 146.0.



4,6-Dichloro-1,3,5-triazin-2-amine.[6] To a 500 mL round-bottom flask was added 53 g (290 mmol) cyanuric chloride and 100 mL acetone and the mixture was stirred. The resulting white suspension was cooled to 0 °C. A mixture of 38 mL 28% w/w aqueous ammonium hydroxide solution (300 mmol) and 100 mL DI water was added to a 200 mL addition funnel. When the suspension in the round-bottom flask reached 0 °C, the ammonium hydroxide solution was added dropwise over 40 min. The resulting white suspension was stirred for an additional 20 min. TLC (9 DCM:1 MeOH, Rf 0.5) was used to monitor formation of product. The mixture was filtered using a Buchner funnel and washed with 250 mL DI water. The product was dried on a lyophilizer to remove excess water to yield 42 g (89%) of the title compound as a white powder. ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.56 (s, 2H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 169.33, 167.10. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 164.97; found 165.0.

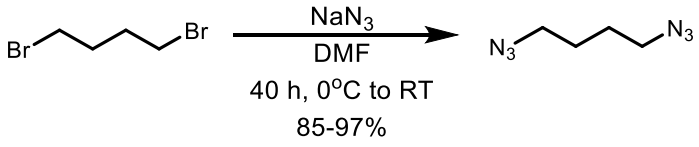


Diethyl terephthalimidate dihydrochloride.[7] All glassware was dried in oven before use. To a 3-necked 250 mL 24-40 round-bottomed flask was added 10.25 g (80.0 mmol) terephthalonitrile, 100 mL chloroform, and 100 mL dry ethanol. Reaction apparatus was assembled to bubble gas through the reaction. In a separate 2-necked, 200 mL RBF was added 85 mL concentrated hydrochloric acid, and 85 mL concentrated sulfuric acid was added dropwise via addition funnel to generate hydrochloride gas. The white solution was stirred under nitrogen atmosphere at room

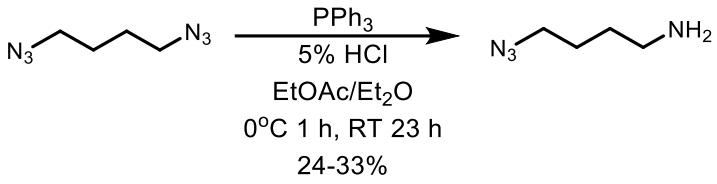
temperature for about 4 d, replacing the acid three times. Crude NMR (DMSO-d_6) was used to monitor reaction progress. The white precipitate was filtered off and washed with dry ethanol (~800 mL) to yield 21.58 g (92%) of the title compound as white solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.24 (s, 4H), 4.58 (q, J = 7.18, 4H), 1.46 (t, J= 6.86, 6H). ^{13}C NMR could not be obtained because the product is water sensitive and will break down with even a small amount of water in the NMR solvent over the time of the ^{13}C NMR experiment.

A.2. TEMPLATE-ASSISTED ASSEMBLY VIA CLICK CHEMISTRY

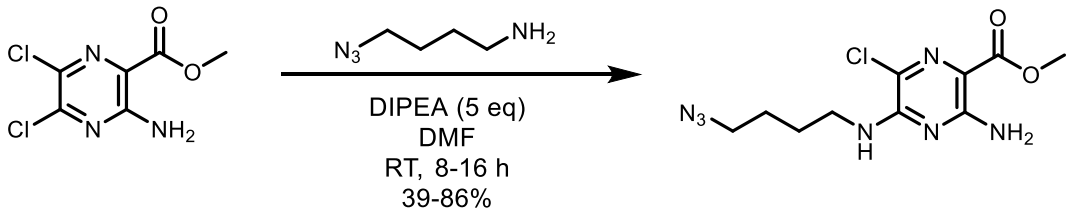
A.2.1. Synthetic Methods



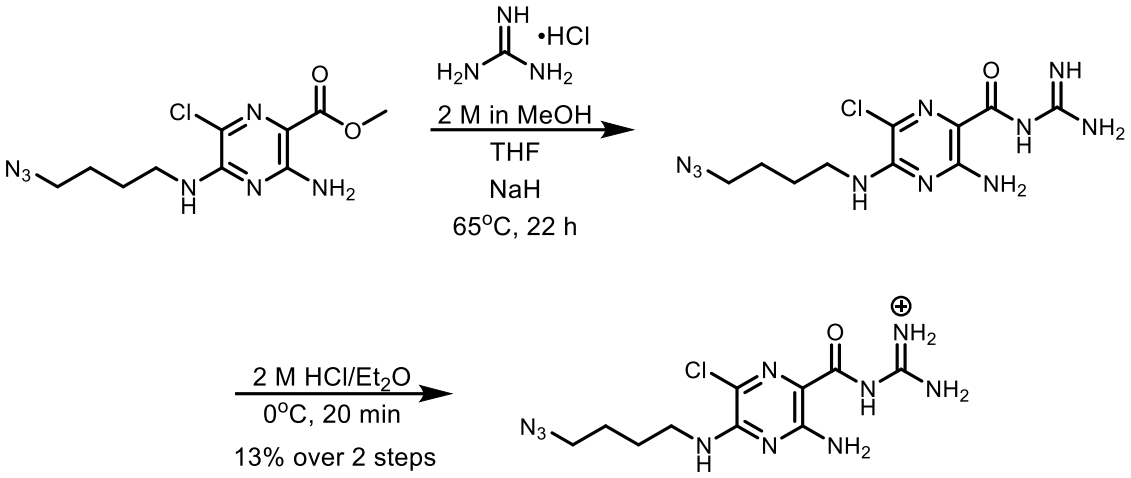
1,4-diazidobutane.[8,9] To a 250 mL round-bottomed flask was added 5.5 mL (46 mmol) 1,4-dibromobutane and 54 mL dry DMF and the mixture was cooled to 0 °C. To the clear solution was added 9.0 g (140 mmol) sodium azide while stirring to form a white suspension that was cooled slowly to room temperature and stirred for up to 40 h until no starting material was observed by TLC (3 EtOAc:2 Hex, PPh_3 and Ninhydrin stains, Rf = 0.85). After the 1,4-dibromobutane had been consumed, 20 mL DI water was added to quench remaining sodium azide and the clear, translucent solution was stirred for an additional 30 min. The product was extracted with 3 x 50 mL hexanes, the organic layers were combined, washed twice with 25 mL DI water (to remove DMF), dried with sodium sulfate, and the solution was concentrated *in vacuo* to afford 6.3 g (97%) of the title compound as a clear liquid. Note: the product is somewhat volatile. ^{1}H NMR (500 MHz, $CDCl_3$) δ 3.34 (t, J = 6.25, 4H), 1.70 (p, J = 3.17, 4 H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 50.95, 26.20. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 141.08; found 141.1.



4-azidobutan-1-amine.[10] To a 500 mL round-bottomed flask was added 9.62 g (68 mmol) 1,4-diazidobutane, 43 mL diethyl ether, and 43 mL ethyl acetate. The mixture was cooled to 0 °C and placed under N_2 atmosphere. To the resulting clear solution 69 mL 5% HCl (aqueous) was added dropwise via addition funnel to obtain a biphasic solution. The solution was stirred gently at 0 °C for 10 min and 17.63 g (67.2 mmol) of triphenylphosphine was added in small portions over 45 min, keeping the mixture at 0 °C. The resulting white, biphasic mixture was slowly warmed to room temperature and stirred for 20-24 h. TLC was used to monitor reaction progress (9 MeOH:1 NH_4OH, Rf = 0.3). The contents of the RBF were decanted into a 500 mL separatory funnel and the aqueous layer was washed with 3x100 mL DCM and 2x100 mL ethyl acetate. The pH of the aqueous layer was adjusted to 13 by slowly adding about 60 mL of 2 M sodium hydroxide (aqueous). The basic aqueous layer was washed 5 times with 100 mL DCM and the organic layers were combined, dried over sodium sulfate, and concentrated *in vacuo* to yield 2.56 g (33%) of the title compound as a tan-colored oil. ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.62 (bs, 2H), 3.30 (t, J = 6.68, 2H), 2.74 (t, J = 7.04, 2H), 1.62 (dddd, J = 13.25, 6.44, 3.39, 1.52, 2H), 1.55 (dddd, J = 0.97, 2.15, 6.14, 11.7, 2H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 51.32, 41.53, 30.46, 26.30. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 115.09; found 115.1.

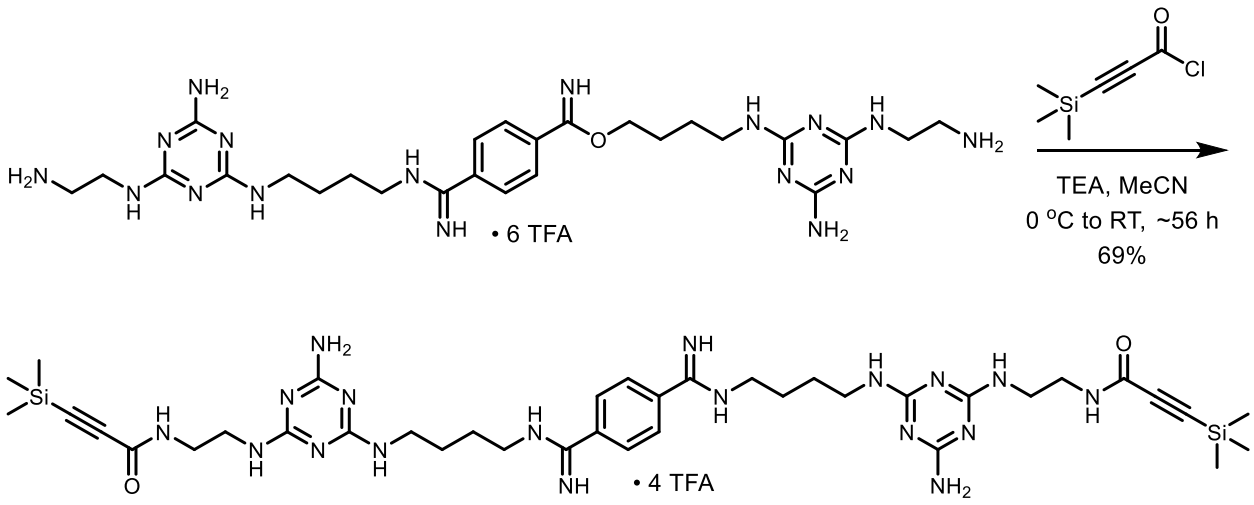


Methyl-3-amino-5-((4-azidobutyl)amino)-6-chloropyrazine-2-carboxylate. To a 15 mL, oven-dried round-bottom flask was added 0.081 g (0.36 mmol) methyl-3-amino-5,6-dichloropyrazine-2-carboxylate followed by 2.0 mL dry DMF. To the resulting dark brown solution was added 0.31 mL (1.8 mmol) *N*,*N*-diisopropyl ethyl amine and the mixture was stirred for 5 min. In a separate vessel was combined 62 mg (0.54 mmol) 4-azidobutan-1-amine and 0.80 mL dry DMF and the resulting light tan solution was added to the 15 mL round-bottom flask dropwise under nitrogen atmosphere. The dark brown solution was stirred at room temperature for about 20 h. Reaction progress was monitored by TLC (5 Hex:1 EtOAc, Rf = 0.2). The dark brown mixture was concentrated *in vacuo* to obtain a brown residue which was partitioned between 50 mL each of DI water and ethyl acetate. The organic layer was washed with water (3 x 50 mL) and saturated brine (2 x 50 mL) and dried over sodium sulfate. The light yellow, transparent liquid was concentrated *in vacuo* to afford 95 mg (86%) of the title compound as a brown residue. ^{1}H NMR (500 MHz, $CDCl_3$) δ 5.53 (brt, *J* = 5.88, 1H), 3.83 (s, 3H), 3.43 (q, *J* = 6.57, 2H), 3.29 (t, *J* = 6.44, 2H), 1.68 (m, 2H), 1.61 (m, 2H), 1.56 (brm, 2H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 166.54, 155.54, 151.55, 121.38, 110.43, 52.05, 51.05, 40.73, 26.47, 26.29. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 300.09; found 300.1.



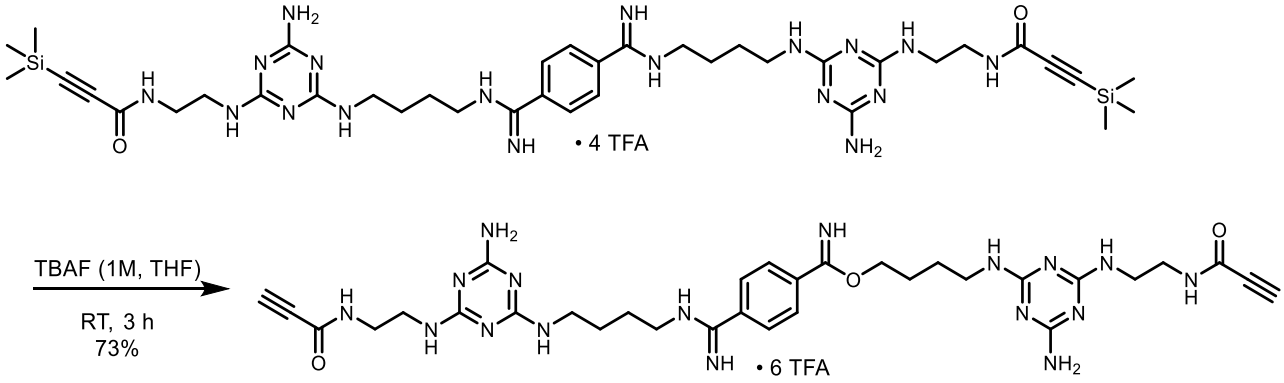
Amino(3-amino-5-((4-azidobutyl)amino)-6-chloropyrazine-2-carboxamido)methaniminium. To a 10 mL round-bottom flask was added 71.1 mg (0.237 mmol) methyl-3-amino-5-((4-azidobutyl)amino)-6-chloropyrazine-2-carboxylate and 1.75 mL THF (0.10 M total reaction concentration). In a separate vessel was mixed 114 mg (1.19 mmol) guanidine hydrochloride and 0.60 mL methanol and added the solution to the 10 mL round-bottom flask. The resulting light orange solution was stirred at 65 °C for 4 h and 54.7 mg (2.28 mmol) sodium hydride was added to deprotonate the guanidine. The light orange solution was stirred at 65 °C for 18 h. TLC was used to monitor reaction progress (9 DCM:1 MeOH, Rf 0.3). The reaction was quenched with 6 mL methanol, stirred for 30 min, diluted with 50 mL DCM, and washed with 50 mL saturated brine to obtain a white precipitate which was removed by filtration. The filtrate was transferred to a separatory funnel, the organic layer was removed, and the aqueous layer was washed with 10 mL DCM. The organic layers were concentrated *in vacuo* to obtain a yellow solid which was purified using 0.5" x 6" silica gel column, wet loading the product dissolved in DCM, in gradient of 95 DCM:5 MeOH to 70 DCM:30 MeOH. The fractions containing product were combined and concentrated *in vacuo*. The resulting pink solid was dissolved in methanol and cooled to 0 °C. The colorless solution was stirred on ice while 1.0 mL 2 M HCl in ether was added dropwise. The

resulting yellow suspension was stirred for 20 min and triturated with cold ether (3x10 mL). The white precipitate was dried *in vacuo* to yield 9.69 mg (12.7% over 2 steps) of the title compound as a white solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.60 (s, 1H), 8.58 (bs, 2H), 8.31 (bs, 2H), 8.01 (t, J = 5.82, 1H), 7.54 (bs, 2H), 3.42 (q, J = 6.67, 2H), 3.38 (t, J = 6.74, 2H), 1.64 (m, 2H), 1.57 (m, 2H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 165.70, 156.26, 155.47, 152.68, 120.62, 108.63, 50.81, 40.60, 26.16, 25.98. ^{1}H NMR (500 MHz, D_2O) δ 3.43 (t, J = 6.70, 2H), 3.31 (t, J = 6.45, 2H), 1.63 (m, 4H). ^{13}C NMR (500 MHz, D_2O) δ 165.55, 155.88, 154.88, 152.72, 122.03, 108.34, 50.82, 40.44, 25.41, 25.37. LR-ESI MS (*m/z*) calcd for [M] 327.12; found 327.1. 95.5% pure by HPLC. Soluble in DMSO and D_2O, stable by NMR over 7 d in D_2O.



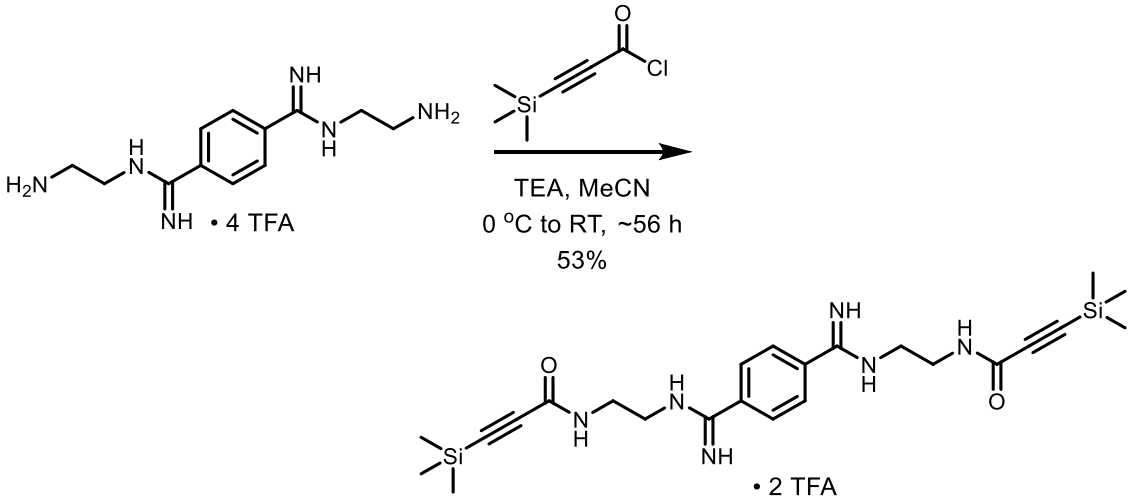
Compound 5a: N,N'-(((((((1,4-phenylenebis(iminomethylene))bis(azanediyl))bis(butane-4,1-diyl))bis(azanediyl))bis(6-amino-1,3,5-triazine-4,2-diyl))bis(azanediyl))bis(ethane-2,1-diyl))bis(3-(trimethylsilyl)propiolamide). To a 20 mL vial was added 19.99 mg (24.1 µmol) N^1,N^4-bis(2-aminoethyl)terephthalimidamide (TFA salt) and 2.5 mL acetonitrile. The resulting white suspension was cooled to 0 °C. To the vial was added 27 µL (0.2 mmol) triethylamine and the suspension was stirred for an additional 10 min. In a separate vessel was mixed 16.96 mg (106 µmol) 3-(trimethylsilyl)propioloyl chloride and 1 mL acetonitrile and the resulting clear solution

was added to the 20 mL vial. The resulting brown solution was warmed to room temperature slowly and stirred for about 56 h under nitrogen atmosphere. TLC was used to monitor reaction progress (9 MeOH:1 NH_4OH, potassium permanganate stain, Rf 0.05). The solvents were removed *in vacuo* to obtain a brown solid that was purified via preparative HPLC (gradient H_2O (0.1% TFA):MeCN (0.1% TFA) from 100:0 to 0:100) to obtain 22.0 mg (69%) of the title compound as a tan, crystalline solid TFA salt. ^{1}H NMR (600 MHz, DMSO-d_6) δ 9.54 (bs, 2H), 9.13 (bs, 2H), 8.73 (bs, 2H), 7.82 (bs, 1H), 7.67 (bs, 1H), 7.47 (bm, 2H), 7.41 (bs, 4H), 7.35 (s, 4H), 2.99 (bm, 2H), 2.87 (bm, 8H), 2.75 (bm, 2H), 2.42 (bm, 4H), 1.06 (bm, 8H), 0.59 (m, 18H). LR-ESI MS (*m/z*) calcd for $C_{24}H_{38}N_6O_2Si_2^{2+}$ [M+H] 857.5; found 857.2.



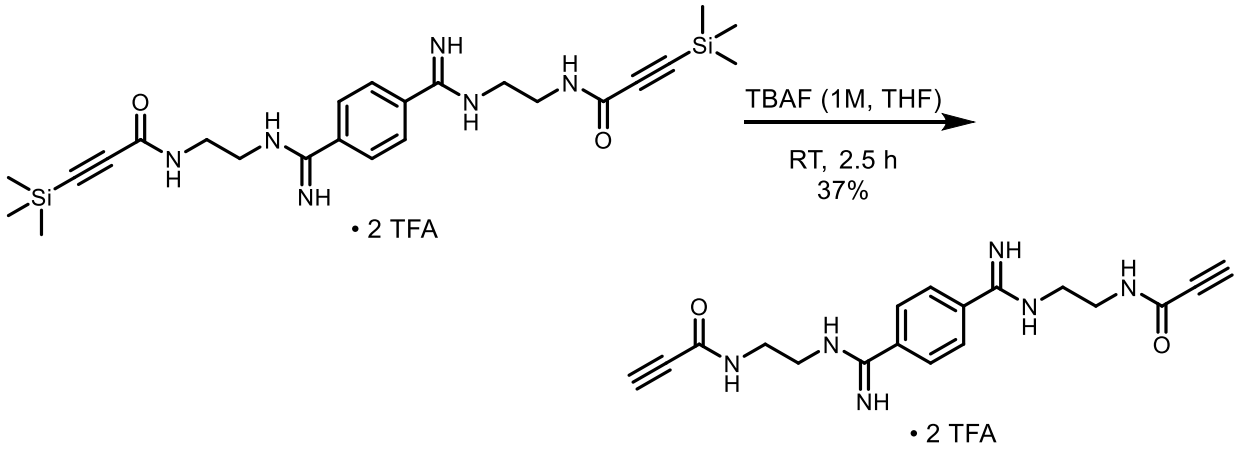
Compound 5: N,N'-(((((((1,4-phenylenebis(iminomethylene))bis(azanediyl))bis(butane-4,1-diyl))bis(azanediyl))bis(6-amino-1,3,5-triazine-4,2-diyl))bis(azanediyl))bis(ethane-2,1-diyl))bis(3-(trimethylsilyl)propiolamide). To a 20 mL vial was added 22.0 mg (16.8 µmol) compound **5a** (TFA salt) and 140 µL (140 µmol) 1 M tetrabutylammonium fluoride in THF and 1.00 mL DI water. The resulting clear solution was stirred at RT for 3 h and dried *in vacuo* to obtain a white solid that was purified via preparative HPLC (gradient H_2O (0.1% TFA):MeCN (0.1% TFA) from 100:0 to 0:100) to yield 19.6 mg (73%) of the title compound as a tan solid TFA salt. ^{1}H NMR (600 MHz, d_2O) δ 7.81 (s, 4H), 3.69 (bm, 4H), 3.50 (bm, 8H), 3.40 (bs, 2H), 3.21

(bm, 4H), 1.77 (bm, 4H), 1.70 (bm, 4H).^{13}C NMR (600 MHz, d_2O) δ 161.80, 161.03, 160.90*, 160.66*, 160.45*, 160.43*, 160.19*, 158.74, 156.02, 153.99, 131.02, 126.28, 116.93*, 115.00*, 113.07*, 111.13*, 52.79, 44.41, 40.43, 37.93, 36.76, 35.88, 23.43, 21.79. *TFA peaks. LR-ESI MS (*m/z*) calcd for $C_{32}H_{46}N_{18}O_2^{2+}$ [M+2H/2] 357.2; found 357.6. 99% pure by analytical HPLC (retention time 3.693 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.)

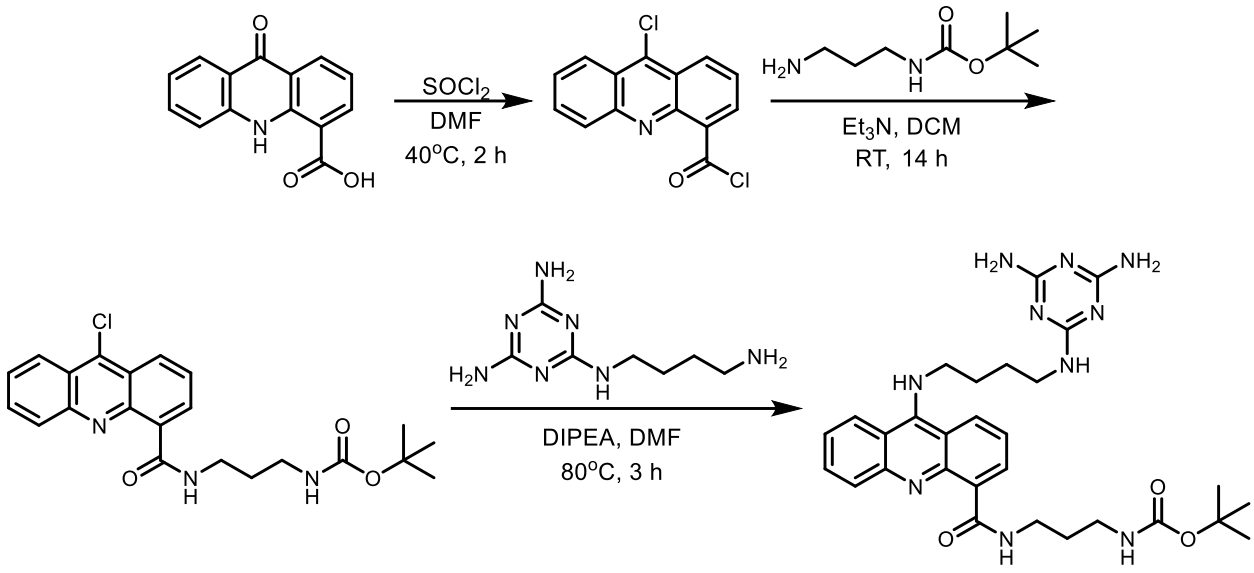


N,N'-(((1,4-phenylenebis(iminomethylene))bis(azanediyl))bis(ethane-2,1-diyl))bis(3-(trimethylsilyl)propiolamide). To a 20 mL vial was added 19.98 mg (28.4 µmol) N^1,N^4-bis(2-aminoethyl)terephthalimidamide (TFA salt) and 1.5 mL acetonitrile. The resulting white suspension was cooled to 0 °C. To the vial was added 28 µL (0.2 mmol) triethylamine and the suspension was stirred for an additional 10 min. In a separate vessel was mixed 18.21 mg (113 µmol) 3-(trimethylsilyl)propioloyl chloride and 1 mL acetonitrile and the resulting clear solution was added to the 20 mL vial. The resulting brown solution was warmed to room temperature slowly and stirred for about 56 h under nitrogen atmosphere. TLC was used to monitor reaction progress (9 MeOH:1 NH_4OH, potassium permanganate stain, Rf 0.05). The solvents were removed *in vacuo* to obtain a brown solid that was purified via preparative HPLC (gradient H_2O (0.1%

TFA):MeCN (0.1% TFA) from 100:0 to 0:100) to obtain 10.5 mg (53%) of the title compound as a white, crystalline solid TFA salt. ^{1}H NMR (600 MHz, DMSO-d_6) δ 10.22 (bs, 2H), 9.96 (bs, 2H), 9.60 (bs, 2H), 8.26 (bs, 2H), 8.08 (s, 4H), 3.74 (bm, 4H), 3.20 (bm, 4H), 1.24 (s, 8H)*. *Partially deprotected product leads to smaller trimethylsilyl peak integration. LR-ESI MS (*m/z*) calcd for $C_{24}H_{38}N_6O_2Si_2^{2+}$ [M+2H/2] 249.13; found 249.1.

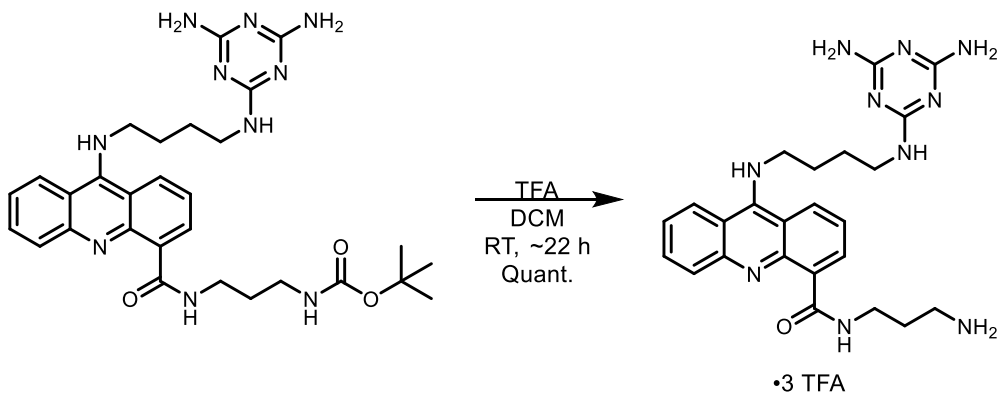


Compound 9: N,N'-(((1,4-phenylenebis(iminomethylene))bis(azanediyl))bis(ethane-2,1-diyl))dipropiolamide). To a 20 mL vial was added 10.5 mg (14.5 µmol) N,N'-(((1,4-phenylenebis(iminomethylene))bis(azanediyl))bis(ethane-2,1-diyl))bis(3-(trimethylsilyl)propiolamide) (TFA salt) and 160 µL (160 µmol) 1 M tetrabutylammonium fluoride in THF and 1.00 mL DI water. The resulting clear solution was stirred at RT for 2.5 h and dried *in vacuo* to obtain a white solid that was purified via preparative HPLC (gradient H_2O (0.1% TFA):MeCN (0.1% TFA) from 100:0 to 0:100) to yield 3.13 mg (37%) of the title compound as a white solid TFA salt. ^{1}H NMR (600 MHz, d_2O) δ 7.89 (s, 4H), 4.03 (bm, 4H), 3.74 (bm, 4H), 3.30 (s, 2H) .^{13}C NMR (600 MHz, d_2O) δ 165.51, 164.30, 163.16*,162.92*, 162.69*, 162.45*, 133.07, 127.44, 119.14*, 117.21*, 115.28*, 113.34*, 44.79, 40.18, 39.28, 36.77. *TFA peaks. LR-ESI MS (*m/z*) calcd for $C_{18}H_{21}N_6O_2^+$ [M+H] 353.41; found 353.6.

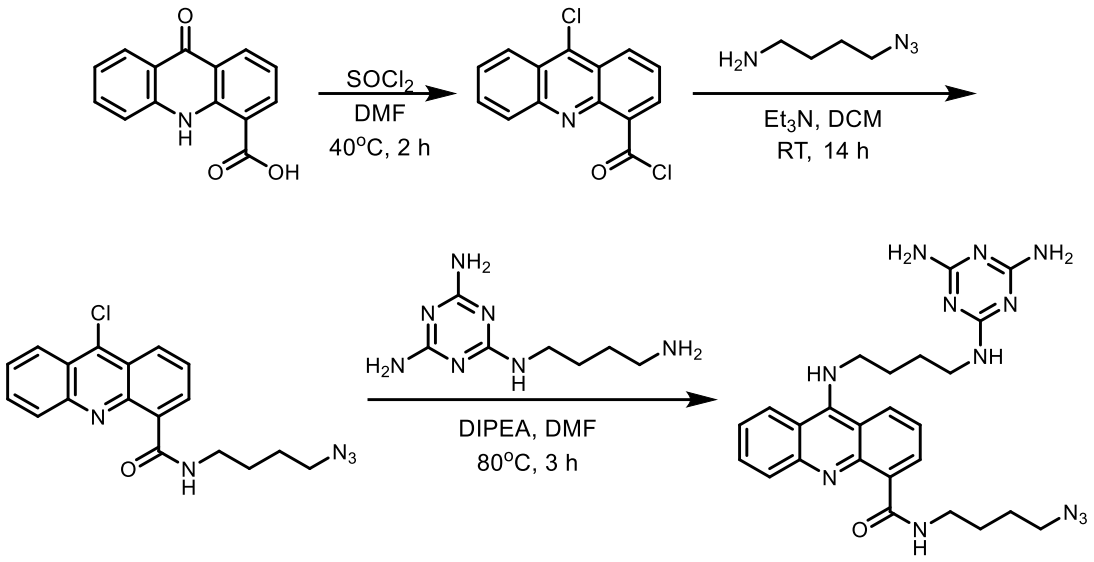


***tert*-butyl (3-(9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-4-carboxamido)propyl)carbamate.** All reactions were run under dry nitrogen and with oven-dried glassware. To an oven-dried 65 mL round-bottom flask was added 0.883 g (3.7 mmol) 9-oxo-9,10-dihydroacridine-4-carboxylic acid (previously dried on Hi-Vac and then lyophilizer) and 6.0 mL freshly distilled thionyl chloride. The resulting yellow suspension was heated to 70 °C and 2 drops anhydrous DMF were added. After about 30 min the solution became homogenous and brownish orange in color. The reaction was stirred at 70 °C for 30-60 min. TLC was used to monitor reaction progress (9 DCM:1 MeOH, Rf 0.9). Reaction times longer than 60 min often resulted in decomposition. The resulting orange solution was cooled to room temperature, 10 mL anhydrous DCM was added, and the solution was distilled to azeotrope the thionyl chloride. Addition of 10 mL DCM and distillation was repeated 2 additional times. The product was dried *in vacuo* for a minimum of 2 h and used directly in the next reaction. The yellow solid product was dissolved in 20 mL DCM and cooled to 0 °C on ice. After stirring the yellow suspension for ~5 min, anhydrous triethylamine was added dropwise until the pH was about 11 (about 18 mL) and 0.80 mL (4.6 mmol) tert-butyl (3-aminopropyl)carbamate was added. Reaction progress was monitored by TLC

(9 DCM:1 MeOH, Rf 0.8) and after 12-24 h, the resulting brown solution was dried *in vacuo* to yield a brown to light tan solid that was further dried *in vacuo*. Half of the dry brown solid (1.53 g, crude) was utilized directly in the next reaction. In a 65 mL, oven-dried RBF was mixed the brown solid and 0.874 g (4.4 mmol) N^2-(4-aminobutyl)-1,3,5-triazine-2,4,6-triamine and 20 mL dry DMF. To the resulting suspension was added 0.77 mL *N*,*N*-diisopropylethylamine dropwise. The light brown suspension was stirred at 80 °C for about 4 h. TLC was used to monitor reaction progress (9 DCM:1 MeOH, Rf 0.3). The resulting brown solution was dried *in vacuo* to obtain a brown oil that was purified using 1.5" x 6" silica gel column with gradient of 9 DCM:1 MeOH:0 NH_4OH to 70 DCM:30 MeOH:5 NH_4OH. The fractions containing product were combined and concentrated *in vacuo* to yield 0.655 g (31% over three steps) of the title compound as an orange solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.68 (d, *J* = 8.55, 1H), 8.52 (d, *J* = 8.63, 1H), 8.45 (d, *J* = 7.37, 1H), 8.00 (t, *J* = 7.64, 1H), 7.93 (d, *J* = 8.46, 1H), 7.61 (q, *J* = 8.04, 2H), 4.22 (brt, *J* = 7.18, 2H), 3.58 (brt, *J* = 6.73, 2H), 3.50 (d, *J* = 7.27, 1H), 3.41 (brt, *J* = 6.63, 2H), 3.37 (s, 1H), 3.22 (brt, *J* = 6.22, 2H), 2.72 (s, 1H), 2.06 (m, 2H), 1.87 (m, 2H), 1.76 (m, 2H), 1.46 (s, 9H), 1.38 (t, *J* = 7.26, 1H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 167.68, 167.09, 166.09, 166.01, 157.16, 153.49, 147.96, 146.99, 140.97, 133.78, 130.66, 128.13, 123.22, 122.71, 120.53, 116.05, 115.28, 107.91, 78.52, 49.80, 48.47, 39.54, 37.94, 36.84, 29.61, 29.34, 27.96, 27.41, 26.65. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 575.31; found 575.3.

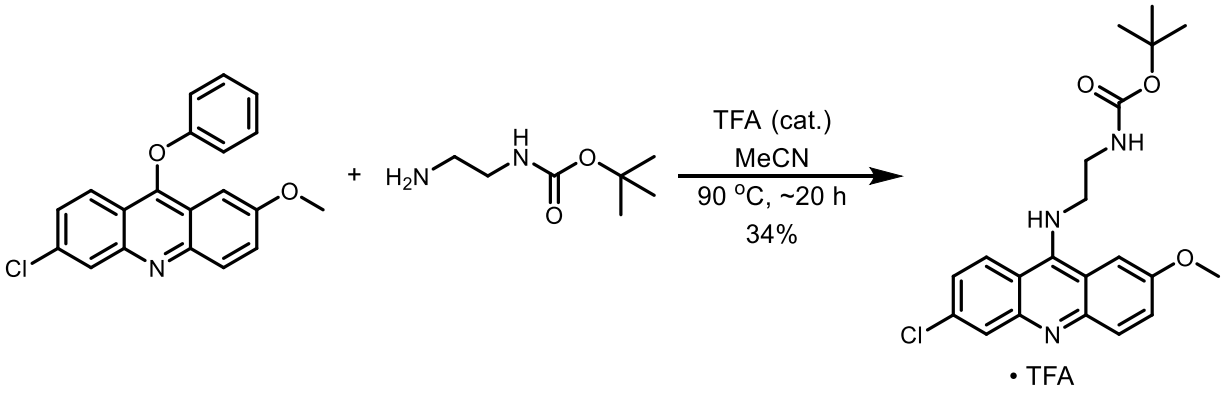


***N*-(3-aminopropyl)-9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-4-carboxamide.** In a 100 mL round bottom flask was mixed 655 mg (1.14 mmol) *tert*-butyl (3-(9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-4-carboxamido)propyl)carbamate, 25 mL trifluoroacetic acid, and 25 mL DCM. The mixture was stirred at room temperature under nitrogen atmosphere for about 22 h. TLC (8 MeOH:2 NH_4OH, Rf pdt = 0.2) was used to monitor reaction progress. The solvents were removed *in vacuo* and the resulting black oil was dissolved in 20 mL MeOH and dried *in vacuo* to yield the product as a fine yellow solid. ^{1}H NMR (600 MHz, DMSO-d_6) δ 13.68 (s, 1H), 10.03 (bs, 1H), 9.36 (bs, 1H), 8.75-8.38 (bm, 1H), .8.48 (bs, 1H), 8.26-8.09 (bm, 2H), 8.07-7.96 (bm, 2H), 7.86 (bs, 2H), 7.80-7.56 (bm, 4H), 4.15-4.07 (bm, 2H), 3.19-3.16 (bm, 2H), 3.12-3.04 (bm, 2H), 2.97-2.87 (bm, 2H), 1.96-1.88 (bm, 2H), 1.70-1.60 (bm, 2H), 1.55-1.40 (bm, 2H), 1.26 (bs, 2H). LR-ESI MS (*m/z*) calcd for [M+H$^+$] 475.27; found 475.1.



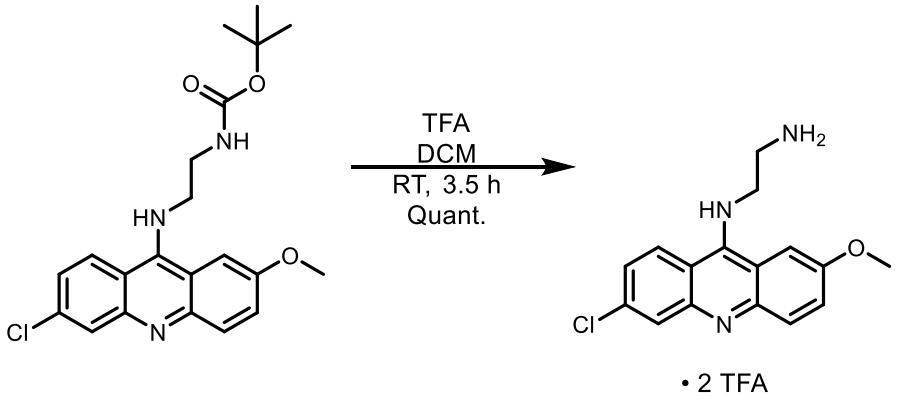
***N*-(4-azidobutyl)-9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-4-carboxamide.** All reactions were run under dry nitrogen and with oven-dried glassware. To an oven-dried 100 mL round-bottom flask was added 2.16 g (9.0 mmol) 9-oxo-9,10-dihydroacridine-4-carboxylic acid (previously dried on Hi-Vac and then lyophilizer) and 75.0 mL freshly distilled thionyl chloride. To the resulting yellow suspension was added 6 drops anhydrous DMF and heated to 70 °C with stirring. After about 50 min the solution became homogenous and brownish orange in color. The reaction was stirred at 70 °C for 30-60 min and stirred at room temperature to obtain a yellow precipitate. TLC was used to monitor reaction progress (9 DCM:1 MeOH, Rf 0.9). Reaction times longer than 60 min often resulted in decomposition. The suspension was transferred to a 300 mL RBF after cooling to room temperature. To the flask was added 50 mL anhydrous DCM and the solution was distilled to azeotrope the thionyl chloride. Addition of 50 mL DCM and distillation was repeated 2 additional times. The product was dried *in vacuo* for a minimum of 2 h and used directly in the next reaction. The yellow solid product was dissolved in 50 mL DCM and cooled to 0 °C on ice. After stirring the yellow suspension for ~5 min, anhydrous triethylamine was added dropwise until the pH was about 11 (about 60 mL) and 1.29 g (11.3 mmol) 4-

azidobutan-1-amine was added. Reaction progress was monitored by TLC (9 DCM:1 MeOH, Rf 0.8) and after 12-24 h, the resulting brown solution was dried *in vacuo* to yield a brown to light tan solid that was utilized directly in the next reaction. In a 100 mL, oven-dried RBF was mixed the brown solid and 1.0 g (5.1 mmol) N^2-(4-aminobutyl)-1,3,5-triazine-2,4,6-triamine and 160 mL dry DMF. To the resulting suspension was added 1.9 mL *N*,*N*-diisopropylethylamine dropwise. The light brown suspension was stirred at 80 °C for about 3 h. TLC was used to monitor reaction progress (9 DCM:1 MeOH, Rf 0.9). The resulting brown solution was dried *in vacuo* to obtain a brown oil that was purified on a Combiflash silica column (12 g) with gradient of 100 DCM:0 MeOH to 70 DCM:30 MeOH. The fractions containing product were combined and concentrated *in vacuo* to yield 0.637 g (14% over three steps) of the title compound as a black solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.78-8.55 (m, 1H), 8.55-8.39 (m, 1H), 8.33-8.13 (m, 1H), 8.08-7.99 (m, 1H), 7.96-7.82 (m, 2H), 7.80-7.72 (m, 1H), 7.53 (t, *J* = 7.73, 1H), 7.39-7.30 (m, 1H), 7.29-7.06 (m, 2H), 3.30-3.23 (m, 4H), 2.80 (s, 2H), 1.77-1.47 (m, 6H), 1.29 (dd, *J* = 14.50, 6.72, 2H), 1.23 (d, *J* = 6.28, 2H).*Amine protons exchnaged. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 515.27; found 515.1.



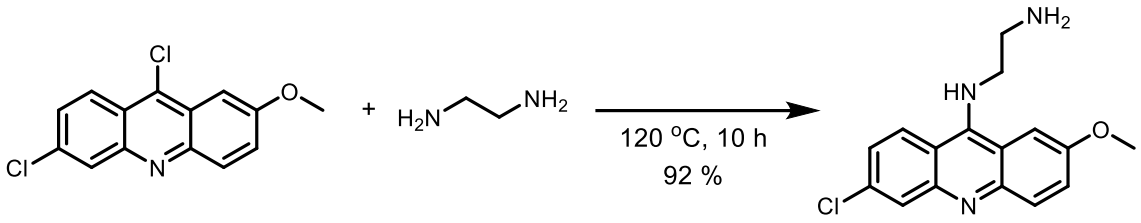
***tert*-butyl (2-((6-chloro-2-methoxyacridin-9-yl)amino)ethyl)carbamate.** To an oven-dried 50 mL round-bottomed flask was added 285.6 mg (0.85 mmol) *tert*-butyl(2-aminoethyl)carbamate and 10 mL dry acetonitrile. The flask was placed under nitrogen in an oil bath at 90 °C. To the

flask was added 147.09 mg (0.92 mmol) 6-chloro-2-methoxy-9-phenoxyacridine dissolved in 3.4 mL acetonitrile. The yellow suspension was heated to reflux and 10 drops of trifluoroacetic acid were added to the flask. The resulting bright orange solution was stirred at 90 °C for about 20 h, cooled to room temperature to form an orange suspension that was filtered with a Buchner funnel. The orange solid was washed with 2x20 mL cold acetonitrile and 2x20 mL cold diethyl ether to yield 154 mg (34%) of the title compound as a bright yellow solid TFA salt. ^{1}H NMR (500 MHz, $CDCl_3$) δ 12.87 (bs, 1H)*, 8.03 (bs, 1H), 7.40 (bs, 1H), 7.13-6.96 (bm, 1H), 6.84 (bs, 1H), 6.79-6.75 (bm, 2H), 6.59 (bs, 1H), 6.24 (bs, 1H), 3.28 (bs, 2H), 3.12 (bs, 2H), 2.69 (s, 3H), 1.66 (s, 9H). *protonated acridine. LR-ESI MS (*m/z*) calcd for [M+H^{+}] 402.16; found 402.2.

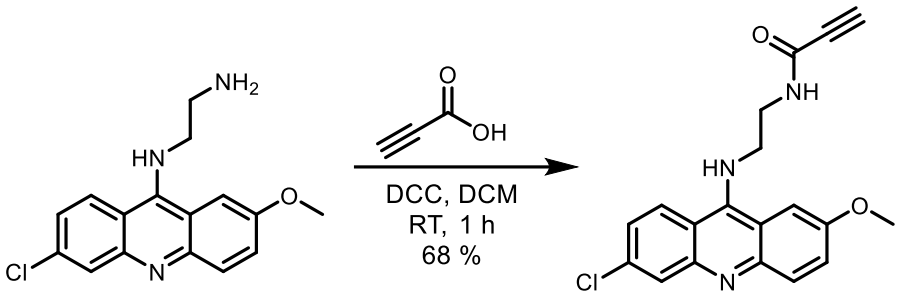


***N*^{1}-(6-chloro-2-methoxyacridin-9-yl)ethane-1,2-diamine.** To a 20 mL vial was added 152 mg (0.38 mmol) of *tert*-butyl (2-((6-chloro-2-methoxyacridin-9-yl)amino)ethyl)carbamate and 2.4 mL DCM. To the resulting yellow suspension was added 2.9 mL trifluoroacetic acid and the resulting orange solution was stirred at room temperature for ~3.5 h. Solvents were removed *in vacuo* to obtain 201 mg (quant.) of the title compound as a yellow solid TFA salt. ^{1}H NMR (500 MHz, $CDCl_3$) δ 13.77 (s, 1H)*, 8.87 (s, 1H), 8.18 (bd, *J* = 9.10, 1H), 7.65 (bs, *J* = 4.85, 1H), 7.54 (bdt, *J* = 9.25, 4.08, 1H), 7.47 (dd, *J* = 9.17, 2.51, 1H), 7.35 (d, *J* = 9.12, 1H), 7.03 (s, 1H), 4.47 (s, 3H), 3.46 (bs, 4H). *protonated acridine. Primary amine protons seem to be exchanged ^{13}C NMR (500

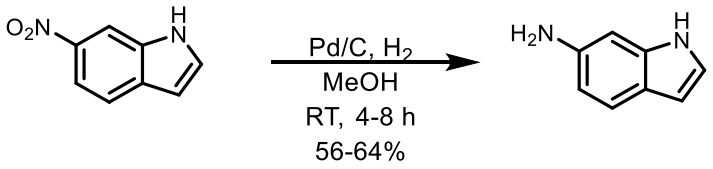
MHz, DMSO-d_6) δ 159.13*, 158.85*, 158.57*, 158.29*, 157.56, 156.37, 139.72, 128.08, 124.12, 121.20, 121.14, 119.97*, 118.70, 117.86*, 117.64, 115.30,* 115.20, 115.16, 112.97*, 110.90, 110.86, 56.18, 50.49, 49.06. *TFA peaks. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 302.78; found 302.0.



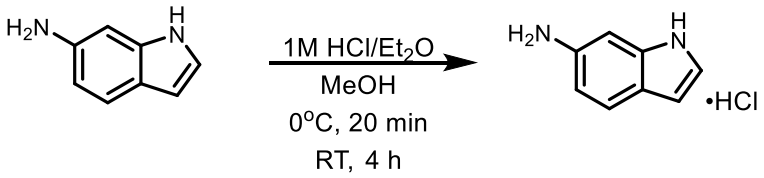
***N*^{1}-(6-chloro-2-methoxyacridin-9-yl)ethane-1,2-diamine.** To a 300 mL round bottomed flask was added 6.03 g (22 mmol) 6,9-dicholoro-2-methoxyacridine and 117 mL (1.75 mol) 1,2-diaminoethane. The yellow suspension was stirred in an oil bath at 120 °C under nitrogen. The yellow suspension transitioned to an orange solution and then a brown solution that was cooled to room temperature after about 10 h. To the flask was added 280 mL DI water to obtain a yellow precipitate in an orange solution. Filtered with a Buchner funnel and washed with 20x6 mL DI water. The solid was suspended in 400 mL diethyl ether, sonicated, filtered again, and washed with ~1200 mL diethyl ether. The solid was dried *in vacuo* to yield 6.00 g (92%) of the title compound as an orange solid. 99% pure by analytical HPLC (retention time 4.647 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.) ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.37 (d, *J* = 9.35, 1H), 7.91-7.76 (bm, 2H), 7.64 (d, *J* = 2.71, 1H), 7.42 (dd, *J* = 9.30, 2.68, 1H), 7.33 (d, *J* = 9.40, 1H), 3.94 (s, 3H), 3.72 (t, *J* = 6.36, 2H), 2.88 (t, *J* = 6.35, 2H). Amine protons seem to be exchanged. ^{13}C NMR (500 MHz, DMSO-d_6) δ 155.50, 151.13, 133.89, 131.10, 127.56, 126.97, 124.70, 123.12, 122.52, 117.68, 115.25, 101.37, 101.17, 56.07, 53.11, 42.82. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 302.11; found 302.0.



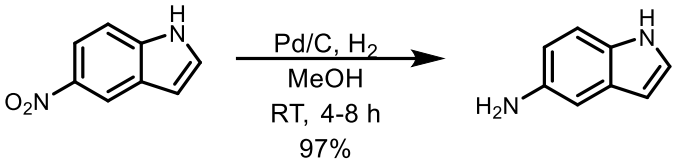
***N*-(2-((6-chloro-2-methoxyacridin-9-yl)amino)ethyl)propiolamide.** To a 65 mL round bottomed flask was added 503 mg (1.7 mmol) N^1-(6-chloro-2-methoxyacridin-9-yl)ethane-1,2-diamine and 22 mL DCM. In a separate vial was mixed 742 mg DCC (3.6 mmol), 0.27 mL (4.3 mmol) propiolic acid, and 7.0 mL DCM. The resulting white solid was filtered off with a Buchner funnel and the brown filtrate was added to the round bottom flask at once. The vial and Buchner funnel were washed with 3 mL DCM and this was added to the flask. The resulting orange suspension was stirred at room temperature for about 1 h. TLC (8 DCM:2 MeOH, Rf pdt = 0.6) was utilized to monitor reaction progress. When no more starting material was observed, the solvents were removed *in vacuo* to afford a yellow solid. The crude product was packed on silica and purified using a Combi-Flash RediSep column (silica, 12 g) with elution gradient 100 DCM:0 MeOH to 80 DCM:20 MeOH. The fractions containing product were combined and concentrated *in vacuo* to yield 395 mg (68%) of the title compound as an orange solid. ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.04 (t, *J* = 5.79, 1H), 8.43 (d, *J* = 9.31, 1H), 7.93 (d, *J* = 2.30, 1H), 7.87 (d, *J* = 9.29, 1H), 7.71 (d, *J* = 2.69, 1H), 7.55 (dd, *J* = 9.35, 2.55, 1H), 7.40 (dd, *J* = 9.30, 2.18, 1H), 4.18 (s, 1H), 4.00 (td, *J* = 7.67, 3.15, 2H), 3.97 (s, 3H), 3.55 (q, *J* = 6.27, 2H), 3.17 (s, 1H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 155.20, 154.54, 153.40, 152.04, 143.32, 139.90, 136.28, 127.28, 125.21, 122.65, 112.68, 111.48, 111.29, 102.35, 77.61, 76.02, 55.52, 48.26, 48.05. HR-ESI MS (*m/z*) calcd for [M+H$^+$] 354.10038; found 354.1016.



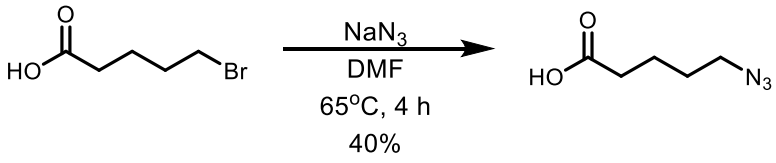
6-aminoindole.[11] To an oven-dried 250 mL round-bottomed flash was added 1.00 g (6.17 mmol) 6-nitroindole. The round-bottomed flask was purged with vacuum and nitrogen and 70 mL methanol was added to obtain a light orange, transparent solution. In a separate vessel was mixed a dime-sized amount of 10 wt. % loading palladium on carbon, 1 mL DCM, and 4 mL methanol. The black palladium on carbon mixture was added to the round-bottomed flask under nitrogen and stirred to obtain a blackish green solution. The nitrogen was turned off and hydrogen gas was bubbled through the reaction until no more starting material was observed via TLC (3 EtOAc:2 Hex, Rf 0.4). The mixture was filtered through celite in methanol and the filtrate was concentrated *in vacuo* to yield 0.523 g (64%) of the title compound as a black solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.42 (s, 1H), 7.15 (d, *J* = 8.28, 1H), 6.94 (dd, *J* = 3.13, 2.17, 1H), 6.54 (brt, *J* = 0.97, 1H), 6.36 (dd, *J* = 8.31, 1.98, 1H), 6.16 (ddd, *J* = 3.06, 1.98, 0.96, 1H), 4.64 (s, 2H).



6-aminoindole hydrochloride. To a 100 mL round-bottom flask was added 0.54 g (4.0 mmol) 6-aminoindole and 4 mL methanol. The mixture was stirred to obtain a dark brown suspension which was cooled to 0 °C and 5.0 mL (5.2 mmol) 1 molar HCl in ether was added dropwise over 20 min. The mixture was slowly warmed to room temperature and stirred for 4 h. The resulting black solution was concentrated *in vacuo* to obtain a brown solid. The product was stored under high vacuum until use in further reactions.

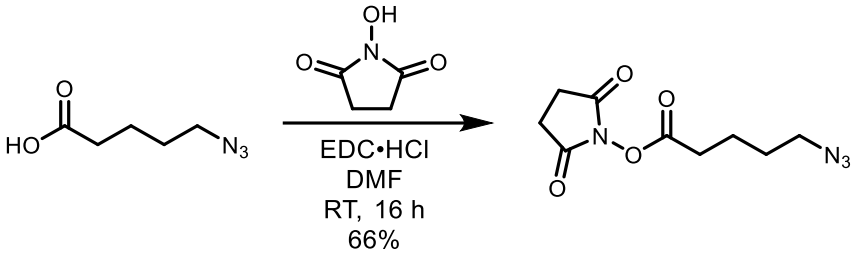


5-aminoindole.[11] To an oven-dried 250 mL round-bottom flask was added 1.00 g (6.17 mmol) 5-nitroindole. The flask was purged under vacuum, filled with nitrogen, and 70 mL dry methanol was added to obtain a light orange transparent solution. In a separate vessel was mixed a dime-sized amount of 10 wt. % loading palladium on carbon, 1 mL DCM, and 4 mL methanol. The black palladium on carbon mixture was added to the round-bottomed flask under nitrogen and stirred to obtain a blackish green solution. The nitrogen was turned off and hydrogen gas was bubbled through the reaction until no more starting material was observed via TLC (3 EtOAc:2 Hex, Rf 0.6). The mixture was filtered through celite in methanol and the filtrate was concentrated *in vacuo* to afford 0.80 g (97%) of the title compound as a brown solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.55 (s, 1H), 7.11 (t, J = 2.66, 1H), 7.07 (d, J = 8.49, 1H), 6.68 (d, J = 1.51, 1H), 6.48 (dd, J = 8.49, 2.01, 1H), 6.16 (bs, 1H), 4.39 (s, 2H), 1.38 (s, 1H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 141.46, 130.19, 128.98, 125.14, 112.25, 111.80, 103.63, 100.05. HR-ESI MS (*m/z*) calcd for [M+H^{+}] 133.07; found 133.0764.



5-azidopentanoic acid.[12] To a 25 mL, 2-necked round-bottom flask was added 7.21 g (40.0 mmol) 5-bromovaleric acid and 2.59 g (39.8 mmol) sodium azide. The flask was purged with vacuum and filled with nitrogen and 6.2 mL DMF was added. The resulting light orange suspension was stirred at 65 °C for about 4 h. TLC was used to monitor reaction progress (2 EtOAc:3 Hex, Rf = 0.6). To the suspension was added 30 mL DI water and the resulting solution was transferred to a 125 mL

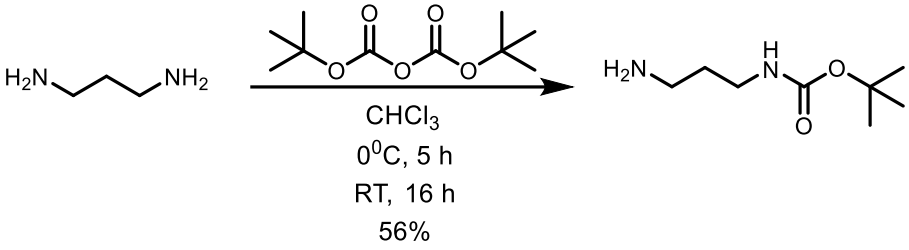
separatory funnel. The product was extracted with ethyl acetate (3 x 30 mL) and the organic layers were combined and washed with 50 mL brine. The organic layer was dried with sodium sulfate and concentrated *in vacuo* to obtain a clear to faint yellow liquid. The crude product was purified by silica gel chromatography (wet-loaded) with a 0.5" x 12" column and a gradient of 95 Hex:5 EtOAc to 80 Hex:20 EtOAc. The fractions containing the desired product were combined and concentrated *in vacuo* to afford 2.25 g (40%) of the title compound as a transparent oil. ^{1}H NMR (500 MHz, $CDCl_3$) δ 11.69 (bs, 1H), 3.31 (t, *J* = 6.59, 2H), 2.41 (t, *J* = 7.14, 2H), 1.72 (m, 2H), 1.66 (m, 2H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 179.80, 51.07, 33.46, 28.22, 21.83. LR-ESI MS (*m/z*) calcd for [M] 143.07, found 142.2.



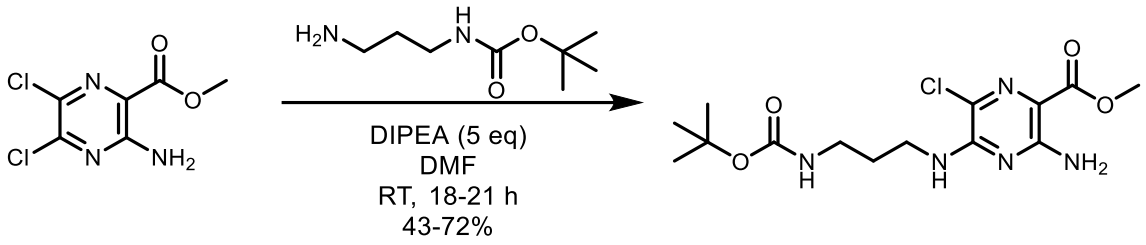
2,5-dioxopyrrolidin-1-yl 5-azidopentanoate.[13] To a 15 mL round-bottom flask was added 0.25 g (1.8 mmol) 5-azidopenatnoic acid and 3.5 mL anhydrous DMF. To the colorless solution was added 0.50 g (2.6 mmol) EDC·HCl and 0.24 g (2.1 mmol) N-hydroxy succinimide. The mixture was placed under nitrogen atmosphere and stirred at room temperature for about 16 h. TLC (3 EtOAc:2 Hex, Rf 0.4) was used to monitor reaction progress. The mixture was diluted with 13 mL ethyl acetate and decanted into a 125 mL separatory funnel to obtain a cloudy white solution which was washed with 13 mL saturated ammonium chloride (aqueous). The clear organic layer was washed with 13 mL DI water and 13 mL saturated brine (aqueous), dried over sodium sulfate, and concentrated *in vacuo* to yield 0.27 g (66%) o the title compound as a transparent to slightly pink oil. Product contains DMF and ethyl acetate. ^{1}H NMR (500 MHz, $CDCl_3$) δ 3.34 (t, *J* = 6.62, 2H),

2.84 (bs, 4H), 2.66 (t, J = 7.23, 2H), 1.85 (p, J = 7.26, 2H), 1.72 (dt, J = 13.53, 6.72, 2H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 169.11, 168.19, 50.83, 30.43, 27.89, 25.60, 21.85. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 241.09; found 241.0. Acylation attempted to form **S1** and product was purified, but it was not acylated at indole nitrogen as expected. A good way to determine which NH_2 was acylated would be with NOE experiment, but the sample is not concentrated enough, and all product was used in the NMR sample.

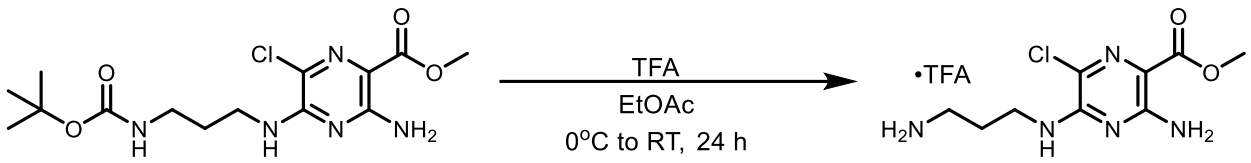
Amiloride derivatization: Adaptations from previously reported intermediates



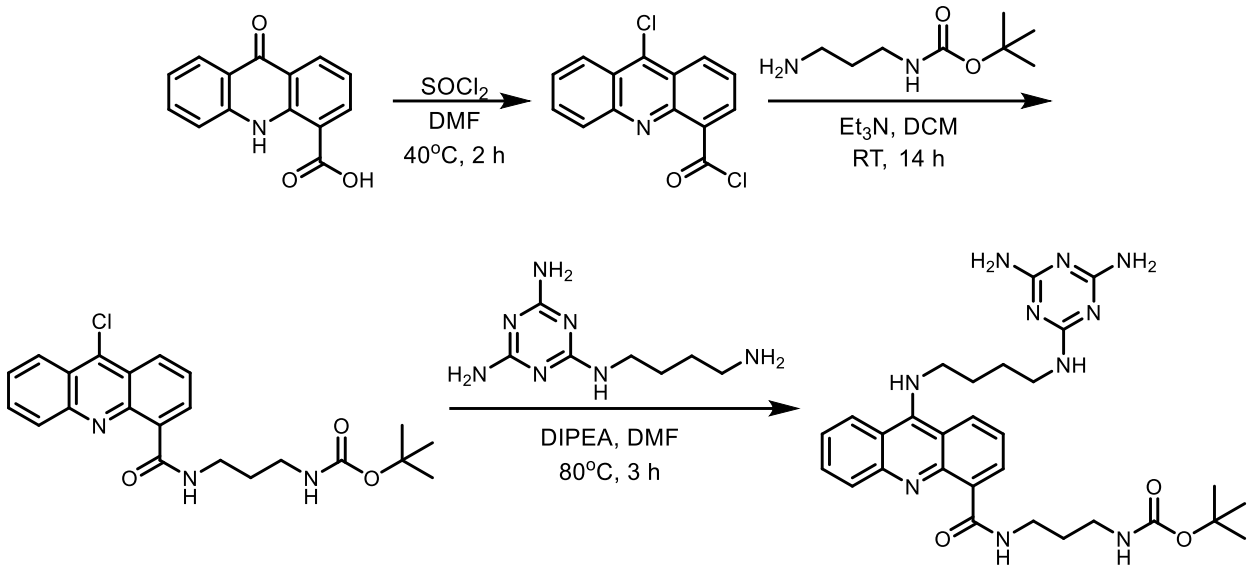
tert-butyl (3-aminopropyl)carbamate.[14] To a 500 mL round-bottom flask was added 19.3 mL (229 mmol) propane-1,3-diamine and 100 mL chloroform and the mixture was cooled to 0 °C on ice. A mixture of 5.05 g (23.1 mmol) di-tert-butyl dicarbonate in 250 mL chloroform was added dropwise via an addition funnel to the 500 mL round-bottom flask on ice over 5 h and the reaction was slowly warmed to room temperature. The resulting white, translucent mixture was stirred at room temperature for about 16 h. Reaction progress was monitored by TLC (9 MeOH:1 NH_4OH, Rf = 0.5). The mixture was transferred to a 500 mL separatory funnel and washed with DI water (6 x 150 mL). The combined organic layers (clear solution) were dried over sodium sulfate and concentrated *in vacuo* to afford 2.25 g (56%) of the title compound as a clear oil. ^{1}H NMR (500 MHz, DMSO-d_6) δ 6.76 (bs, 1H), 3.29 (bs, 2H), 2.95 (q, J = 6.62, 2H), 2.51 (t, J = 6.65, 2H), 1.42 (p, J = 6.92, 2H), 1.37 (s, 9H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 156.09, 77.74, 39.60, 38.12, 33.91, 28.74. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 175.14; found 175.1.



methyl 3-amino-5-((3-tert-butoxycarbonyl)amino)propyl)amino)-6-chloropyrazine-2-carboxylate. To an oven-dried 15 mL round-bottom flask was added 80.0 mg (0.36 mmol) methyl 3-amino-5,6-dicholoropyrazine-2-carboxylate and 2.5 mL anhydrous DMF. The dark brown solution was placed under nitrogen atmosphere and 0.31 mL *N*,*N*-diisopropylethylamine was added. The mixture was stirred for 5 min and 69.6 mg (0.40 mmol) tert-butyl (3-aminopropyl) carbamate dissolved in 0.70 mL anhydrous DMF was added dropwise to the round-bottom flask over five min. The resulting dark brown solution was stirred at room temperature for about 18 h. TLC was used to monitor reaction progress (5 Hex:1 EtOAc, Rf = 0.55). The mixture was concentrated *in vacuo* to obtain a brown residue, which was dissolved in 50 mL ethyl acetate and transferred to a separatory funnel with 50 mL DI water. The ethyl acetate layer was washed with DI water (3 x 50 mL) and saturated brine (aqueous, 2 x 50 mL). The organic layers were combined, dried over sodium sulfate to obtain a yellow liquid, and concentrated *in vacuo* to afford 93 g (72%) of the title compound as a thin orange film. ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.51 (brt, J =5.81, 1H), 7.24 (bs, 2H), 6.84 (brt, J = 5.91, 1H), 3.74 (s, 3H), 3.37 (q, J = 6.32, 2H), 2.98 (q, J = 6.47, 2H), 1.68 (p, J = 6.72, 2H), 1.39 (s, 9H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 165.98, 155.71, 154.80, 151.37, 119.78, 108.30, 77.60, 51.19, 38.28, 37.09, 28.83, 28.30. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 360.14; found 360.1.



methyl 3-amino-5-((3-aminopropyl)amino)-6-chloropyrazine-2-carboxylate. To an oven-dried 10 mL round-bottom flask was added 93 mg (0.26 mmol) methyl-3-amino-5-((3-((tert-butoxycarbonyl)amino)propyl)amino)-6-chloropyrazine-2-carboxylate and 1.5 mL anhydrous ethyl acetate. The resulting light orange solution was placed under nitrogen and cooled to 0 °C on ice. To the round-bottom flask was added 0.70 mL TFA dropwise while stirring and the mixture was slowly warmed to room temperature. The mixture was stirred for about 24 h. TLC was used to monitor reaction progress (9 MeOH:1 NH_4OH, Rf 0.8). The light orange solution was concentrated *in vacuo* to yield 125 mg of the title compound as a red oil (contains EtOAc). ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.72 (bs, 2H), 7.66 (t, J = 5.87, 1H), 7.26 (bs, 2H), 3.42 (q, J = 6.18, 2H), 2.82 (brq, J = 6.49, 2H), 1.84 (p, J = 6.74, 2H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 166.34, 159.12, 158.84, 158.56, 158.28, 155.99, 151.88, 120.15, 108.98, 51.65, 37.95, 37.26, 27.13. LR-ESI MS (*m/z*) calcd (without TFA) for [M+H$^+$] 260.08; found 260.1.



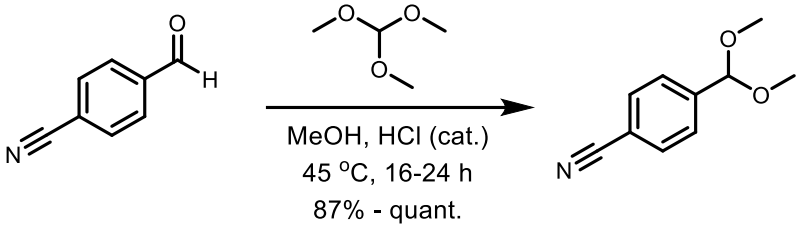
tert-butyl (3-(9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-4-carboxamido)propyl)carbamate. All reactions were run under dry nitrogen and with oven-dried glassware. To an oven-dried 50 mL round-bottom flask was added 0.50 g (2.1 mmol) 9-oxo-9,10-dihydroacridine-4-carboxylic acid (previously dried on Hi-Vac and then lyophilizer) and 5.0 mL freshly distilled thionyl chloride. To the resulting yellow suspension was added 5 drops anhydrous DMF, and after about 10 s the solution became transparent orange in color. The reaction was heated at 80 °C for 30-60 min. TLC was used to monitor reaction progress (9 DCM:1 MeOH, Rf 0.9). Reaction times longer than 60 min often resulted in decomposition. The resulting orange solution was cooled to room temperature, 5 mL anhydrous DCM was added, and the solution was distilled to azeotrope the thionyl chloride. Addition of 5.0 mL DCM and distillation was repeated 2 additional times. The product was dried on Hi-Vacuum for a minimum of 2 h. The yellow solid product was dissolved in 11.5 mL DCM and cooled to 0 °C. After stirring the yellow suspension for 10 min, anhydrous triethylamine was added dropwise until the pH was about 11 and 0.45 mL (2.61 mmol) tert-butyl (3-aminopropyl)carbamate was added. Reaction progress was monitored by TLC (9 DCM:1 MeOH, Rf 0.8) and after 12-14 h, the resulting brown solution was dried *in*

vacuo to yield a brown to light tan solid that was further dried *in vacuo*. To the dry brown solid was added 0.62 g (3.2 mmol) N^2-(4-aminobutyl)-1,3,5-triazine-2,4,6-triamine and 11.3 mL DMF. To the resulting suspension, 0.90 mL *N*,*N*-diisopropylethylamine was added dropwise. The light brown suspension was stirred at 80 °C for about 4 h. TLC was used to monitor reaction progress (9 DCM:1 MeOH, Rf 0.3). The resulting brown solution was dried *in vacuo* to obtain a brown oil that was purified using 1" x 6" silica gel column, wet loading the brown oil, in gradient of 98 DCM:2 MeOH:0 NH_4OH to 70 DCM:30 MeOH:5 NH_4OH. The fractions containing product were combined and concentrated *in vacuo* to yield 0.33 g (27% over three steps) of the title compound as an orange solid. ^{1}H NMR (500 MHz, CD_3OD) δ 8.68 (d, *J* = 8.55, 1H), 8.52 (d, *J* = 8.63, 1H), 8.45 (d, *J* = 7.37, 1H), 8.00 (t, *J* = 7.64, 1H), 7.93 (d, *J* = 8.46, 1H), 7.61 (q, *J* = 8.04, 2H), 4.22 (brt, *J* = 7.18, 2H), 3.58 (brt, *J* = 6.73, 2H), 3.50 (d, *J* = 7.27, 1H), 3.41 (brt, *J* = 6.63, 2H), 3.37 (s, 1H), 3.22 (brt, *J* = 6.22, 2H), 2.72 (s, 1H), 2.06 (m, 2H), 1.87 (m, 2H), 1.76 (m, 2H), 1.46 (s, 9H), 1.38 (t, *J* = 7.26, 1H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 167.68, 167.09, 166.09, 166.01, 157.16, 153.49, 147.96, 146.99, 140.97, 133.78, 130.66, 128.13, 123.22, 122.71, 120.53, 116.05, 115.28, 107.91, 78.52, 49.80, 48.47, 39.54, 37.94, 36.84, 29.61, 29.34, 27.96, 27.41, 26.65. LR-ESI MS (*m/z*) calcd for [$M+H^+$] 575.31; found 575.3.

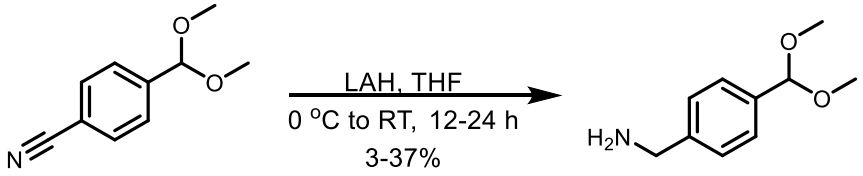
A.3. REVERSIBLE TEMPLATE-ASSISTED LIGATION

A.3.1. Synthetic Methods

The following were synthesized according to previously reported procedures: compound **1**,[15] diethyl terephthalimidate dihydrochloride,[7] and 4,6-dichloro-1,3,5-triazin-2-amine.[6]

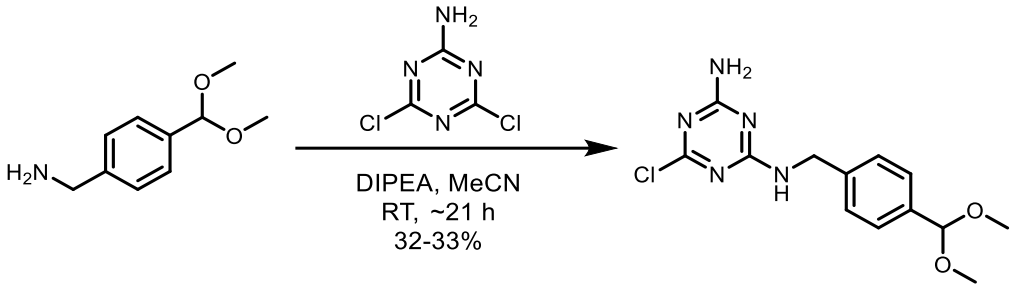


4-(dimethoxymethyl)benzonitrile.[16] To a 250 mL round-bottom flask was added 8.04 g (61.3 mmol) 4-formylbenzonitrile and 106 mL anhydrous methanol. The resulting clear solution was warmed to 45 °C in an oil bath and 40.0 mL (366 mmol) trimethoxymethane was added followed by 0.5 mL conc. HCl was added dropwise. The mixture was stirred for about 16-24 h at 45 °C. TLC (8 pet. ether:2 ethyl acetate, Rf 0.75) was used to monitor reaction progress. The resulting tan solution was cooled to room temperature and dried *in vacuo*. The crude yellow oil was dissolved in 150 mL ethyl acetate and washed with 3x150 mL saturated $NaHCO_3$ and 3x150 mL brine. The organic layer was dried over sodium sulfate and dried *in vacuo* to yield 10.8 g (100%) of the title compound as an orange oil. ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.86 (bd, J = 8.35, 2H), 7.56 (bd, J = 8.32, 2H), 5.47 (s, 1H), 3.26 (s, 6H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 143.85, 132.68, 128.00, 119.09, 111.73, 102.10, 53.25. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 178.09; found 178.1.

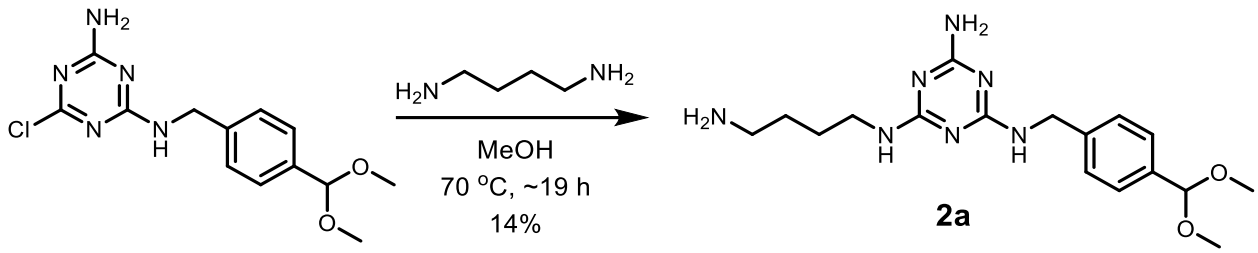


(4-(dimethoxymethyl)phenyl)methanamine.[16] To an oven-dried 300 mL round-bottom flask was added 0.84 g (22 mmol) lithium aluminum hydride and 25 mL anhydrous THF. The resulting grey suspension was cooled on ice and kept under nitrogen. To the suspension was added slowly, dropwise a solution of 1.97 g (11.1 mmol) 4-dimethoxymethyl)benzonitrile dissolved in 2.5 mL anhydrous THF. The resulting green suspension was warmed slowly to RT and stirred for 24 h.

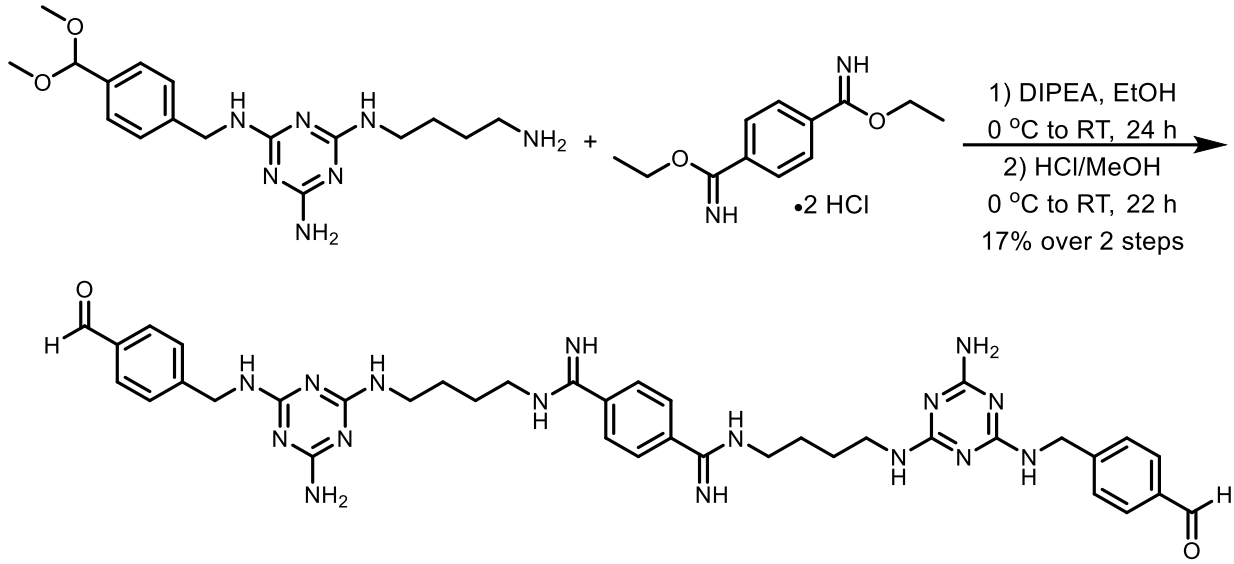
The suspension was cooled on ice and 11 mL of 1M NaOH (aqueous) was added slowly, dropwise and stirred overnight, warming slowly to room temperature. The resulting white suspension was filtered over celite and washed with 2 x 200 mL ethyl acetate. The yellow filtrate was washed with 2 x 75 mL saturated NaHCO3 and 2 x 50 mL brine. The organic layer was dried over sodium sulfate and dried *in vacuo* to obtain 0.76 g (37%) of the title compound as a yellow oil. ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.33 (q, J = 7.55, 4H), 5.36 (s, 1H), 3.73 (s, 2H), 3.23 (s, 6H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 144.73, 136.55, 127.20, 126.75, 103.10, 52.85, 45.80. LR-ESI MS (*m/z*) calcd for [M+H^{+}] 182.12; found 182.1



6-chloro-*N*^{2}-(4-(dimethoxymethyl)benzyl)-1,3,5-triazine-2,4-diamine. To a 200 mL round-bottom flask was added 1.25 g (7.57 mmol) 4,6-dicholoro-1,3,5-triazine-2-amine and 42 mL acetonitrile. The resulting clear solution was placed under nitrogen and a solution of 1.50 g (8.28 mmol) (4-(dimethoxymethyl)phenyl)methanamine dissolved in 5.2 mL acetonitrile was added dropwise. The resulting cloudy white solution was stirred at room temperature for about 21 h. TLC (9 DCM:1 MeOH, Rf 0.7) was used to monitor reaction progress. The white suspension was filtered and washed with 50 mL acetonitrile to obtain 0.710 g (33%) of the title compound as a white solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.32 (bs, 3H), 7.50 (d, J = 7.99, 2H), 7.43 (d, J = 7.90, 2H), 5.41 (s, 1H), 4.40 (s, 1H), 3.25 (s, 6H). ^{13}C NMR (500 MHz, DMSO-d6) δ 193.29*, 141.05, 136.48, 130.13, 129.93, 53.24, 49.07, 42.38. *Melamine peaks seem to be coincident. LR-ESI MS (*m/z*) calcd for [M+H+] 310.11; found 310.1.



Compound 2a. To a 50 mL round-bottom flask was added 0.70 mL (6.9 mmol) 1,4-diaminobutane and 1.0 mL methanol. To the clear solution was added dropwise 0.710 g (2.29 mmol) 6-chloro-N^2-(4-(dimethoxymethyl)benzyl)-1,3,5-triazine-2,4-diamine dissolved in 4.5 mL methanol while stirring. The solution was stirred at 70 °C for about 19 h and cooled to room temperature. TLC (8 DCM:2 MeOH, Rf 0.4) was used to monitor reaction progress. The solvents were removed *in vacuo* to obtain a yellow oil that was further purified on a CombiFlash silica column (24 g) with gradient 95 DCM:5 MeOH to 70 DCM:30 MeOH. The fractions containing product were combined and dried *in vacuo* to yield about 115 mg (14%) of the title compound as a tan crystalline solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.07 (bs, 2H), 7.31 (bm, 4H), 7.22 (bs, 1H), 6.42 (bs, 1H), 5.35 (bs, 1H), 4.45 (bs, 2H), 3.22 (s, 6H), 3.17 (bm, 2H), 2.77 (bm, 2H), 1.51 (bm, 4H). ^{13}C NMR (500 MHz, DMSO-d6) δ 166.06, 136.92, 136.81, 127.51, 127.13, 126.82, 126.77, 103.07, 52.94, 49.05, 43.20 , 38.96, 26.72, 24.89. LR-ESI MS (*m/z*) calcd for [M+H+] 362.23; found 362.23.



Compound 2. To a 35 mL oven-dried round-bottom flask was added 0.049 g (0.167 mmol) diethyl terephthalimidate dihydrochloride and 2.3 mL anhydrous ethanol and the flask was cooled to 0 °C and kept under nitrogen atmosphere. To the white suspension was added 0.74 mL (4.2 mmol) *N*,*N*-diisopropylethylamine to obtain a colorless solution. To the round-bottom flask was added 115 mg (0.317 mmol) N^2-(4-aminobutyl)-N4-(4-(dimethoxymethyl)benzyl)-1,3,5-triazine-2,4,6-triamine dissolved in 2.0 mL anhydrous ethanol. The colorless suspension was stirred at 0 °C for ~2 h, warmed slowly to 25 °C and stirred for about 22 h. TLC was used to monitor progress with ninhydrin stain (8 MeOH:2 NH_4OH, Rf 0.4). The resulting solution was dried *in vacuo* to obtain a white solid that was directly deprotected. To the white solid was added 5 mL methanol to obtain a clear solution that was cooled to 0 °C on ice. To the solution was added 1.0 mL conc. HCl dropwise over 5 min. The resulting clear solution was warmed slowly to room temperature and stirred for about 22 h under nitrogen. The crude product was purified using preparative HPLC (gradient H2O (0.1% TFA):MeCN (0.1% TFA) from 100:0 to 0:100 to yield 20.1 mg (17% over two steps) of the title compound as a white solid. 96% pure by analytical HPLC (retention time 5.022 min, gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100%

over 5 min.). ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.91 (d, J = 10.32, 2H), 9.84 (bs, 1H), 9.79 (bs, 1H), 9.51 (bs, 2H), 9.14-9.05 (bm, 2H), 8.65 (s, 1H), 8.08 (d, J = 8.17, 4H), 7.85-7.75 (m, 8H), 7.46 (d, J = 7.87, 4H), 4.57-4.50 (bm, 4H), 3.84 (s, 4H), 3.26-3.18 (bm, 4H), 1.66-1.54 (bm, 4H), 1.50 (bp, J = 7.21, 2H), 1.43 (bq, J = 7.06, 2H). ^{13}C NMR (500 MHz, DMSO-d6) δ 195.67, 167.86, 163.38, 163.35,* 163.10,* 162.82,* 162.54,* 146.49, 134.61, 133.98, 132.67, 130.34, 130.04, 129.33, 127.79, 127.51, 119.80,* 117.48,* 115.16,* 112.84,* 53.02, 44.03, 42.47, 39.76, 37.85, 25.65, 23.62. COSY NMR (500 MHz, d_2O) Figure A.3. HSQC (500 MHz, d_2O) Figure A.4. HMBC (500 MHz, d_2O) Figure A.5. MALDI MS (m/z) calcd for [M+H^{+}] 759.41; found 759.642.

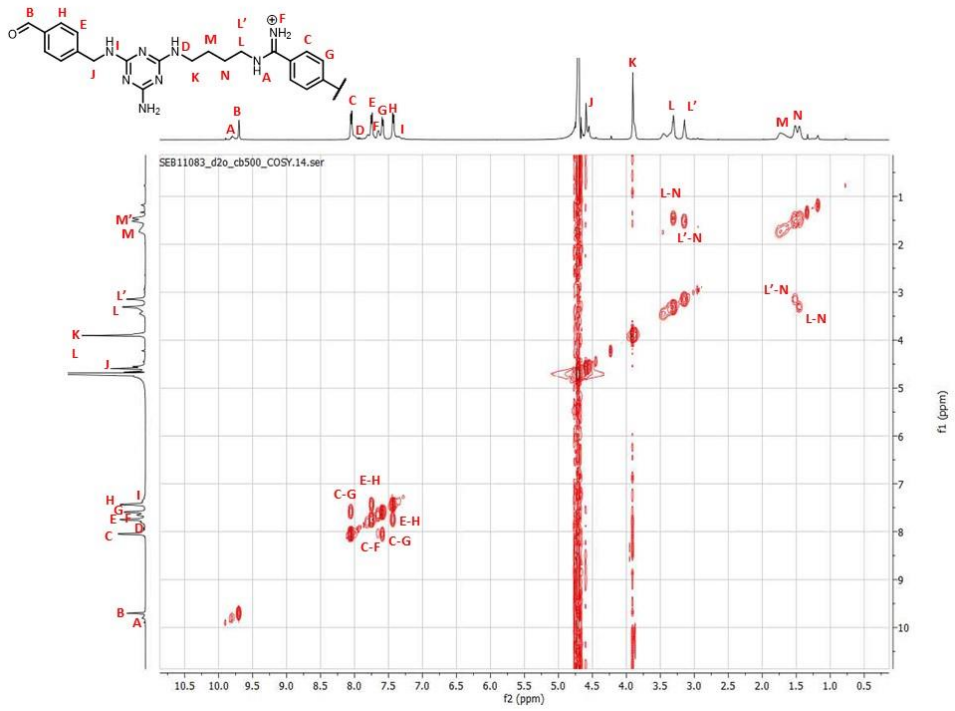


Figure A.3. COSY NMR in d_2O of **2**.

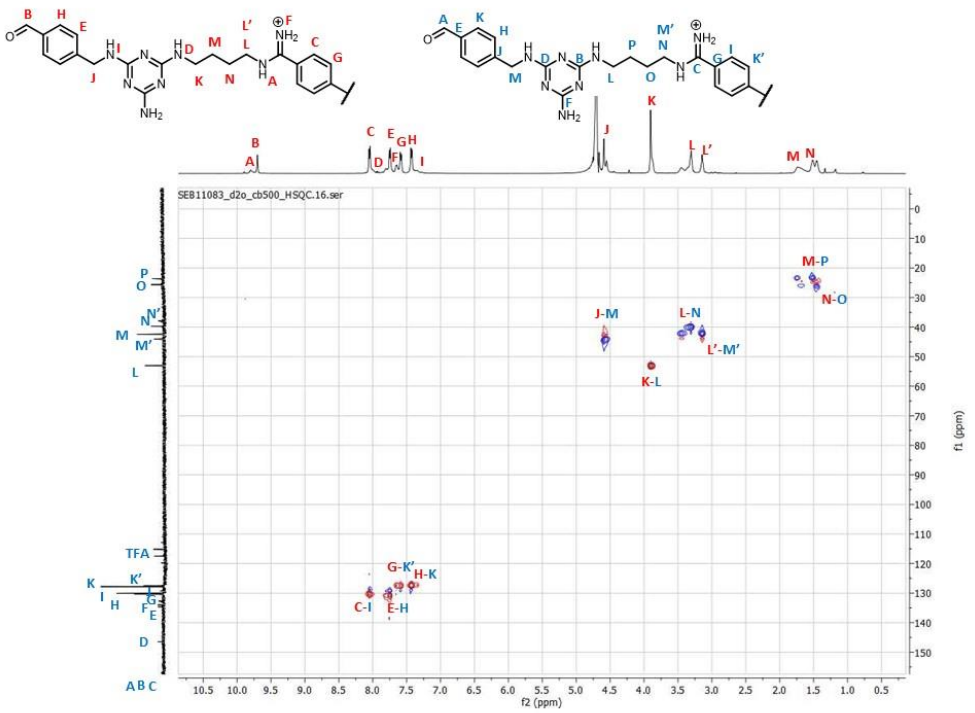


Figure A.4. HSQC NMR in d_2O of **2**.

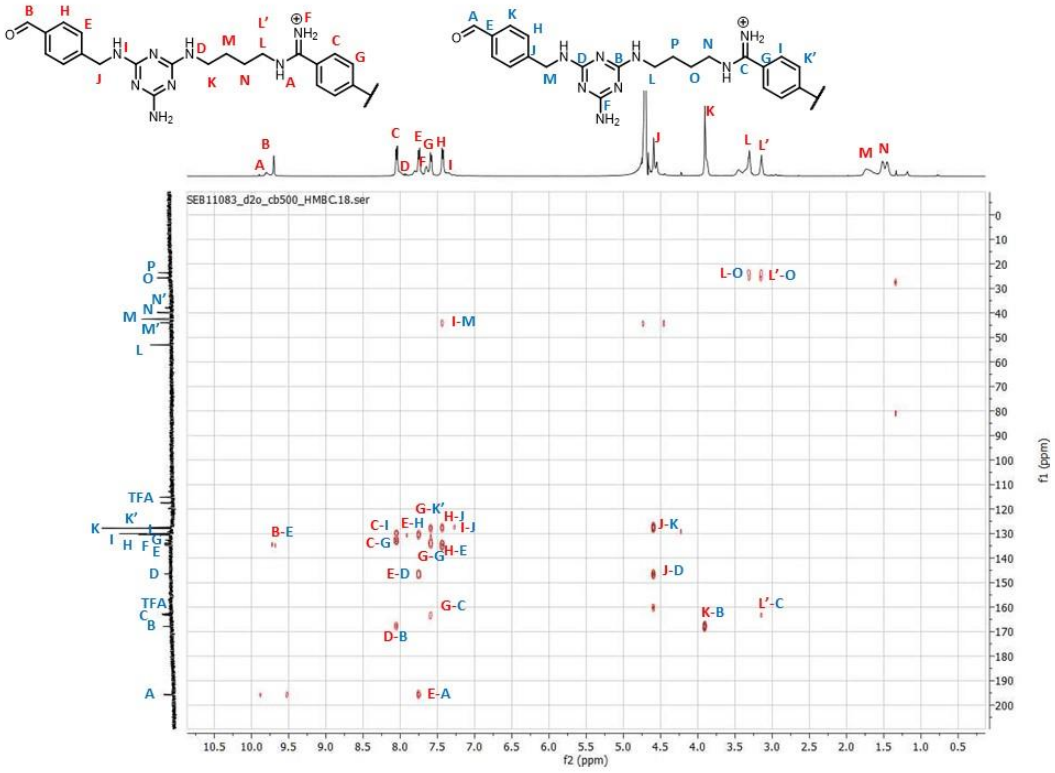
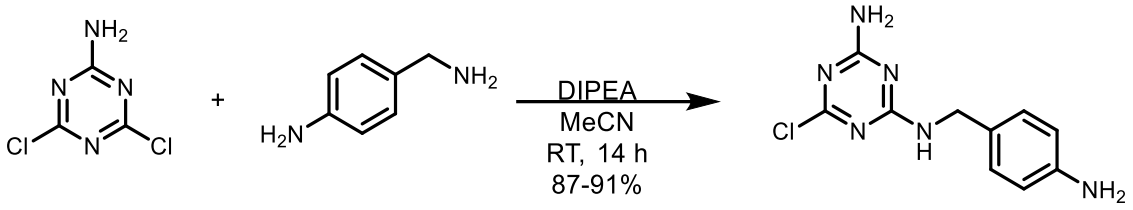
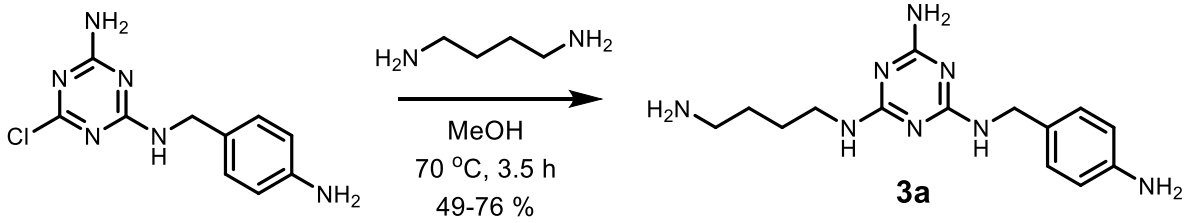


Figure A.5. HMBC NMR in d_2O of **2**.

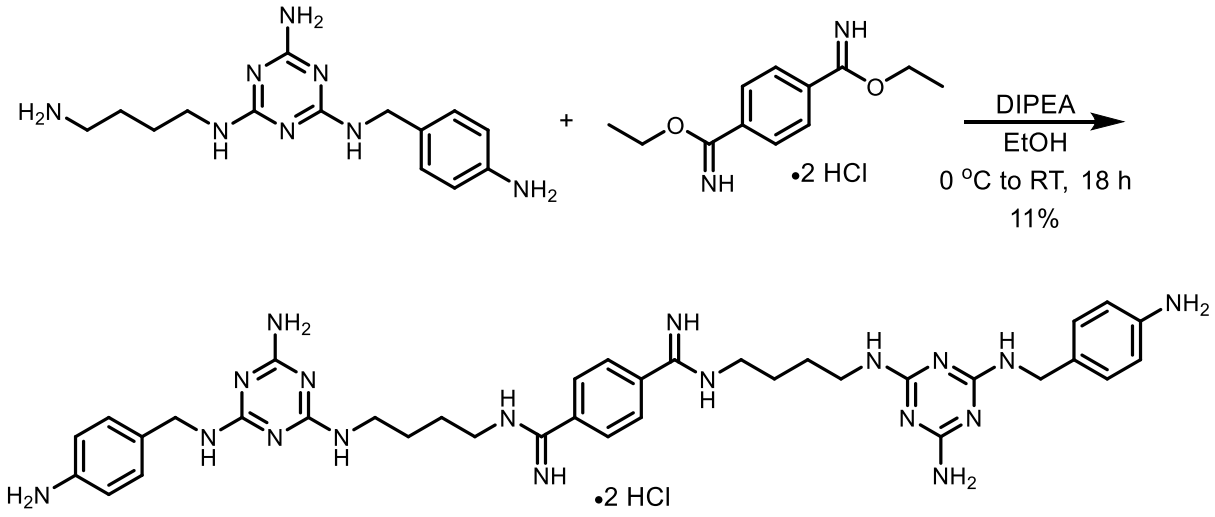


***N*^{2}-(4-aminobenzyl)-6-chloro-1,3,5-triazine-2,4-diamine.** To a 100 mL round-bottom flask was added 0.50 g (3.0 mmol) 4,6-dicholoro-1,3,5-triazin-2-amine, 19 mL acetonitrile, and 1.06 mL (6.1 mmol) *N*,*N*-diisopropylethylamine. The resulting yellow solution was placed under nitrogen and 0.69 mL (6 mmol) 4-aminobenzylamine was added. The white suspension was stirred at room temperature for about 14 h. TLC (9 DCM:1 MeOH, ninhydrin stain, Rf 0.7). The white precipitate was filtered with a Buchner funnel and dried under vacuum for about 1.5 h to yield 691 mg (91%) of the title compound as a white solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.04 (t, *J* = 6.24, 1H), 7.29 (bd, *J* = 5.91, 1H), 6.93 (t, *J* = 8.16, 2H), 6.50 (d, *J* = 8.30, 2H), 4.95 (s, 2H), 4.25 (dd, *J* = 6.22, 11.90, 2H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 168.63, 167.41, 166.08, 148.02, 128.62, 126.42, 114.13, 43.52. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 251.70; found 251.1.



Compound 3a. To a 50 mL round-bottom flask equipped with a reflux condenser was added 0.60 mL (0.60 mmol) 1,4-diaminobutane and 2.0 mL methanol. The clear solution was stirred at 70 °C and to the flask was slowly added 0.301 g (1.2 mmol) *N*^{2}-(4-aminobenzyl)-6-chloro-1,3,5-triazine-2,4-diamine dissolved in 5 mL methanol. The resulting white suspension was stirred at 70 °C for about 3.5 h and dried *in vacuo*. TLC was used to monitor reaction progress (9 MeOH:1 NH_4OH,

ninhydrin stain, Rf 0.4). The crude product was purified on silica gel column with elution gradient 90 DCM:10 MeOH to 70 DCM:30 MeOH + 5 % NH_4OH. The fractions containing product were combined and dried *in vacuo* to obtain 362 mg (76 %) of the title compound as a yellow powder. The product was further purified using preparative HPLC (gradient H_2O (0.1% TFA):MeCN (0.1% TFA) from 100:0 to 0:100 to obtain the title compound as a yellow, crystalline solid. 99% pure by analytical HPLC (retention time 3.952 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.) ^{1}H NMR (500 MHz, DMSO-d_6) δ 6.95 (bs, 2H), 6.81 (bm, 2H), 6.48 (d, J = 8.26, 2H), 6.41 (bs, 1H), 5.91 (bm, 2H), 4.88 (bs, 2H), 4.24 (bd, J = 6.18, 2H), 3.18 (bm, 4H), 2.70 (bs, 2H), 1.48 (bs, 4H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 166.54, 162.79, 155.99, 155.58, 147.63, 128.27, 114.05, 49.07, 36.26,* 27.02*. *Likely result of two overlapping methylene groups on the aliphatic diaminobutane chain. HR-ESI MS (*m/z*) calcd for [M+H$^+$] 303.39; found 303.2056.



Compound 3. To a 100 mL oven-dried round-bottom flask was added 0.180 g (0.614 mmol) diethyl terephthalimidate dihydrochloride and 6.0 mL freshly distilled ethanol. The resulting white suspension was cooled to 0 °C and kept under nitrogen atmosphere. To the white suspension was

added 0.30 mL (1.7 mmol) *N*,*N*-diisopropylethylamine to obtain a colorless solution. A solution of 361 mg (1.19 mmol) N^2-(4-Aminobutyl)-N4-(2,2,dimethoxyethyl)-1,3,5-triazine-2,4,6-triamine dissolved in 10 mL freshly distilled ethanol was added to the round-bottom flask. The colorless suspension was stirred at 0 °C for ~2 h, warmed slowly to 25 °C and stirred for about 18 h. The solvents were removed in vacuo to obtain a tan solid. TLC was used to monitor progress with ninhydrin stain (8 MeOH:2 NH_4OH, Rf 0.6). The crude product was purified on a silica column with elution gradient 9 DCM:1 MeOH to 7 DCM:3 MeOH + 5 % NH_4OH. The fractions containing product were combined and dried *in vacuo* to obtain the title compound as a yellow solid. The product was further purified using preparative HPLC (gradient H_2O (0.1% TFA):MeCN (0.1% TFA) from 90:10 to 0:100 to obtain 27.41 mg (11%) of the title compound as a yellow, crystalline solid. 99% pure by analytical HPLC (retention time 4.160 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.) ^{1}H NMR (500 MHz, DMSO-d6) δ 10.02 (s, 2H), 9.67 (s, 2H), 9.22 (s, 2H), 8.61 (s, 1H)*, 8.32 (s, 2H)*, 7.93 (s, 4H), 7.81 (s, 2H)*, 7.16 (bm, 4H), 6.82 (bm, 4H), 4.41 (bd, *J* = 5.57, 4H), 3.39 (bm, 8H), 1.66 (bm, 8H). *Partially exchanged amines. ^{13}C NMR (500 MHz, DMSO-d6) δ 162.47, 159.13, 158.87, 158.60, 133.54, 129.20, 129.08, 118.29, 115.93, 43.00, 26.40, 25.07. Suspected that additional 2 methylene peaks are suppressed under the DMSO solvent peak. HR-ESI MS (*m/z*) calcd for [M+H] 733.44; found 733.4385.

A.3.2. in vitro Transcription Inhibition Assay

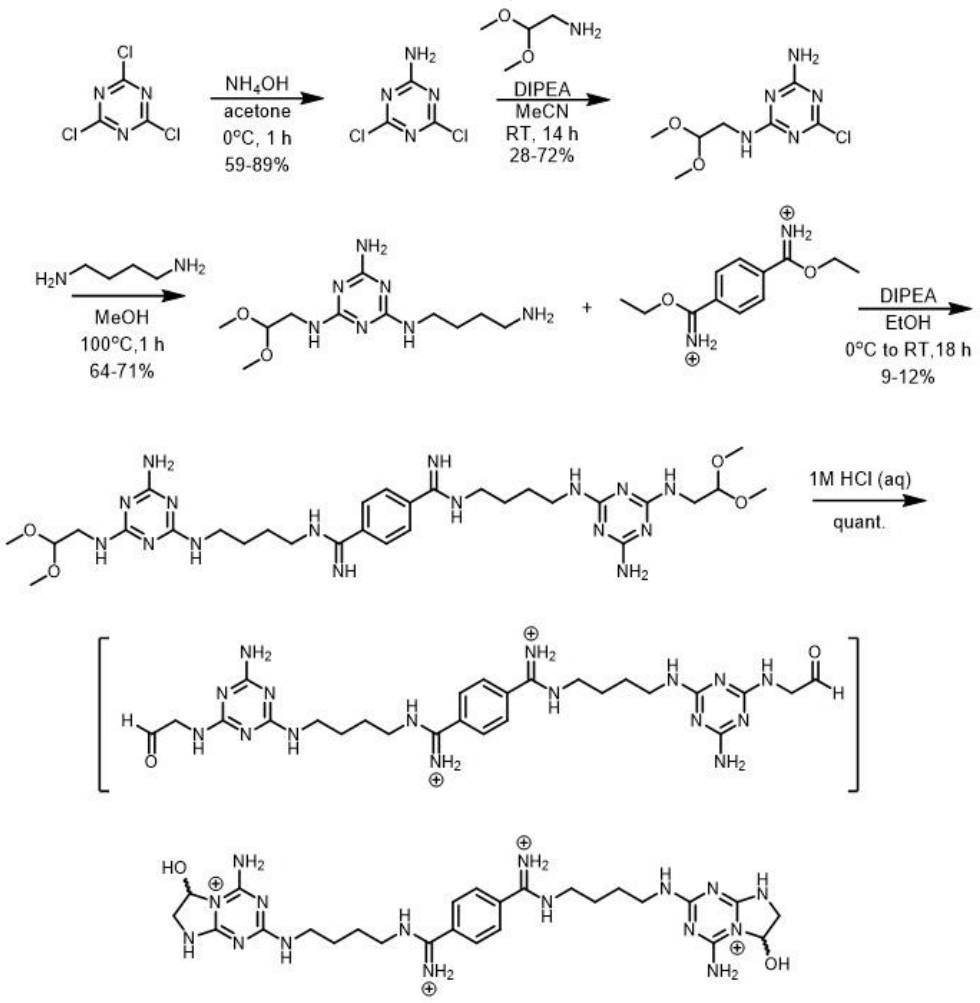
See General Methods section A.1.8 above.

A.4. DYNAMIC COVALENT SCREEN FOR NUCLEIC ACID TARGETING AGENTS

A.4.1. Synthetic Methods

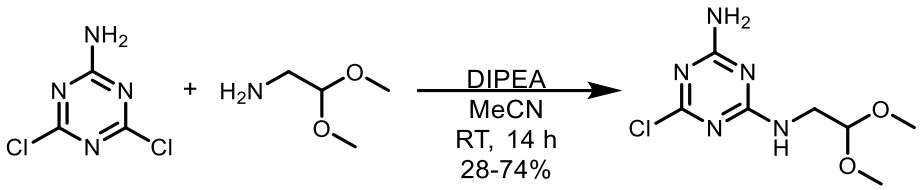
Compounds **A3**, **A4**, **N2**, **N5**, **N6**, and **N7** were purchased from commercial sources, purified to achieve at least 95% purity, and utilized in the screen.

The following were synthesized according to previously reported procedures: diethyl terephthalimidate dihydrochloride,[7] **N9**,[17] and 6-cholor-2-methoxy-9-phenoxyacridine.[18]

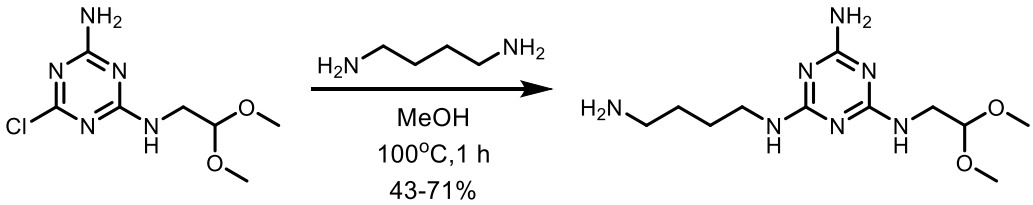


Scheme A.1. Synthetic route to hemiaminal A1. Notably, upon deprotection of the acetal, the cyclized aldehyde product was formed.

*Synthesis of Hemiaminal **A1***

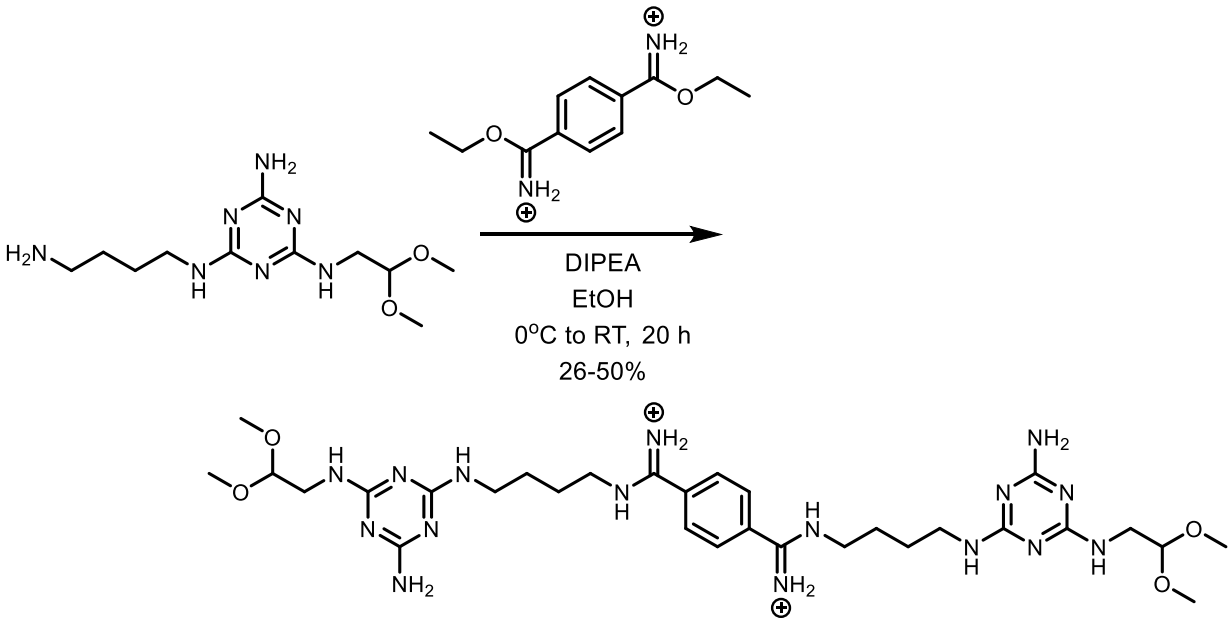


6-Chloro-*N*^{2}-(2,2-dimethoxyehtyl)-1,3,5-triazine-2,4-diamine. To a 500 mL round-bottom flask was added 1.50 g (9.09 mmol) 4,6-dichloro-1,3,5-triazin-2-amine and 25 mL acetonitrile. To the resulting white suspension was added 3.20 mL (18.2 mmol) *N, N*-diisopropylethylamine and the mixture was stirred for about 15 min. A mixture of 2.0 mL (18 mmol) 2,2-dimethoxylethan-1-amine and 32 mL acetonitrile was added dropwise via addition funnel to obtain a faint yellow suspension. The mixture was stirred for about 20 h, filtered with a Buchner funnel, dried under nitrogen, and further dried on a lyophilizer to obtain a white powder that was stored at 0 °C until further use. TLC (9 DCM:1 MeOH, Rf 0.8) was used to monitor reaction progress and the reaction yielded 1.6 g (74%) of the title compound as a white solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.77 (brt, *J* = 6.02, 1H), 7.62 (t, *J* = 6.07, 1H), 7.34 (bs, 2H), 7.19 (bd, *J* = 50.44, 1H) 4.46 (dt, *J* = 17.49, 5.52, 1H), 3.35 (m, 2H), 3.27 (s, 6H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 168.65, 167.34, 166.23, 101.78, 53.45, 42.18. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 234.08; found 234.1.



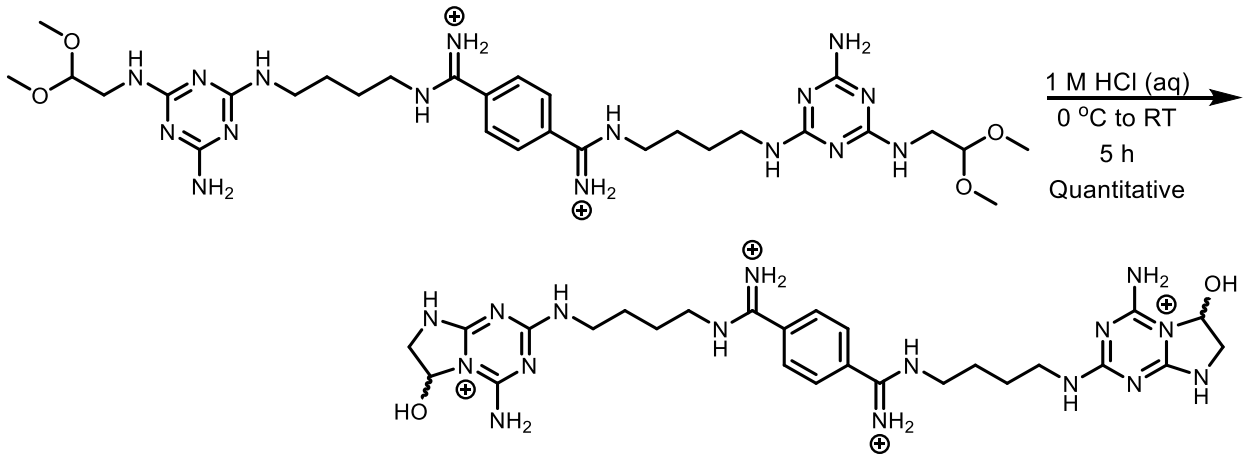
***N*^{2}-(4-Aminobutyl)-*N*^{4}-(2,2,dimethoxyethyl)-1,3,5-triazine-2,4,6-triamine.** To a 50 mL round-bottom flask was added 2.0 mL (20 mmol) 1,4-diaminobutane and 4.5 mL methanol. The mixture was stirred at 100 °C to obtain clear solution and a mixture of 1.0 g (4.4 mmol) 6-chloro-*N*^{2}-(2,2-dimethoxyethyl)-1,3,5-traizine-2,4-diamine in 5 mL methanol was added slowly over 20 min. The

reaction was stirred for an additional 5 h and dried *in vacuo* to obtain colorless oil. TLC was used to monitor progress with ninhydrin stain (9 MeOH:1 NH_4OH, Rf triazine 0.95, Rf pdt 0.4, Rf 1,4-DAB 0.05). The crude product was purified on Combi-Flash RediSep column (silica, 24 g) with elution gradient 100 DCM:0 MeOH to 70 DCM:30 MeOH. The fractions containing product were combined and dried *in vacuo* to obtain 1.25 g (71%) of the title compound as a thin white powder. ^{1}H NMR (500 MHz, DMSO-d_6) δ 6.57 (bs, 1H), 6.42 (bd, J = 17.54, 1H), 6.25 (bd, J = 29.42, 1H), 5.89 (bd, J = 76.21, 1H), 4.47 (m, 1H), 3.29 (m, 2H), 3.26 (s, 6H), 3.17 (m, 2H), 2.57 (t, J = 7.00, 2H), 1.47 (m, 2H), 1.37 (p, J = 6.49, 2H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 167.16, 166.47, 166.37, 102.80, 53.57, 53.36, 41.90, 41.45, 30.04, 27.32. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 286.20; found 286.2.



***N*^{1},*N*^{4}-bis(4-((4-Amino-6-((2,2,dimethoxyethyl)amino)-1,3,5-triazine-2-yl)amino)butyl)terephthalimidamide.** To an oven-dried 100 mL round-bottom flask was added 0.43 g (1.5 mmol) diethyl terephthalimidate dihydrochloride and 25 mL ethanol stored on 4Å molecular sieves. The resulting white suspension was cooled to 0 °C and kept under nitrogen

atmosphere. To the white suspension was added 0.76 mL (4.4 mmol) *N*,*N*-diisopropylethylamine to obtain a colorless solution. A solution of 0.83 g (2.9 mmol) N^2-(4-Aminobutyl)-N^4-(2,2,dimethoxyethyl)-1,3,5-triazine-2,4,6-triamine in 14 mL ethanol stored on 4Å molecular sieves was added to the round-bottom flask dropwise with a syringe over 5 min. The colorless solution was warmed slowly to 25 °C and stirred for about 18 h. TLC was used to monitor progress with ninhydrin stain (8 MeOH:2 NH_4OH, Rf 0.2). The solvents were removed *in vacuo* to obtain a white to tan solid. The crude product was purified on a silica column with elution gradient 9 DCM:1 MeOH to 7.5 DCM:2.5 MeOH + 5% NH_4OH. The fractions containing product were combined and dried *in vacuo* to obtain 508 mg (50%) of the title compound as a white solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.18 (s, 2H), 9.78 (s, 2H), 9.37 (s, 2H), 7.97 (s, 4H), 6.91-6.34 (m, 6H), 4.48 (m, 2H), 3.48 (brq, J = 5.27, 4H), 3.41 (m, 4H), 3.30 (m, 4H), 3.27 (s, 12H), 1.68 (brp, J = 7.05, 4H), 1.60 (m, 4H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 165.43, 162.17, 162.12, 162.09, 133.40, 129.12, 102.31, 53.43, 43.10, 42.11, 41.93, 26.97, 25.23. HR-ESI MS (*m/z*) calcd for [M+H$^+$] 699.43; found 699.4283.



Compound A1. To a 2 mL vial was added 49.1 mg (70.2 µmol) N^1,N^4-bis(4-((4-Amino-6-((2,2,dimethoxyethyl)amino)-1,3,5-triazine-2-yl)amino)butyl)terephthalimidamide. The vial was placed on ice and 0.40 mL (0.43 mmol) 1 M aqueous hydrochloric acid was added dropwise over

5 min while stirring. The vial was slowly warmed to room temperature and stirred for about 5 h. The solvents were removed *in vacuo* to obtain 43 mg (100%) of the title compound as an off-white, crystalline solid. ^{1}H NMR (500 MHz, d_2O) δ 7.81 (s, 4H), 6.05 (q, J = 6.21, 2H), 3.95 (dq, J = 12.43, 7.09, 2H), 3.58 (t, J = 11.42, 2H), 3.41 (m, 8H), 1.70 (m, 8H). ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.29 (s, 2H), 9.81 (s, 1H), 9.43 (m, 4H), 8.79 (d, J = 72.07, 1H), 8.41 (m, 2H), 7.99 (s, 4H), 7.93 (s, 1H), 6.11 (m, 1H), 3.89 (brq, J = 11.21, 2H), 3.50 (bs, 4H), 3.41 (t, J = 10.44, 2H), 3.33 (m, 4H), 1.69 (m, 8H). ^{13}C NMR (500 MHz, d_2O) δ 164.09 (d), 163.26, 157.36 (d), 153.86 (t), 133.30, 128.56, 79.45 (d), 48.80 (t), 42.66 (d), 40.28 (t), 25.72 (t), 24.03 (t) *note: double and triple peaks are likely from diastereomeric products. COSY NMR (500 MHz, d_2O) Figure A.6. HSQC (500 MHz, d_2O) Figure A.7. HMBC NMR (500 MHz, d_2O) Figure A.8. HR-ESI MS (*m/z*) see Figure A.9. LC/MS analysis (positive ESI polarity, gradient water: MeCN from 100:0 to 0:100) suggests that all peaks in HPLC are same MW (diastereomeric products). Product was made into 100 mM stock solution in molecular biology grade water for further experiments.

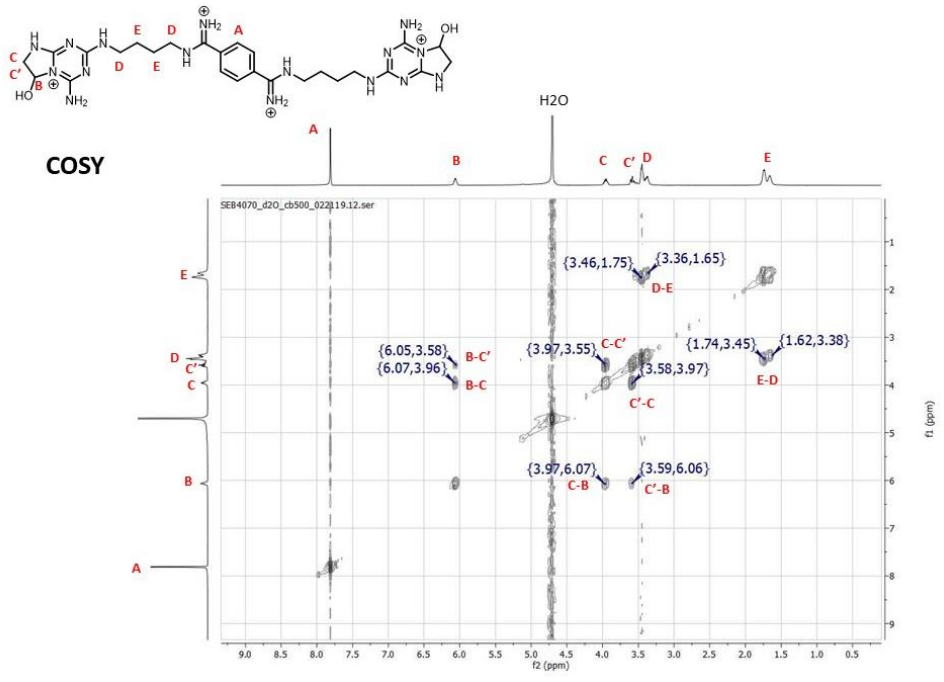


Figure A.6. COSY NMR in d_2O of hemiaminal **A1**.

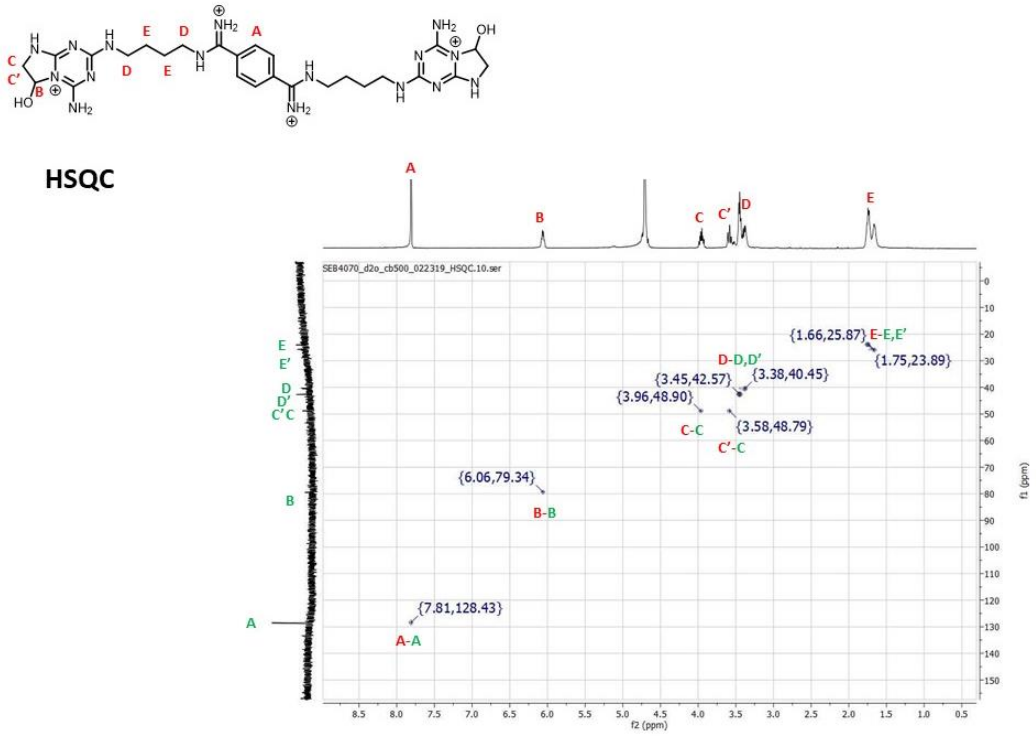


Figure A.7. HSQC NMR in d_2O of hemiaminal **A1**.

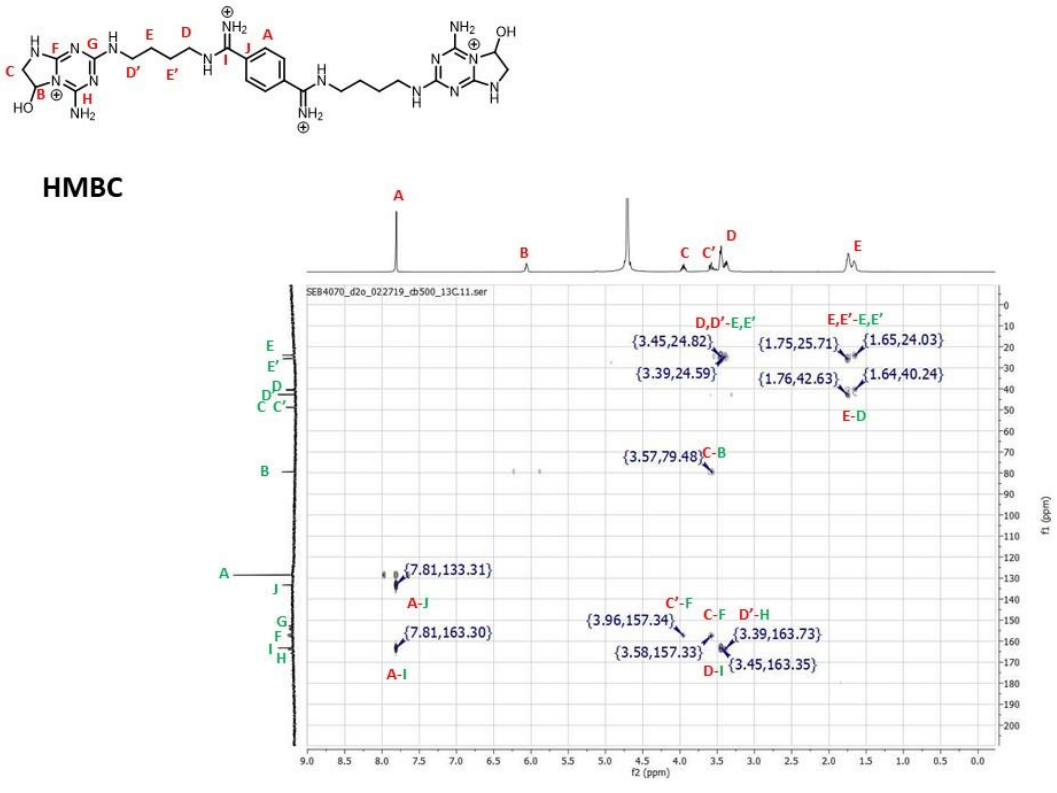
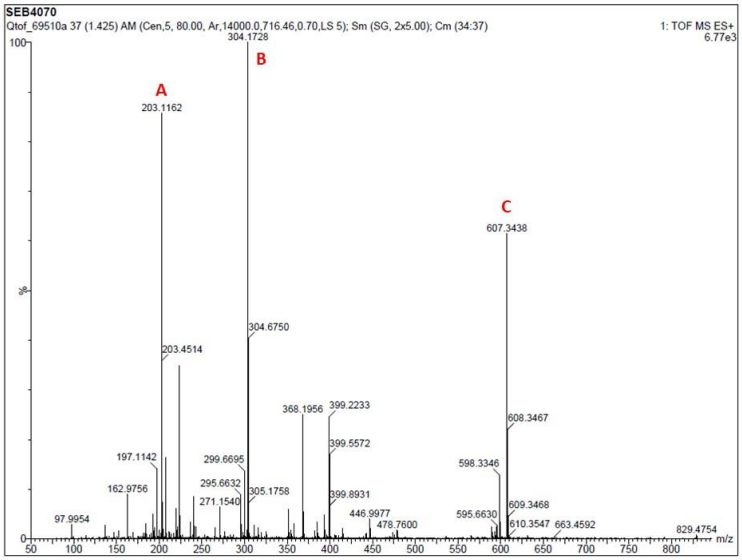
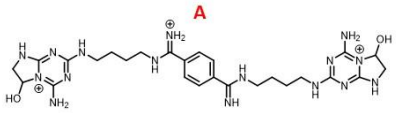


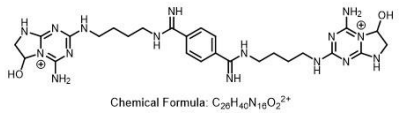
Figure A.8. HSQC NMR in d_2O of hemiaminal **A1**.





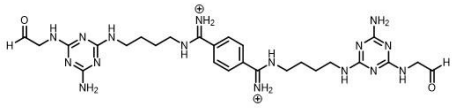
A

Chemical Formula: $C_{26}H_{41}N_{16}O_2^{3+}$
Exact Mass: 609.36
m/z: 203.12 (100.0%), 203.45 (28.1%), 203.45 (4.4%), 203.79 (2.7%), 203.45 (1.5%), 203.79 (1.2%), 203.79 (1.1%)

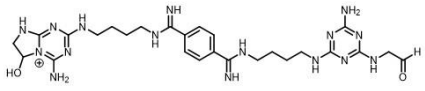


Chemical Formula: $C_{26}H_{40}N_{16}O_2^{2+}$
Exact Mass: 608.35
m/z: 304.18 (100.0%), 304.68 (28.1%), 304.67 (5.9%), 305.18 (3.8%), 305.18 (1.7%)

B

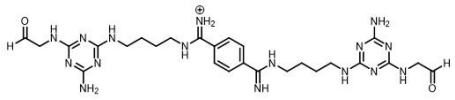


Chemical Formula: $C_{26}H_{40}N_{16}O_2^{2+}$
Exact Mass: 608.35
m/z: 304.18 (100.0%), 304.68 (28.1%), 304.67 (5.9%), 305.18 (3.8%), 305.18 (1.7%)



Chemical Formula: $C_{26}H_{39}N_{16}O_2^{+}$
Exact Mass: 607.34
m/z: 607.34 (100.0%), 608.35 (28.1%), 608.34 (4.4%), 609.35 (2.7%), 608.34 (1.5%), 609.34 (1.2%), 609.35 (1.1%)

C



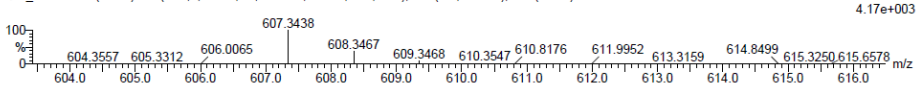
Chemical Formula: $C_{26}H_{39}N_{16}O_2^{+}$
Exact Mass: 607.34
m/z: 607.34 (100.0%), 608.35 (28.1%), 608.34 (4.4%), 609.35 (2.7%), 608.34 (1.5%), 609.34 (1.2%), 609.35 (1.1%)

Elemental Composition Report **Page 1**

Single Mass Analysis
Tolerance = 5.0 PPM / DBE: min = -1.5, max = 100.0
Element prediction: Off
Number of isotope peaks used for i-FIT = 3

Monoisotopic Mass, Even Electron Ions
518 formula(e) evaluated with 3 results within limits (up to 50 closest results for each mass)
Elements Used:
C: 0-200 H: 0-200 N: 12-18 O: 0-5 Na: 0-1

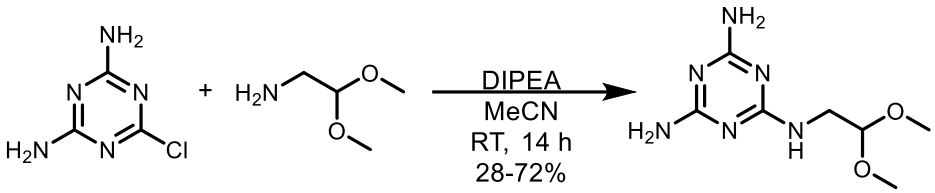
SEB4070
Qtof_69510a 37 (1.425) AM (Cen,5, 80.00, Ar,14000.0,716.46,0.70,LS 5); Sm (SG, 2x5.00); Cm (34:37)
1: TOF MS ES+
4.17e+003



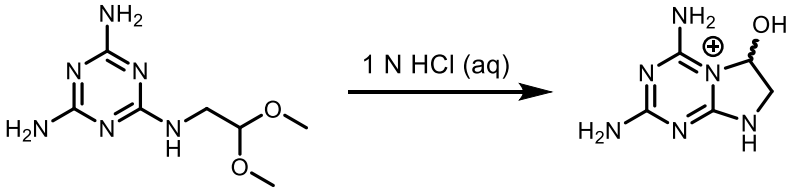
Minimum:				-1.5		
Maximum:		5.0	5.0	100.0		
Mass	**Calc. Mass**	**mDa**	**PPM**	**DBE**	**i-FIT**	**Formula**
607.3438	607.3442	-0.4	-0.7	15.5	2.4	C26 H39 N16 O2
	607.3458	-2.0	-3.3	16.5	2.8	C29 H40 N14 Na
	607.3418	2.0	3.3	12.5	10.1	C24 H40 N16 O2 Na

Figure A.9. HR-ESI MS and Elemental Composition Report for hemiaminal **A1**.

*Synthesis of Hemiaminal **A2***



***N*^{2}-(2,2-dimethoxyehtyl)-1,3,5-triazine-2,4,6-triamine.** To a 100 mL round-bottom flask was added 0.90 mL (8.2 mmol) 2,2-dimethoxylethan-1-amine and 24 mL deionized water. To the resulting clear solution was added 1.2 g (15 mmol) sodium bicarbonate and 1.1 g (7.4 mmol) 6-chloro-1,3,5-triazine-2,4-diamine. The white suspension was stirred at 90 °C for about 36 h, cooled to room temperature, and rotovapped to remove about ½ of the water. The white solid was filtered with a Buchner funnel, washed with 250 mL cold (0 °C) water, and dried under nitrogen. TLC (8 DCM:2 MeOH, Rf 0.6) was used to monitor reaction progress and the reaction yielded 0.53 g (36%) of the title compound as a white solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 6.46 (t, J = 6.08, 1H), 6.13 (m, 4H), 4.46 (t, J = 5.52, 1H), 3.29 (t, J = 5.86, 2H), 3.26 (s, 6H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 167.48 (d), 166.77, 151.26, 102.48, 53.41, 41.93. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 215.13; found 215.1.



Compound A2.[19] To a 2 mL vial was added 202 mg (0.943 mmol) *N*^{2}-(2,2-dimethoxyehtyl)-1,3,5-triazine-2,4,6-triamine. The vial was placed on ice and 3.0 mL (2.8 mmol) 1 M aqueous hydrochloric acid was added dropwise over 5 min while stirring. The vial was slowly warmed to

room temperature and stirred for about 3.5 h. The solvents were removed *in vacuo* to obtain 43 mg (>100%) of the title compound as an off-white, crystalline solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 11.16 (s, 2H), 9.33 (s, 1H), 8.52 (s, 1H), 8.25 (s, 1H), 7.78 (d, J = 18.62, 2H), 6.60 (dd, J = 7.45, 2.27, 1H), 3.87 (dd, J = 11.37, 7.55, 1H), 3.40 (dd, J = 11.31, 2.10, 1H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 166.30, 158.05, 154.49, 150.39, 79.42, 48.59. Note: 1 extra signal is observed in both ^{1}H and ^{13}C NMR that is not currently understood. This experiment has been repeated twice by two independent researchers and the same extra peaks were seen in the NMRs. Although the proton signal could be from some aldehyde in the equilibrium, the ^{13}C spectra have no peaks in the aldehyde region (around 200 ppm). Product is 99% pure by HPLC. COSY NMR (500 MHz, d_2O) Figure A.10. HSQC (500 MHz, d_2O) Figure A.11. HR-ESI MS (*m/z*) calcd [M^+] 169.08; found 169.0830. Elemental Composition 169.0838 $C_5H_9N_6O$ (desired [M^+] $C_5H_9N_6O$).

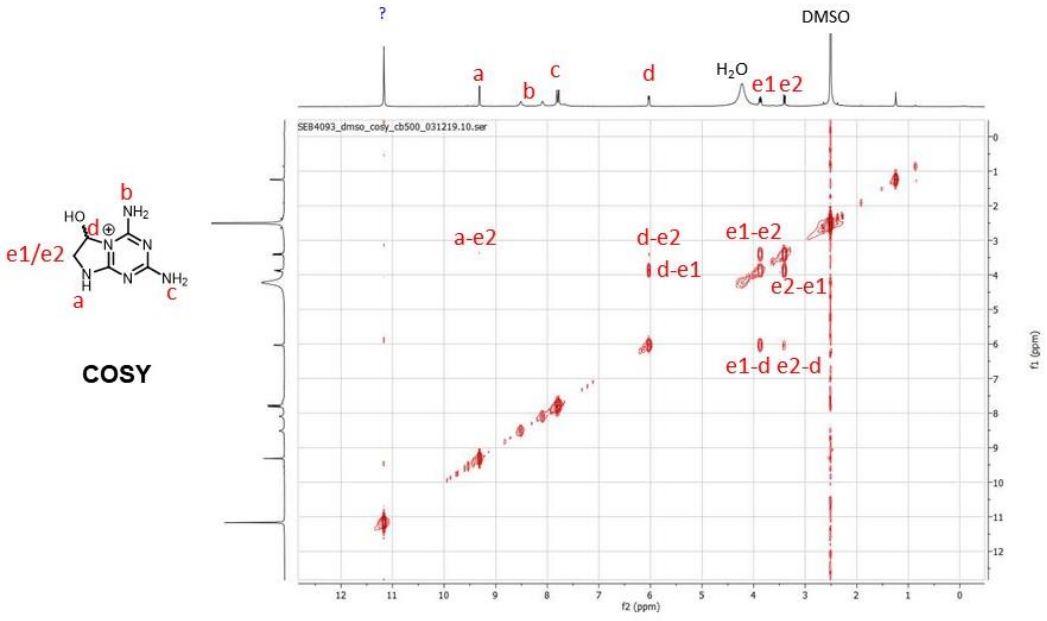


Figure A.10. COSY NMR is DMSO-d_6 of **A2**.

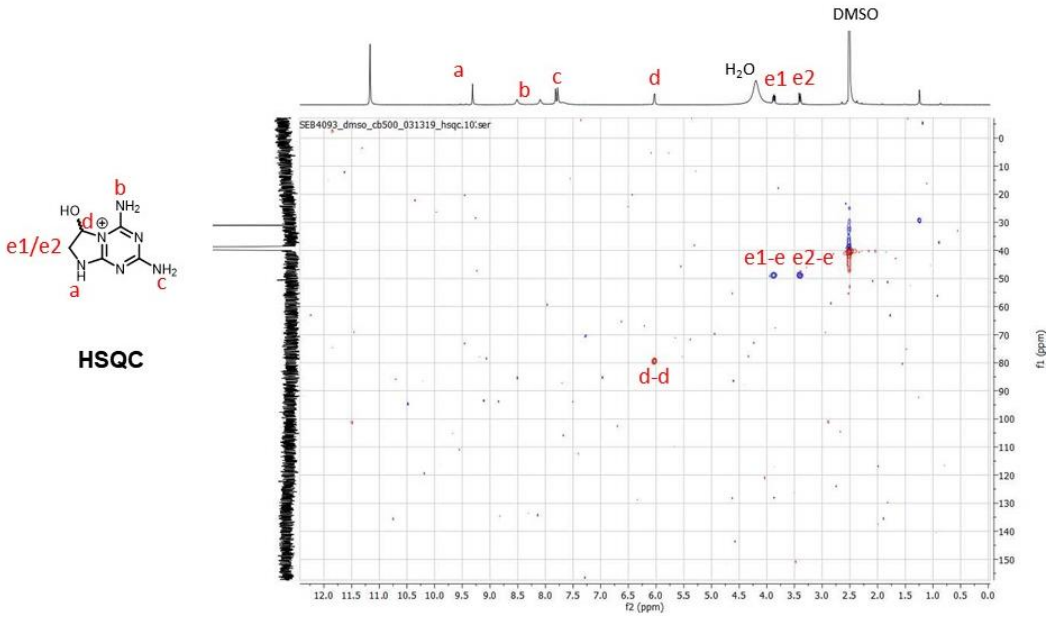
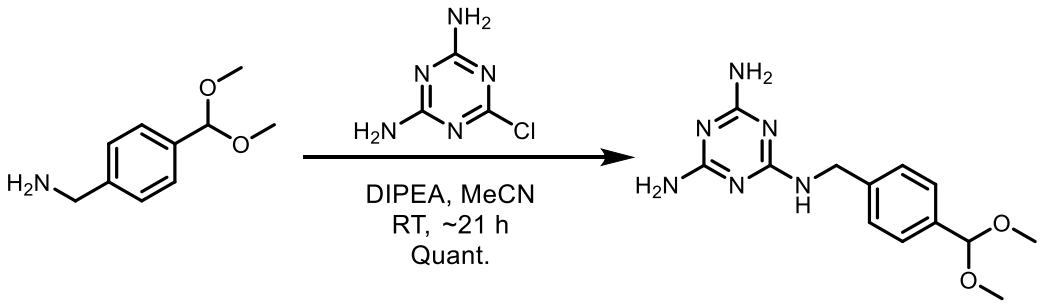


Figure A.11. HSQC NMR is DMSO-d_6 of **A2**.

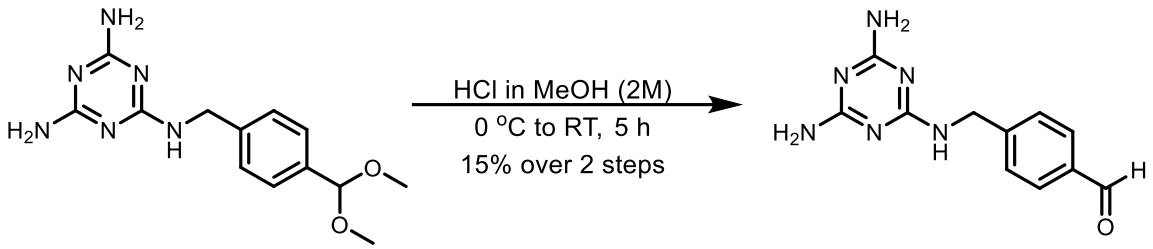
*Synthesis of Compound **A5***

See procedures in Section A.3.1, compound **2** above.

Synthesis of **A6**

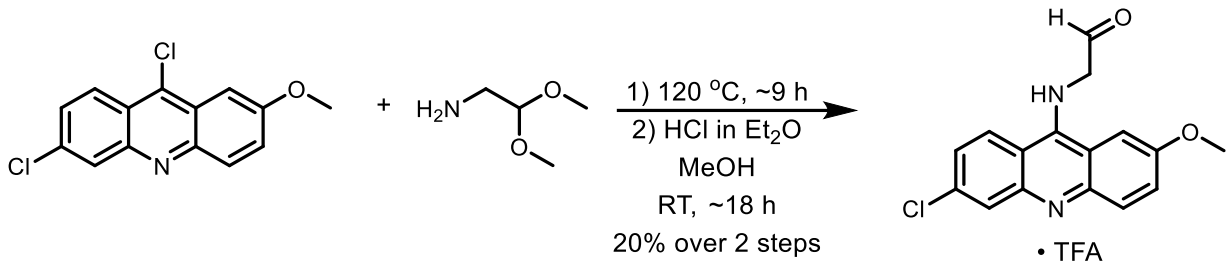


***N*^{2}-(4-(dimethoxymethyl)benzyl)-1,3,5-triazine-2,4,6-triamine.** To a 25 mL round-bottom flask was added 0.439 g (2.42 mmol) (4-(dimethoxymethyl)phenyl)methanamine and 1.0 mL methanol. The resulting tan solution was placed under nitrogen and a solution of 0.282 g (1.94 mmol) 6-chloro-1,3,5-triazine-2,4-diamine dissolved in 3.0 mL acetonitrile was added dropwise. The resulting cloudy white solution was stirred at room temperature for about 21 h. TLC (9 DCM:1 MeOH, ninhydrin stain, Rf 0.7) was used to monitor reaction progress. The white solution was dried *in vacuo* to obtain a white solid that was purified after deprotection. ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.44 (d, J = 7.90, 2H), 7.39 (d, J = 7.85, 2H), 7.21 (s, 2H), 7.13 (s, 2H), 6.05 (m, 1H), 5.39 (s, 1H), 3.93 (s, 2H), 3.24 (s, 6H). ^{13}C NMR (500 MHz, DMSO-d6) δ 169.11, 167.54, 138.22, 138.10, 128.48, 127.07, 126.78, 102.89, 52.97, 43.63. LR-ESI MS (m/z) calcd for $C_{13}H_{19}N_6O_2^+$, [M+H$^+$] 291.16; found 291.2.



Compound A6. To a 20 mL scintillation vial was added 0.559 g (1.93 mmol) N^2-(4-(dimethoxymethyl)benzyl)-1,3,5-triazine-2,4,6-triamine and the vial was chilled on ice. To the vial was added 8.0 mL of 2 M HCl in MeOH dropwise. The vial was slowly warmed to room temperature and stirred for 5 h. TLC (8 DCM:2 MeOH, ninhydrin stain, Rf 0.3) was used to monitor reaction progress. When no more starting material was observed, the suspension was dried *in vacuo* to obtain a tan solid. Preparative HPLC was used to purify 49.22 mg of the crude product (gradient H2O (0.1% TFA):MeCN (0.1% TFA) from 100:0 to 0:100 to yield 6.3 mg (15% over two steps) of the title compound as a white solid. Product is 96% pure by analytical HPLC (retention time 3.729 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.). ^{1}H NMR (500 MHz, d_2O) δ 9.84 (s, 1H), 7.86 (d, J = 8.32, 2H), 7.47 (d, J = 7.87, 2H), 4.61 (s, 2H). Amine protons likely exchanged in d_2O solvent. ^{13}C NMR (500 MHz, DMSO-d6) δ 195.99, 160.40, 157.58, 145.64, 134.89, 130.45, 127.57, 43.83. HR-ESI MS (m/z) calcd for $C_{11}H_{13}N_6O^+$, [M+H$^+$] 245.1145; found 245.1144.

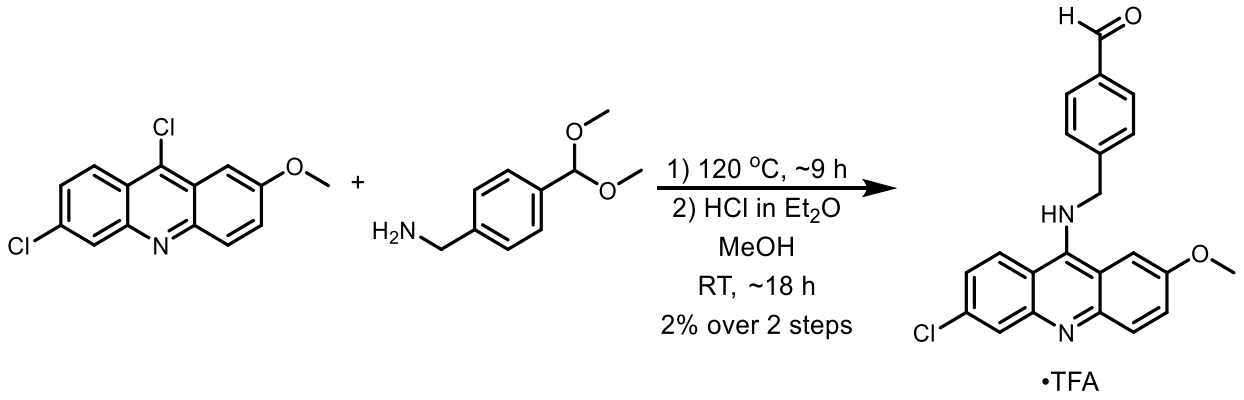
Synthesis of **A7**



Compound A7. To a 50 mL round-bottom flask was added 1.995 g (19.0 mmol) 2,2-dimethoxyethan-1-amine and 250 mg (0.90 mmol) 6,9-dichloro-2-methoxyacridine. The yellow solution was heated to 120 °C in an oil bath under nitrogen. The resulting brown solution was stirred under was reflux for about 9 h. Reaction progress was monitored by TLC (9 DCM:1 MeOH, Rf 0.8). Solvents were removed *in vacuo* to yield the crude product as a brown solid that was directly deprotected. To the flask containing the crude product was added 15 mL methanol to obtain an orange solution followed by 4.0 mL 2 M HCl in diethyl ether dropwise while stirring. The resulting yellow suspension was stirred overnight. To the flask was added an additional 4.0 mL HCl in diethyl ether and 1.0 mL DI water. TLC was used to monitor reaction progress (9 DCM:1 MeOH, Rf 0.3). Solvents were removed *in vacuo* and the crude orange solid was purified using a CombiFlash silica column (12 g) with gradient from 100 DCM:0 MeOH to 70 DCM:30 MeOH. Fractions containing product were combined. The product was further purified by preparatory HPLC using gradient from 0 to 100% acetonitrile in water, 0.1% TFA to obtain 14.68 mg (20%) the title compound as a yellow solid TFA salt. 99% pure by HPLC. (retention time 5.613 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.) ^{1}H NMR (600 MHz, DMSO-d_6) δ 8.43 (d, J = 9.19, 1H), 7.81 (d, J = 2.16, 1H), 7.78 (d, J = 9.30, 1H), 7.72 (bs, 1H), 7.44 (dd, J = 2.30, 9.44, 1H), 7.37 (dd, J =1.69, 9.97), 6.55 (s, 1H), 3.93 (s, 2H), 2.55 (s, 2H). NH seems to be exchanged. ^{13}C NMR (600 MHz,

DMSO-d_6) δ 206.87, 160.66, 158.24, 155.40, 149.31, 148.35, 143.78, 137.82, 126.05, 122.96, 119.71, 117.25, 116.88, 113.63, 63.10, 56.26. LR-ESI MS (*m/z*) calcd for [M+H] 301.07; found 259.1*. HR-ESI MS (*m/z*) calcd for [M+H] 301.0738; found 259.0632*. MALDI MS (*m/z*) calcd for [M+H] 301.07; found 259.0*. *fragmentation between acridine NH and CH_2 group α to aldehyde.

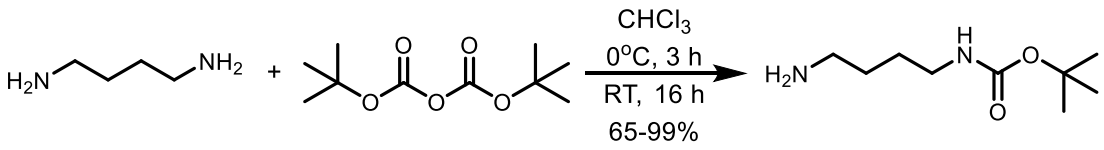
Synthesis of **A8**



Compound A8. To a 35 mL round-bottom flask was added 674.2 mg (3.72 mmol) (4-dimethoxymethyl)phenyl)methanamine and 305.9 mg (1.10 mmol) 6,9-dichloro-2-methoxyacridine. The yellow solution was heated to 120 °C in an oil bath under nitrogen. To the resulting orange solution was added 2.0 mL ethanol and the mixture was refluxed for about 9 h. Reaction progress was monitored by TLC (9 DCM:1 MeOH, Rf 0.6). Solvents were removed *in vacuo* to yield the crude product as a yellow solid that was directly deprotected. To the flask containing the crude product was added 10 mL methanol to obtain a yellow suspension followed by 2.0 mL 2 M HCl in diethyl ether dropwise while stirring. The suspension was stirred overnight. To the flask was added an additional 4.0 mL HCl in diethyl ether and 1.0 mL DI water. TLC was used to monitor reaction progress (9 DCM:1 MeOH, Rf 0.3). Solvents were removed *in vacuo* and

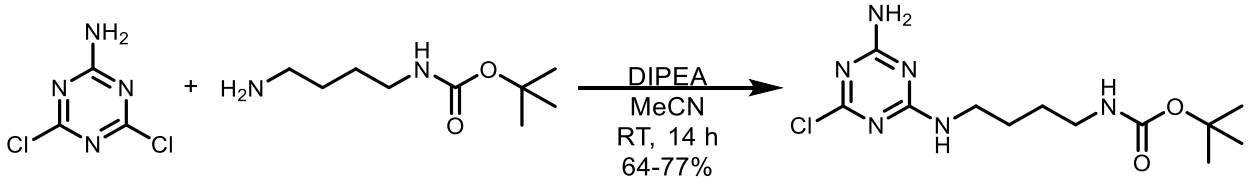
the crude product was purified using a CombiFlash silica column (12 g) with gradient from 100 DCM:0 MeOH to 70 DCM:30 MeOH. Fractions containing product were combined. The product was further purified by preparatory HPLC using gradient from 0 to 100% acetonitrile in water, 0.1% TFA to obtain the title compound as a yellow solid TFA salt. 99% pure by HPLC. (retention time 6.435 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.) ^{1}H NMR (600 MHz, DMSO-d_6) δ 10.03 (s, 1H), 8.39 (s, 1H), 7.98 (d, J = 8.18, 2H), 7.91 (d, J = 2.18, 1H), 7.88 (d, J = 9.27, 1H), 7.76 (d, J = 8.02, 2H), 7.71 (d, J = 8.80, 1H), 7.49 (bs, 2H), 5.43 (s, 2H), 3.78 (s, 3H). ^{13}C NMR (600 MHz, DMSO-d_6) δ 193.25, 158.32, 158.12, 157.92, 144.36, 139.44, 136.21, 136.16, 130.59, 128.10, 127.91, 124.30, 124.24, 124.23, 118.87, 113.96, 104.19, 104.04, 56.37, 51.51. LR-ESI MS (*m/z*) calcd for [M+H] 377.11; found 377.1.

Synthesis of Diamine **N1**

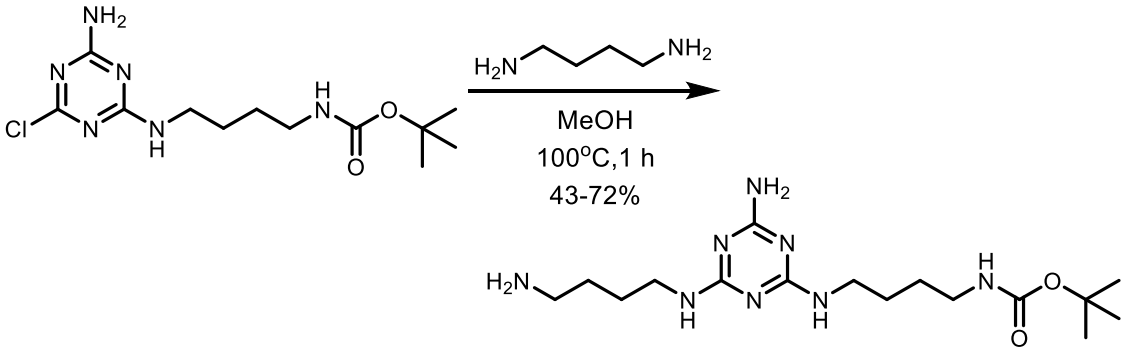


tert-butyl (4-aminobutyl)carbamate:[14] To a 1000 mL round-bottom flask was added 28.3 mL (282 mmol) 1,4-diaminobutane and 355 mL chloroform. The resulting clear solution was cooled to 0 °C in an ice bath. In a 500 mL Erlenmeyer flask was mixed 13.09 g (60.0 mmol) di-tert-butyl dicarbonate and 300 mL chloroform to form a clear solution that was added dropwise to the round-bottom flask using an addition funnel while stirring. The resulting milky white mixture was kept at 0 °C for 3 h, removed from the ice bath to slowly warm to room temperature and stirred for an additional 16 h. Thin layer chromatography on silica gel plates (9 MeOH:1 NH_4OH, rf 0.6, ninhydrin stain) was used to monitor reaction progress. The white solution was decanted into a

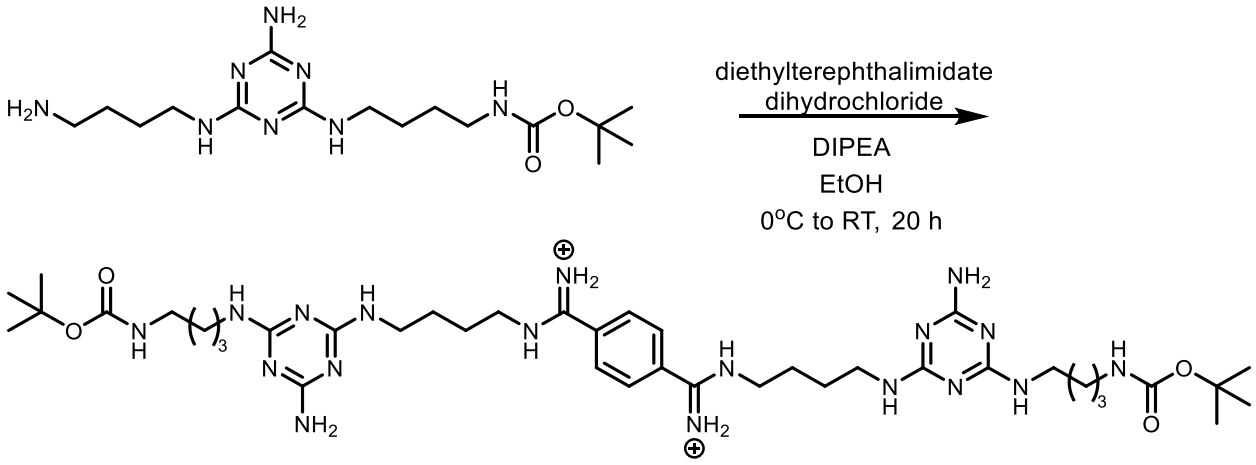
1000 mL separatory funnel, washed with DI water (10 x 200 mL) and the organic layer was checked for remaining 1,4-diaminobutane by TLC (9 MeOH:1 NH_4OH, 1,4-diaminobutane rf 0.05, pdt rf 0.6). The organic layer was dried over sodium sulfate and concentrated *in vacuo* to obtain 11.3 g (99 %) of the title compound as a colorless oil.^{1}H NMR (500 MHz, d_6-DMSO) δ 6.78 (s, 1H), 2.88 (q, J = 6.62 Hz, 2H), 2.50 (t, J = 6.82, 2H), 1.37 (s, 9H), 1.365 (m, 2H), 1.29 (m, 2H), 1.23 (s, 2H). ^{13}C NMR (500 MHz, d_6-DMSO) δ 155.53, 77.22, 41.45, 40.98, 30.73, 28.27, 27.03. LR-ESI MS (*m/z*) calculated 189.16, found [M+H$^+$] 189.2.



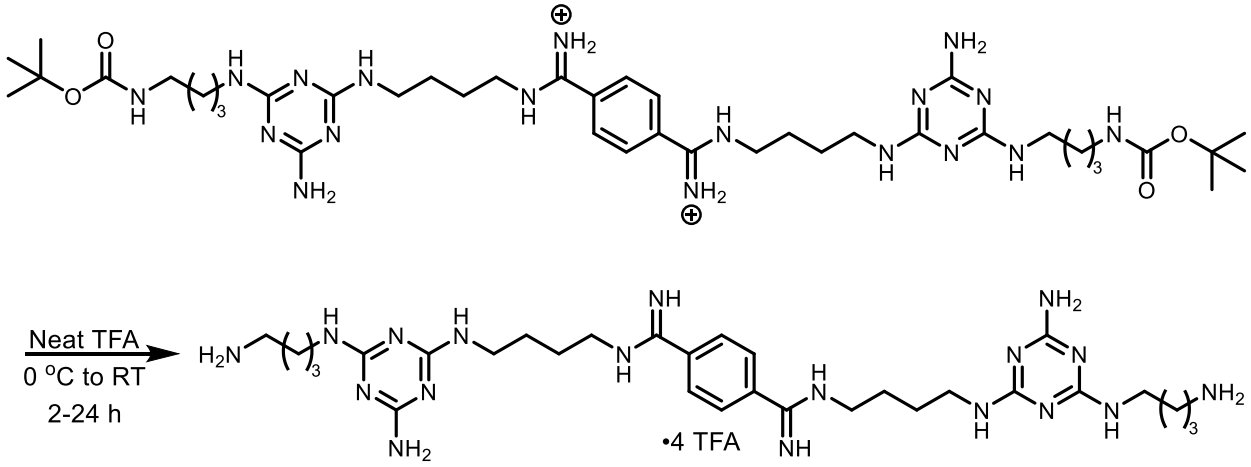
tert-butyl (4-((4-Amino-6-chloro-1,3,5-triazin-2-yl)amino)butyl)carbamate. To a 250 mL round-bottom flask was added 2.00 g (12.2 mmol) 4,6-dichloro-1,3,5-triazin-2-amine and 50 mL acetonitrile. To the resulting white suspension was added 2.10 mL (12.1 mmol) N, N-diisopropylethylamine and the mixture was stirred. A mixture of 2.27 g (12.1 mmol) tert-butyl (4-aminobutyl)carbamate and 25 mL acetonitrile was added dropwise via syringe to obtain a faint white suspension. The mixture was stirred for about 14 h, filtered with a Buchner funnel, and dried on a lyophilizer to obtain a white powder that was stored at 0 °C until further use. TLC (9 DCM:1 MeOH, Rf 0.8) was used to monitor reaction progress and the reaction yielded 2.93 g (77%) of the title compound as a white solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.73 (t, *J* = 7.73, 1H), 7.26 (s, 2H), 6.78 (t, *J* = 5.64, 1H), 3.17 (p, *J* = 6.99, 2H), 2.90 (q, *J* = 6.51, 2H), 1.43 (m, 2H), 1.39 (m, 2H), 1.37 (s, 9H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 168.00, 166.94, 165.53, 155.56, 77.35, 39.95, 39.78, 28.27, 26.999, 26.08. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 317.14; found 317.1.



tert-butyl(4-((4-Amino-6-((4-aminobutyl)amino)-1,3,5-triazin-2-yl)amino)butyl)carbamate. To a 100 mL round-bottom flask was added 0.88 g (9.9 mmol) 1,4-diaminobutane and 5 mL methanol. The mixture was stirred at 100 °C to obtain clear solution and 0.72 g (2.3 mmol) tert-butyl (4-((4-Amino-6-chloro-1,3,5-triazin-2-yl)amino)butyl)carbamate was added slowly over 15 min. The reaction was stirred for an additional 45 min and dried *in vacuo* to obtain colorless oil. TLC was used to monitor progress with ninhydrin stain (8 MeOH:2 NH_4OH, Rf triazine 0.95, Rf pdt 0.5, Rf 1,4-DAB 0.05). The crude product was purified on Combi-Flash RediSep column (silica, 24 g) with elution gradient 100 DCM:0 MeOH to 70 DCM:30 MeOH. The fractions containing product were combined and dried *in vacuo* to obtain 0.60 g (72%) of the title compound as a thin, tan solid. Yield: 0.60 g (72%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 6.83 (t, J = 5.71, 1H), 6.15 (m, 4H), 3.14 (m, 4H), 2.89 (q, J = 6.53, 2H), 2.52 (t, J = 6.88, 2H), 2.03 (s, 2H), 1.43 (m, 4H), 1.37 (s, 9H), 1.33 (m, 4H). HR-ESI MS (*m/z*) calcd for [M+H$^+$] 369.49; found 369.2735.

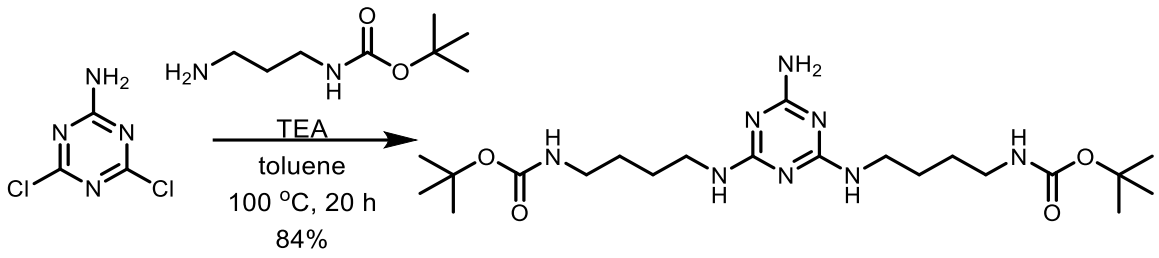


1,4-Phenylenebis(((4-((4-amino-6-((4-((tert-butoxycarbonyl)amino)butyl)amino)-1,3,5-triazin-2-yl)amino) butyl)amino)methaniminium). To an oven-dried 100 mL round-bottom flask was added diethyl terephthalimidate and 20 mL anhydrous ethanol. The resulting white suspension was cooled to 0 °C and kept under nitrogen atmosphere. To the white suspension was added *N*,*N*-diisopropylethylamine to obtain a colorless solution. A solution of N^2-(4-Aminobutyl)-N^4-(2,2,dimethoxyethyl)-1,3,5-triazine-2,4,6-triamine in 10 mL ethanol was added to the round-bottom flask. The colorless suspension was stirred at 0 °C for 2 h, warmed slowly to 25 °C and stirred for about 14 h. The solvents were removed *in vacuo* to obtain a white solid. TLC was used to monitor progress with ninhydrin stain (8 MeOH:2 NH_4OH, Rf 0.43). The crude product was purified on silica column with elution gradient 72.2 DCM:25.3 MeOH:2.5 NH_4OH to 37.9 DCM:53.0 MeOH:9.1 NH_4OH. The fractions containing product were combined and dried *in vacuo* to obtain the title compound as a tan solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.80 (s, 4H), 6.79 (bs, 2H), 6.51 (m, 2H), 6.33 (m, 2H), 5.93 (m, 4H), 3.18 (m, 12H), 2.90 (q, J = 6.51, 4H), 1.60 (m, 8H), 1.41-1.37 (m+s, 26H). Note: 2 amidinium protons missing. ^{13}C NMR (500 MHz, DMSO-d_6) δ 167.24, 166.54, 166.41, 160.96, 156.05, 127.18, 127.10, 77.81, 40.59, 40.42, 40.26, 28.76, 27.82, 27.58, 27.37. HR-ESI MS (*m/z*) calcd for [M+2H$^+$]/2 433.29; found 433.2874, calcd [M+2H$^+$]/2 – 1 Boc 383.26, found 383.2615, calcd [M+2H$^+$]/2 – 2 Boc 333.24, found 333.2355.

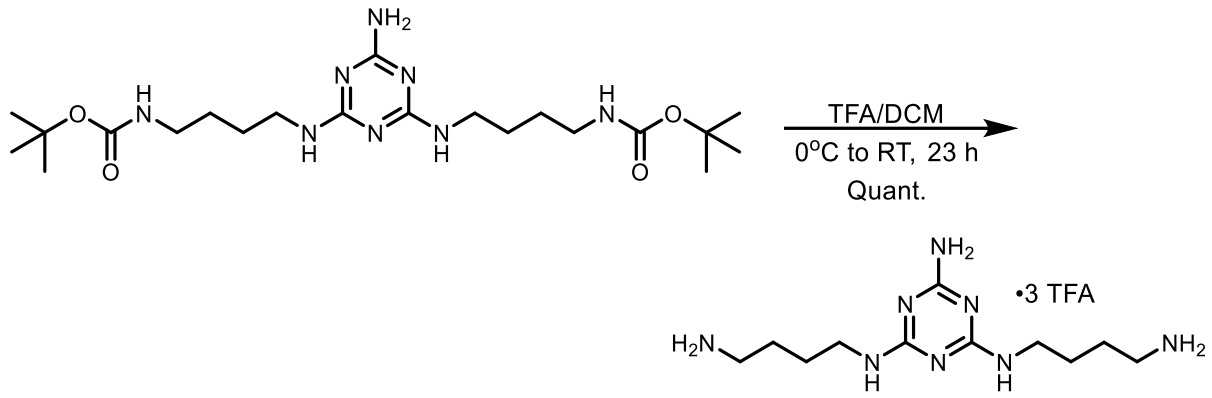


Compound N1. To a 2 mL vial was added 20.9 mg (0.24 mmol) 1,4-Phenylenebis(((4-((4-amino-6-((4-((tert-butoxycarbonyl)amino)butyl)amino)-1,3,5-triazin-2-yl)amino)butyl)amino)methaniminium). The vial was placed on ice and 0.15 mL (2.0 mmol) trifluoroacetic acid was added dropwise over 5 min while stirring. The vial was slowly warmed to room temperature and stirred for about 12 h. The solvents were removed in vacuo and the product was purified by preparative HPLC (gradient H_2O (0.1% TFA):MeCN (0.1% TFA) from 98:2 to 0:100 to obtain the title compound as a white, crystalline solid. No yield is reported because the compound was purified in small portions. ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.09 (s, 2H), 9.71 (s, 2H), 9.28 (s, 2H), 8.34 (m, 2H), 7.94 (s, 4H), 7.85 (m, 6H), 3.50-3.25 (m, 12H), 2.82 (bs, 4H), 1.68 (m, 8H), 1.56 (bs, 8H). ^{1}H NMR (500 MHz, d_2O) δ 7.73 (s, 4H), 3.64 (m, 2H), 3.50 (dd, *J* = 11.74, 4.34, 4H) 3.44-3.24 (m, 16H), 2.88 (bs, 4H), 1.73-1.48 (m, 18H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 162.45, {159.56, 159.28, 159.00, 158.71}, 157.71, 157.64, 156.52, 133.54, 129.05, {119.92, 117.60, 115.27, 112.94}, 55.45, 42.98, 38.99, 38.94, 26.36 (d), 26.01 (d), 25.03 (d), 24.80 (d). ^{13}C NMR (500 MHz, d_2O) δ 163.19 {162.99, 162.71, 162.42, 162.14}, 158.54, 158.37, 155.59, 133.19, 128.46, {119.62, 117.30, 114.99, 112.67}, 72.00, 62.42, 42.64, 39.94, 39.07, 25.61, 25.27, 23.93. Note: {TFA ^{13}C NMR peaks}. HR-ESI MS (m/z) calcd [M+H$^+$] 665.47, found 665.4704, [M+2H$^+$]/2 333.24, found 333.2357. 99% pure by HPLC.

Synthesis of **N3**

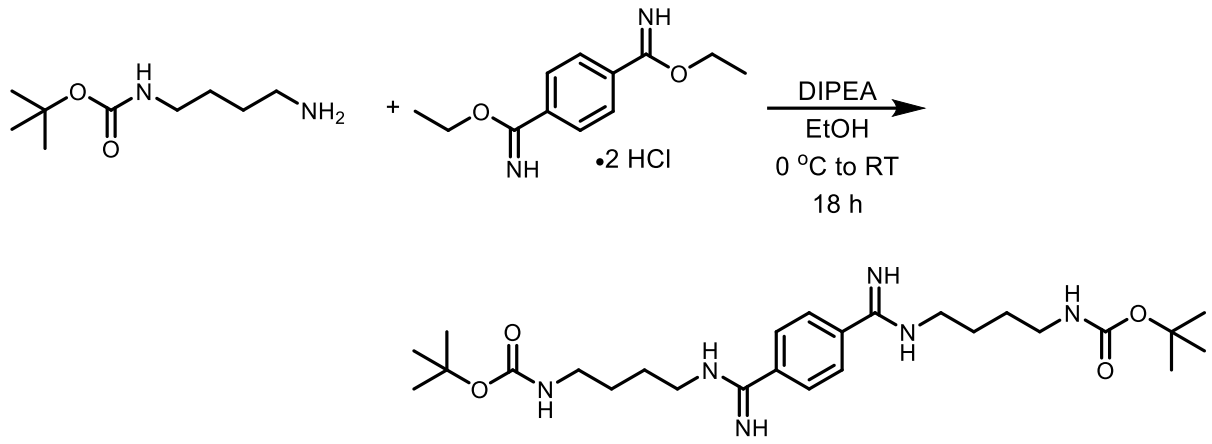


di-tert-butyl (((6-amino-1,3,5-triazine-2,4-diyl)bis(azanediyl))bis(butane-4,1-diyl))dicarbamate. To a 100 mL round-bottom flask was added 1.0 g (6.1 mmol) 4,6-dichloro-1,3,5-triazin-2-amine and 22 mL toluene. The resulting white suspension was stirred for 5 min before addition of 2.4 mL (17 mmol) *N, N*-triethylamine and 2.86 g (15.2 mmol) tert-butyl (4-aminobutyl)carbamate dissolved in 10 mL toluene. The mixture was stirred for about 5 min at room temperature and heated to 100 °C to obtain a tan suspension that was stirred for about 20 h. TLC (9 MeOH:1 NH_4OH, Rf 0.9) was used to monitor reaction progress. The resulting tan suspension was filtered to remove the triethylamine salt and dried *in vacuo* to obtain a tan oil. The crude product was purified on a Combi-Flash RediSep column (silica, 40 g) with elution gradient 100 DCM:0 MeOH to 70 DCM:30 MeOH. The fractions containing product were combined and dried *in vacuo* to obtain 2.84 g (72%) of the title compound as a thin, tan solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 6.78 (bm, 2H), 6.42 (bm, 1H), 6.13 (bs, 1H), 3.17 (bm, 4H), 2.91 (bq, J = 6.48, 4H), 1.42 (bm, 8H), 1.32 (s, 18H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 168.07, 166.10, 156.04, 77.80, 28.75, 27.55, 27.30, 26.57. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 469.32; found 469.3.

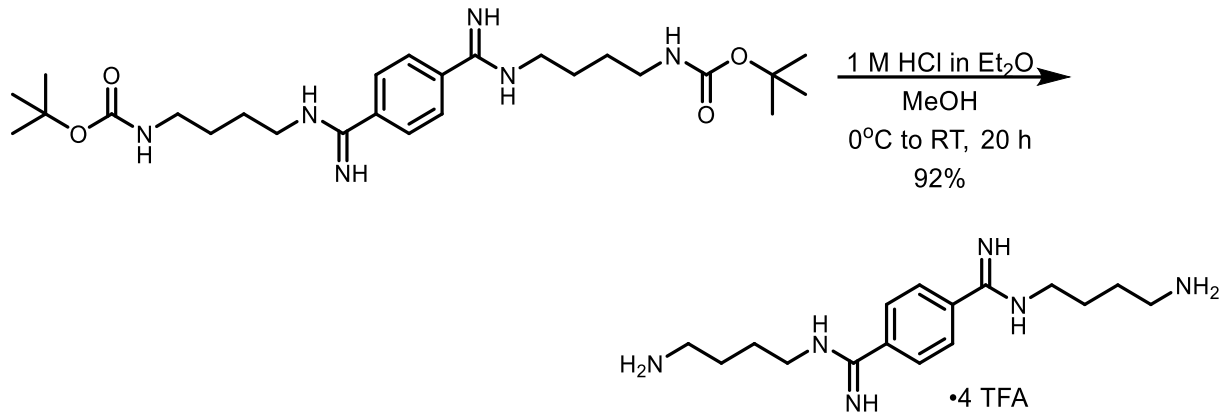


Compound N3. To a 65 mL round-bottom flask was added 998 mg (2.13 mmol) di-tert-butyl (((6-amino-1,3,5-triazine-2,4-diyl)bis(azanediyl))bis(butane-4,1-diyl))dicarbamate and 5.5 mL DCM. The solution was cooled on ice and 2.0 mL TFA was added dropwise. The tan solution was slowly warmed to room temperature and stirred for about 23 h. The tan, cloudy solution was concentrated *in vacuo* to yield the title compound as a tan oil. The product was purified using preparative HPLC (gradient H_2O (0.1% TFA):MeCN (0.1% TFA) from 100:0 to 0:100 to obtain the title compound as a white, crystalline solid. 96% pure by analytical HPLC (retention time 1.074 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.) ^{1}H NMR (500 MHz, d_2O) δ 7.51 (bs, 2H)*, 3.41 (bt, J = 6.21, 2H), 3.32 (bt, J = 6.25, 2H), 2.94 (bm, 4H) *partially exchanged amine. ^{13}C NMR (500 MHz, d_2O) δ 163.35*, 163.07*, 162.78*, 162.50*, 161.83, 159.21, 119.78*, 117.46*, 115.14*, 112.82*, 39.89, 39.15, 25.38, 25.23 *TFA. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 269.22; found 269.2.

Synthesis of **N4**

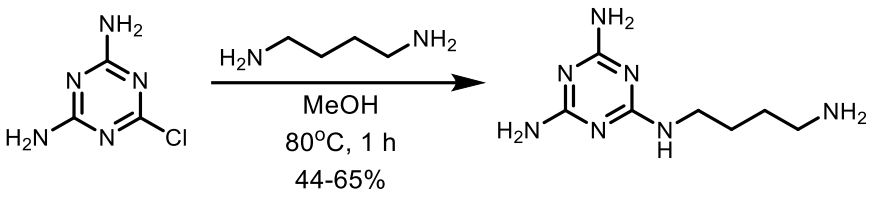


di-tert-butyl (((1,4-phenylenebis(iminomethylene))bis(azanediyl))bis(butane-4,1-diyl))dicarbamate. To a 100 mL oven-dried round-bottom flask was added 0.50 g (1.7 mmol) diethyl terephthalimidate dihydrochloride and 40 mL anhydrous ethanol. The resulting white suspension was cooled to 0 °C and kept under nitrogen atmosphere. To the white suspension was added 0.89 mL (5.1 mmol) *N*,*N*-diisopropylethylamine to obtain a colorless solution. A solution of 665 mg (3.53 mmol) tert-butyl (4-aminobuytl)carbamate dissolved in 6.0 mL anhydrous ethanol was added to the round-bottom flask dropwise. The colorless suspension was stirred at 0 °C for ~2 h, warmed slowly to 25 °C and stirred for about 24 h. TLC was used to monitor progress with ninhydrin stain (8 MeOH:2 NH_4OH, Rf 0.3).The resulting clear solution was dried *in vacuo* to obtain a tan oil. The crude product was purified on a silica column with elution gradient 9 DCM:1 MeOH to 7 DCM:3 MeOH + 5 % NH_4OH. The fractions containing product were combined and dried *in vacuo* to obtain the title compound as a white solid. ^{1}H NMR (500 MHz, DMSO-d6) δ 9.73 (bm, 6H), 7.96 (s, 4H), 6.88 (t, *J* = 5.49, 2H), 3.44 (t, *J* = 7.04, 4H), 2.97 (q, *J* = 6.42, 4H), 1.64 (bp, *J* = 7.27, 4H), 1.51 (bp, *J* = 7.18, 4H), 1.39 (s, 18H). ^{13}C NMR (500 MHz, DMSO-d6) δ 162.12, 156.13, 133.48, 129.07, 77.94, 55.40, 43.05, 28.75, 27.24, 25.15. LR-ESI MS (*m/z*) calcd for [M+H] 505.35; found 505.3.



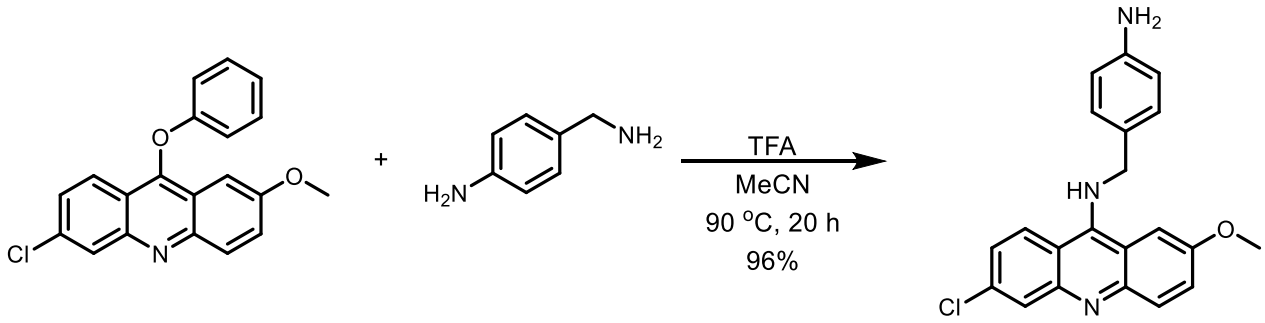
Compound N4. A 20 mL vial containing 72.0 mg (0.143 mmol) di-tert-butyl (((1,4-phenylenebis(iminomethylene))bis(azanediyl))bis(butane-4,1-diyl))dicarbamate and 4.0 mL methanol was cooled on ice and 1.0 mL 2M HCl in diethyl ether was added dropwise. The white solution was slowly warmed to room temperature and stirred for about 20 h. TLC was used to monitor reaction progress (8 MeOH:2 NH_4OH, Rf 0.8). The thick white solution was concentrated *in vacuo* to yield 59 mg (92%) of the title compound as a white solid. The product was purified using preparative HPLC (gradient H_2O (0.1% TFA):MeCN (0.1% TFA) from 100:0 to 0:100 to obtain the title compound as a white, crystalline solid. 99% pure by analytical HPLC (retention time 0.769 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.) ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.01 (bt, J = 5.86, 2H), 9.68 (bs, 2H), 9.23 (bs, 2H), 7.90 (s, 4H), 7.82 (bs, 4H), 3.42 (bq, J = 6.48, 2H), 2.84 (bq, J = 6.65, 2H), 1.67 (m, 8H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 163.03, 159.22*, 158.97*, 158.72*, 158.46*, 133.52, 129.05, 118.62, 116.25, 113.88, 42.66, 36.34, 31.30, 24.79. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 305.24; found 305.2.

Synthesis of **N8**



Compound N8.[20] To a 200 mL round-bottom flask was added 15.3 mL (152 mmol) 1,4-diaminobutane and 78 mL methanol. The mixture was stirred and heated to 100 °C in an oil bath. To the clear solution was slowly added 5.16 g (35.5 mmol) 6-chloro-1,3,5-triazine-2,4-diamine to obtain a white suspension. The suspension was stirred for about 1 h to obtain a clear solution that was stirred for an additional 3 h before cooling to room temperature. TLC (9 DCM:1 MeOH, Rf 0.05, 9 MeOH:1 NH_4OH, Rf 0.3) with ninhydrin stain was used to monitor product formation. The resulting clear solution was dried *in vacuo.* The resulting white powder was packed on silica in methanol and purified on a 2.5" x 6" silica gel column with elution gradient 9 DCM:1 MeOH to 6 DCM:3 MeOH:1 NH_4OH. The fractions containing product were combined and dried *in vacuo* to obtain 7.0 g (65 %) of the title compound as a white solid that was used in subsequent reactions without further purification. For use in DCC screen, the product was further purified using preparative HPLC (gradient H_2O (0.1% TFA):MeCN (0.1% TFA) from 100:0 to 0:100 to obtain the title compound as a white solid. 99% pure by analytical HPLC (retention time 0.809 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.) ^{1}H NMR (500 MHz, DMSO-d_6) δ 6.39 (t, J = 5.77, 1H), 6.04 (s, 2H), 5.89 (d, J = 18.13, 2H), 4.09 (s, 2H), 3.17 (m, 2H), 2.53 (t, J = 6.98, 2H), 1.46 (tt, J = 7.41, 6.99, 2H), 1.35 (m, 2H). LR-ESI MS (*m/z*) calculated 198.25, found [M+H$^+$] 198.1.

Synthesis of **N10**



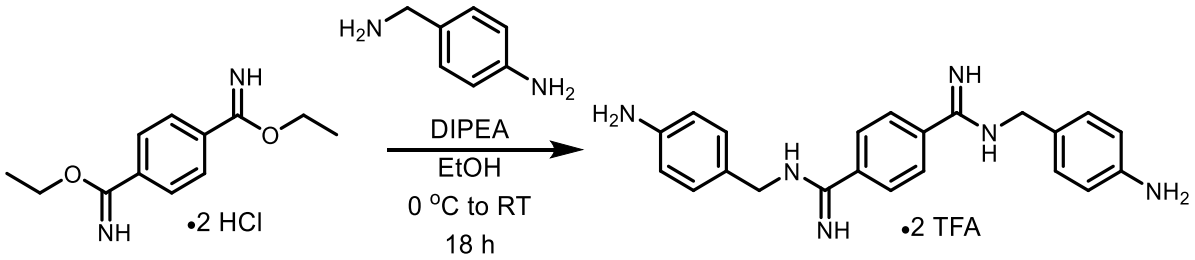
Compound N10. To an oven-dried 65 mL round-bottom flask was added 298 mg (0.887 mmol) 6-cholor-2-methoxy-9-phenoxyacridine and 13.4 mL acetonitrile. To the RBF was added 0.10 mL (0.89 mmol) 4-aminobenzylamine. The resulting yellow suspension was stirred at 90 °C and after 30 mins was added 10 drops trifluoroacetic acid. After ~20 h, the RBF was cooled to room temperature. TLC (DCM, Rf 0.1). The bright orange suspension was filtered and washed with 2x20 mL cold acetonitrile and 2x20 mL cold diethyl ether to obtain 406 mg (96%) of the title compound as a fluorescent orange solid. 99% pure by analytical HPLC (retention time 5.457 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.) ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.33 (d, J = 9.29, 1H), 7.84 (bd, J = 2.23, 1H), 7.80 (d, J = 9.27, 1H), 7.63 (d, J = 2.73, 1H), 7.38 (dd, J = 9.36, 2.68, 1H), 7.26 (dd, J = 9.30, 2.20, 1H), 7.06 (bd, J = 8.28, 2H), 7.00 (bs, 2H), 6.51 (bd, J = 8.32, 2H), 4.97 (s, 2H), 4.80 (s, 2H), 3.79 (s, 3H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 187.17, 155.34, 151.66, 150.86, 148.14, 133.85, 128.03, 127.05, 124.63, 122.93, 114.34, 55.93, 52.35. Note: Some aromatic peaks are missing and seem to be coincident. .) ^{1}H NMR (500 MHz, d_2O) δ 7.79 (d, J = 9.31, 1H), 7.61 (bd, J = 10.39, 2H), 7.53 (dd, J = 9.38, 2.59, 1H), 7.49 (bd, J = 8.51, 2H), 7.29 (dd, J = 9.31, 2.11, 1H), 7.24 (bd, J = 8.43, 2H), 6.91 (bd, J = 2.59, 1H), 4.16 (s, 2H), 3.52 (s, 3H). ^{13}C NMR (500 MHz, d_2O) δ 155.61, 152.83, 140.95, 140.83, 138.80, 131.85, 130.69, 128.82, 126.18, 125.60, 124.82, 120.83,

117.91, 117.53, 115.21, 114.47, 103.14, 55.51, 42.43. HR-ESI MS (*m/z*) calcd for [M+H$^+$] 364.12; found 364.1206. Elemental Composition 364.1217 $C_{21}H_{19}N_3OCl$ (desired [M+H$^+$] $C_{21}H_{19}N_3OCl$).

Synthesis of **N11**

See procedures in Section A.3.1, compound **3** above.

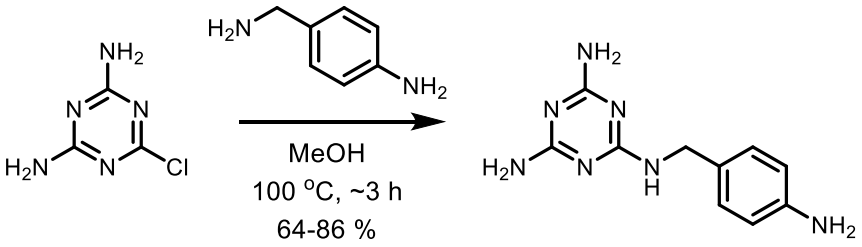
Synthesis of **N12**



Compound N12. To a 100 mL oven-dried round-bottom flask was added 0.50 g (1.7 mmol) diethyl terephthalimidate dihydrochloride and 40 mL anhydrous ethanol. The resulting white suspension was cooled to 0 °C and kept under nitrogen atmosphere. To the white suspension was added 0.89 mL (5.1 mmol) *N*,*N*-diisopropylethylamine to obtain a colorless solution. A solution of 0.40 mL (3.5 mmol) 4-(aminomethyl)aniline dissolved in 6.0 mL anhydrous ethanol was added to the round-bottom flask dropwise. The colorless suspension was stirred at 0 °C for ~2 h, warmed slowly to 25 °C and stirred for about 24 h. TLC was used to monitor progress with ninhydrin stain (8 MeOH:2 NH_4OH, Rf 0.7).The resulting yellow suspension was centrifuged at 5000 rpm for 5 min and the clear supernatant was decanted. The yellow solid was washed with 45 mL ethanol, 30 mL DCM, and 30 mL diethyl ether. The crude product was purified on a silica column with elution gradient 100% DCM to 7 DCM:3 MeOH + 5 % NH_4OH. The fractions containing product were combined and dried *in vacuo* to obtain the title compound as a yellow solid. The product was

further purified using preparative HPLC (gradient H_2O (0.1% TFA):MeCN (0.1% TFA) from 90:10 to 0:100 to obtain the title compound as a yellow, crystalline solid. 99% pure by analytical HPLC (retention time 0.865 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.) ^{1}H NMR (500 MHz, DMSO-d6) δ 10.35 (bs, 2H), 9.71 (bs, 2H), 9.31 (bs, 2H), 7.94 (s, 4H), 7.12 (bdt, *J* = 8.97, 2.42, 4H), 6.58 (bdt, *J* = 8.92, 2.41, 4H), 5.17 (bs, 4H), 4.49 (bs, 4H). ^{13}C NMR (500 MHz, DMSO-d6) δ 149.11, 129.51, 129.45, 129.15, 114.22, 46.44. Amindium carbon signal not observed, two of the aromatic peaks appear to be overlapping. HR-ESI MS (*m/z*) calcd for [M+H] 373.21; found 373.2141.

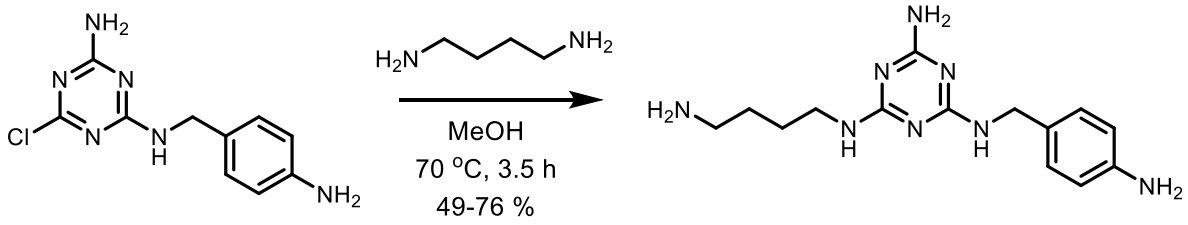
Synthesis of **N13**



Compound N13. To a 25 mL round-bottom flask equipped with a reflux condenser was added 1.20 mL (10.6 mmol) 4-(aminomethyl)aniline and 7.50 mL methanol. The clear solution was stirred at 100 °C and to the flask was slowly added 0.50 g (3.4 mmol) 6-chloro-1,3,5-triazine-2,4-diamine dissolved in 5 mL methanol. The resulting yellow suspension was stirred at 100 °C for about 3 h. TLC was used to monitor reaction progress (8 MeOH:2 NH_4OH, ninhydrin stain, Rf 0.6). The suspension was filtered to obtain a white solid that was washed with 15 mL methanol, 30 mL DCM, and 30 mL diethyl ether. The solid was dried *in vacuo* to obtain 686 mg (86 %) of the title compound as a white crystalline solid. The product was further purified using preparative HPLC (gradient H_2O (0.1% TFA):MeCN (0.1% TFA) from 90:10 to 0:100 to obtain the title compound as a white solid. 95% pure by analytical HPLC (retention time 0.987 min, analytical

HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.) ^{1}H NMR (500 MHz, DMSO-d_6) δ 6.93 (bd, J = 8.28, 2H), 6.67 (t, J = 6.32, 1H), 6.48 (dt, J = 8.77, 2.12, 2H), 6.08 (bs, 2H), 5.94 (bs, 2H), 4.87 (s, 2H), 4.25 (bd, J = 6.30, 2H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 167.62, 166.93, 147.62, 128.39, 128.21, 114.08, 43.11. HR-ESI MS (m/z) calcd for [M+H$^+$] 232.13; found 232.1305.

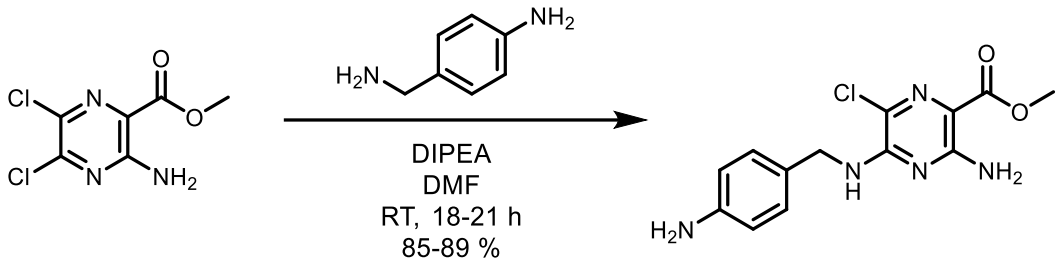
Synthesis of **N14**



Compound N14. To a 50 mL round-bottom flask equipped with a reflux condenser was added 0.60 mL (0.60 mmol) 1,4-diaminobutane and 2.0 mL methanol. The clear solution was stirred at 70 °C and to the flask was slowly added 0.301 g (1.2 mmol) N^2-(4-aminobenzyl)-6-chloro-1,3,5-triazine-2,4-diamine dissolved in 5 mL methanol. The resulting white suspension was stirred at 70 °C for about 3.5 h and dried *in vacuo*. TLC was used to monitor reaction progress (9 MeOH:1 NH_4OH, ninhydrin stain, Rf 0.4). The crude product was purified on silica gel column with elution gradient 90 DCM:10 MeOH to 70 DCM:30 MeOH + 5 % NH_4OH. The fractions containing product were combined and dried *in vacuo* to obtain 362 mg (76 %) of the title compound as a yellow powder. The product was further purified using preparative HPLC (gradient H_2O (0.1% TFA):MeCN (0.1% TFA) from 100:0 to 0:100 to obtain the title compound as a yellow, crystalline solid. 99% pure by analytical HPLC (retention time 3.952 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.) ^{1}H NMR (500 MHz, DMSO-d_6) δ 6.95 (bs, 2H), 6.81 (bm, 2H), 6.48 (d, J = 8.26, 2H), 6.41 (bs, 1H), 5.91

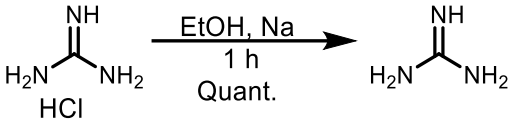
(bm, 2H), 4.88 (bs, 2H), 4.24 (bd, *J* = 6.18, 2H), 3.18 (bm, 4H), 2.70 (bs, 2H), 1.48 (bs, 4H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 166.54, 162.79, 155.99, 155.58, 147.63, 128.27, 114.05, 49.07, 36.26,* 27.02*. *Likely result of two overlapping methylene groups on the aliphatic diaminobutane chain. HR-ESI MS (*m/z*) calcd for [M+H$^+$] 303.39; found 303.2056.

Synthesis of **N15**

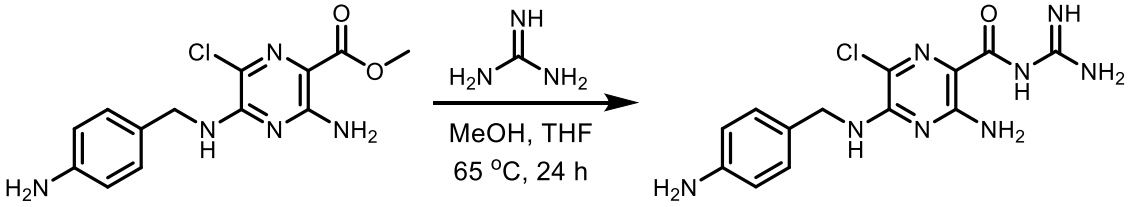


methyl 3-amino-5-((4-aminobenzyl)amino)-6-chloropyrazine-2-carboxylate. To an oven-dried 50 mL round-bottom flask was added 250 mg (1.13 mmol) methyl 3-amino-5,6-dicholoropyrazine-2-carboxylate and 8.7 mL anhydrous DMF. The dark brown solution was placed under nitrogen atmosphere and 1.0 mL (5.7 mmol) *N*,*N*-diisopropylethylamine was added. The mixture was stirred for 5 min and 0.14 mL (1.24 mmol) 4-(aminomethyl)aniline was added dropwise. The resulting dark brown solution was stirred at room temperature for about 18 h. TLC was used to monitor reaction progress (9 DCM:1 MeOH, Rf = 0.9). The mixture was concentrated *in vacuo* to obtain a brown residue that was dissolved in 200 mL ethyl acetate and transferred to a separatory funnel with 200 mL DI water. The ethyl acetate layer was washed with DI water (3 x 100 mL) and saturated brine (aqueous, 2 x 100 mL). The organic layers were combined, dried over sodium sulfate, and concentrated *in vacuo* to afford 0.31 g (89 %) of the title compound as an orange solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.91 (bt, *J* =6.11, 1H), 7.25 (bs, 2H), 7.04 (bd, *J* = 8.32, 2H), 6.49 (bdt, *J* = 8.81, 2.12, 2H), 4.96 (s, 2H), 4.39 (d, *J* = 6.12, 2H), 3.73 (s, 3H). ^{13}C

NMR (500 MHz, DMSO-d_6) δ 166.38, 156.10, 151.60, 148.12, 129.16, 126.25, 120.10, 114.06, 108.83, 51.61, 43.92. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 308.09; found 308.1.



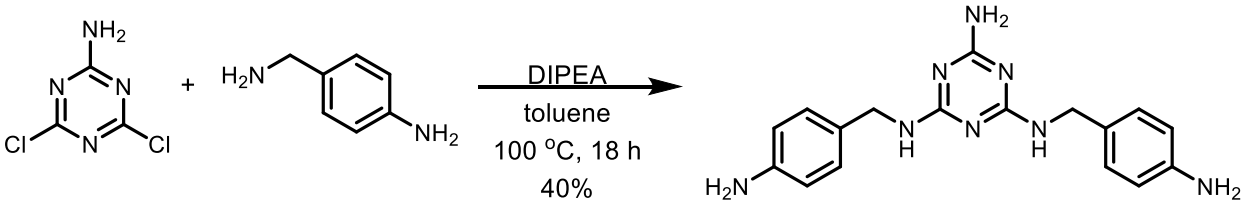
Guanidine. To a 100 mL round-bottom flask was added 20 mL dry ethanol followed by slow addition of 1.20 g (52.3 mmol) sodium metal. To the clear solution was added 5.00 g (52.3 mmol) guanidine hydrochloride slowly followed by 5.0 mL dry ethanol. The resulting white suspension was stirred for about 1 h, the salt was filtered off, and the clear filtrate was dried *in vacuo* to obtain the title compound as a tan oil. TLC (9 DCM:1 MeOH, Rf 0.8) was used to monitor reaction progress and the reaction yielded 1.6 g (74%) of the title compound as a white solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 5.14 (bm, 4H), 2.55 (s, 1H).



Compound N15. To a 15 mL round-bottom flask was added 84.2 mg (0.274 mmol) methyl 3-amino-5-((4-aminobenzyl)amino)-6-chloropyrazine-2-carboxylate and 2.4 mL THF. The dark orange solution was placed under nitrogen atmosphere and 0.81 mL 2M guanidine in methanol solution was added. The resulting orange solution was stirred at 65 °C for about 24 h. TLC was used to monitor reaction progress (8 DCM:2 MeOH, Rf = 0.4). The mixture was concentrated *in vacuo* to obtain an orange oil that was purified using preparative HPLC (gradient H_2O (0.1% TFA):MeCN (0.1% TFA) from 100:0 to 0:100; analytical HPLC (retention time 4.274 min,

analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.). ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.03 (bd, J = 8.23, 2H), 7.00 (s, 1H), 6.50 (bd, J = 8.36, 2H), 5.51 (bs, 5H)*, 4.93 (s, 2H), 4.37 (s, 2H). *partially exchanged amine ^{13}C NMR (500 MHz, DMSO-d_6) δ 160.25, 158.80, 155.23, 150.08, 147.94, 128.99, 127.02, 118.161, 114.09, 43.74. LR-ESI MS (*m/z*) calcd for [M+H^{+}] 335.11; found 335.1.

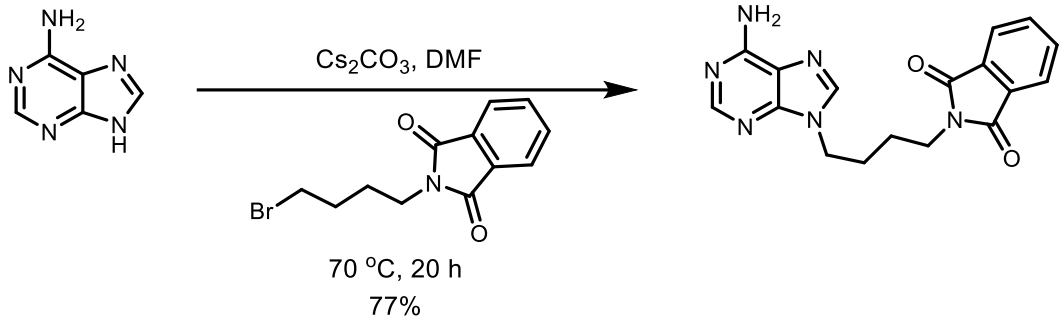
Synthesis of **N16**



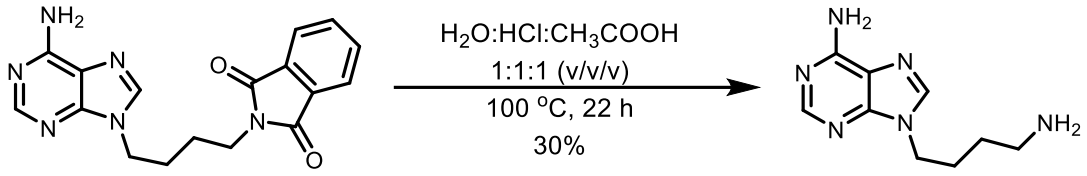
Compound N16. To a 100 mL round-bottom flask was added 1.03 g (6.24 mmol) 4,6-dichloro-1,3,5-triazin-2-amine and 32 mL toluene. To the resulting white suspension was added 2.4 mL (17 mmol) *N*, *N*-triethylamine and 1.7 mL (15 mmol) 4-(aminomethyl)aniline. The mixture was stirred for 5 min at room temperature and heated to 100 °C to obtain a tan suspension that was stirred for about 18 h. TLC (9 MeOH:1 NH_4OH, Rf 0.9) was used to monitor reaction progress. The thick orange suspension was sonicated, dissolved in methanol, and stirred at 100 °C for an additional 3 h. The suspension was filtered and the filtrate was dried *in vacuo* to obtain a thick orange solid. The crude product was purified on a silica column with gradient 100 DCM:0 MeOH to 70 DCM:30 MeOH. Fractions containing product were dried *in vacuo* to yield 0.82 g (40%) of the title compound as an orange solid. The product was purified using preparative HPLC (gradient H_2O (0.1% TFA):MeCN (0.1% TFA) from 100:0 to 0:100 to obtain the title compound as a white solid. 99% pure by analytical HPLC (retention time 3.459 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.) ^{1}H NMR (500 MHz,

DMSO-d_6) δ 8.74 (bs, 1H)*, 8.69 (bt, *J* = 6.10, 1H)*, 8,47 (bs, 1H)*, 7.88 (bs, 1H)*, 7.29 (bt, *J* = 8.41, 1H), 7.24 (d, *J* = 8.02, 1H), 7.07 (bt, *J* = 8.99, 1H), 7.02 (d, *J* = 7.99, 1H), 4.46 (bdd, *J* = 15.95, 5.89, 2H). *Partially exchanged amine. ^{13}C NMR (500 MHz, DMSO-d_6) δ 159.21, 158.95, 156.69, 129.21, 129.10, 120.35, 43.77. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 337.19; found 337.2.

Synthesis of **N17**

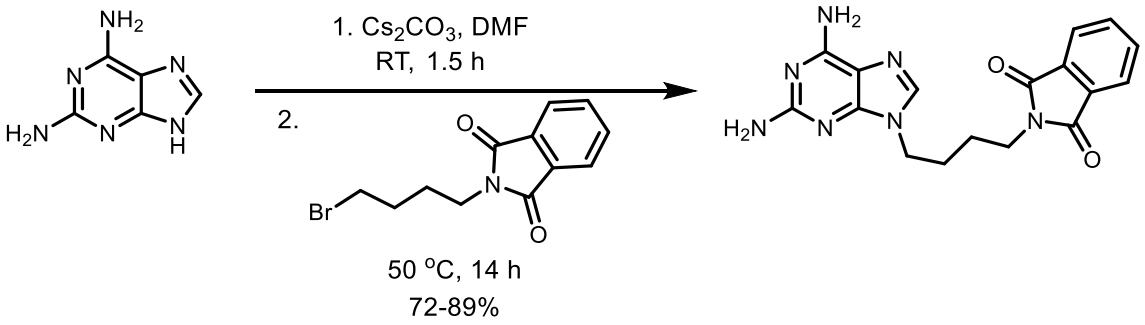


2-(4-(6-amino-9H-purin-9-yl)butyl)isoindoline-1,3-dione. To an oven-dried 200 mL 24/40 round-bottom flask was added 3.03 g (22.4 mmol) 9*H*-purine-6-amine and 83 mL anhydrous DMF stored over 4Å molecular sieves. To the resulting tan suspension was added 8.16 g (25.0 mmol) cesium carbonate and 7.78 g (27.6 mmol) 2-(4-bromobutyl)isoindoline-1,3-dione . The mixture was stirred at 70 °C under nitrogen atmosphere for about 20 h to obtain a yellow suspension. TLC (8 DCM:2 MeOH, Rf 0.8) was used to monitor reaction progress. The white solid was filtered off and the yellow filtrate was partitioned between water and ethyl acetate (100 mL each). The resulting yellow suspension was filtered and washed with 2x500 mL DI water and 200 mL ethyl ether. The solid was dried *in vacuo* to obtain 5.75 g (77%) of the title compound as a yellow solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.12 (s, 1H), 8.07 (s, 1H), 7.84 (m, 4H), 7.17 (s, 2H), 4.16 (t, *J* = 6.96, 2H), 3.60 (t, *J* = 6.92, 2H), 1.83 (bp, *J* = 7.39, 2H), 1.55 (bp, *J* = 7.38, 2H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 168.45, 156.39, 152.92, 152.79, 149.98, 141.29, 134.81, 132.10, 123.46, 119.18, 42.84, 37.31, 25.54, 25.44. HR-ESI MS (*m/z*) calcd for [M+H$^+$] 337.14; found 337.1427.

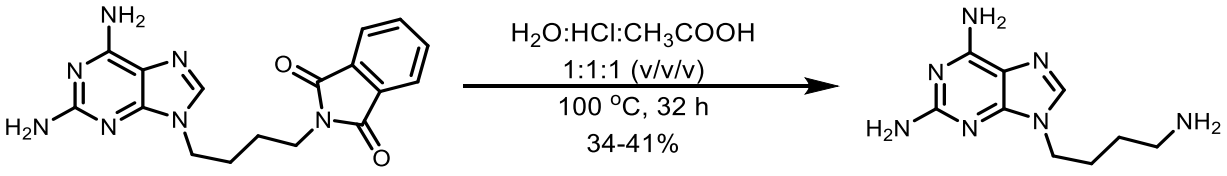


Compound N17. To a 65 mL round-bottomed flask was added 3.60 g (10.7 mmol) 2-(4-(6-amino-9H-purin-9-yl)butyl)isoindoline-1,3-dione and 15.0 mL of a 1:1:1 (v/v/v) solution of water : glacial acetic acid : concentrated hydrochloric acid. The solution was stirred under nitrogen atmosphere at 100 °C for about 22 h and cooled to room temperature while stirring. TLC (9 MeOH:1 NH_4OH, Rf 0.05) was used to monitor reaction progress. The white precipitate was filtered off and the filtrate was recycled until clear. The remaining solvents were removed *in vacuo* to obtain a tan colored solid. When a white precipitate formed in the flask during *in vacuo* evaporation, it was filtered off. The crude product was dissolved in 250 mL DI water and basified with 3M aqueous KOH (about 26 mL) to pH 12 and saturated with NaCl powder. The suspension was transferred to a separatory funnel and extracted 6 times with 100 mL of 10:1 (v/v) CH_3Cl:*n*-BuOH. The organic layers were combined and concentrated *in vacuo*. The resulting white solid was frozen and dried on a lyophilizer to yield 673 mg (30%) of the title compound as a white solid that was used in subsequent reactions without further purification. For use in DCC screen, the product was further purified using preparative HPLC (gradient H_2O (0.1% TFA):MeCN (0.1% TFA) from 100:0 to 0:100 to obtain the title compound as a white solid. 99% pure by analytical HPLC (retention time 0.857 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.14 (s, 1H), 8.13 (s, 1H), 7.17 (s, 2H), 4.13 (t, J = 7.12, 2H), 2.53 (t, J = 6.89, 2H), 1.82 (bp, J = 7.42, 2H), 1.29 (bp, J = 7.31, 2H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 156.41, 152.80, 150.02, 141.30, 119.21, 43.32, 41.58, 30.76, 27.45. HR-ESI MS (*m/z*) calcd for [M+H$^+$] 207.14; found 207.1352.

Synthesis of **N18**



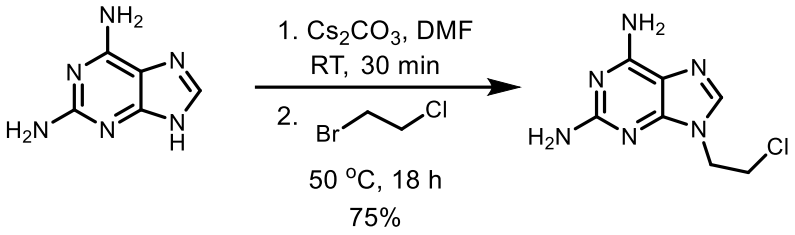
2-(4-(2,6-Diamino-9*H*-purin-9-yl)butyl)isoindoline-1,3-dione.[21] To an oven-dried 250 mL 24/40 round-bottom flask was added 5.03 g (33.5 mmol) 9*H*-purine-2,6-diamine and 125 mL anhydrous DMF stored over 4Å molecular sieves. To the resulting tan suspension was added 12.4 g (38 mmol) cesium carbonate and the mixture was stirred at 50 °C under nitrogen atmosphere for 1.5 h. To the tan suspension was added 11.7 g (41 mmol) 2-(4-bromobutyl)isoindoline-1,3-dione and the mixture was stirred at 50 °C for an additional 14 h. TLC (8 DCM:2 MeOH, Rf 0.8) was used to monitor reaction progress and the reaction yielded 10.5 g (89%) of the title compound as a yellow solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.85 (m, 4H), 7.69 (s, 1H), 6.60 (bs, 2H), 5.74 (bs, 2H), 3.97 (t, J = 7.01, 2H), 3.60 (t, J = 6.99, 2H), 1.76 (brp, J = 7.42, 2H), 1.56 (brp, J = 7.40, 2H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 168.45, 160.71, 156.55, 152.20, 137.89, 134.85, 132.09, 123.50, 113.67, 42.31, 37.39, 27.32, 25.71. HR-ESI MS (*m/z*) calcd for [M+H$^+$] 352.1516(5); found 352.1515.



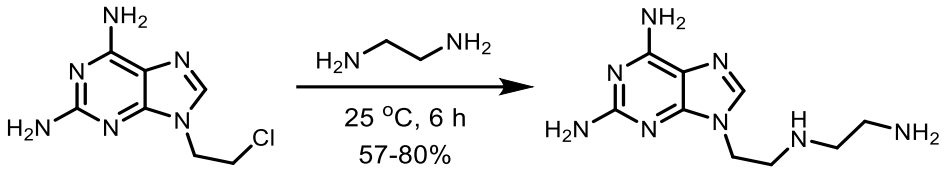
Compound N18. To a 50 mL 24-40 round-bottomed flask was added 439 mg (1.25 mmol) 2-(4-(2,6-diamino-9*H*-purin-9-yl)butyl)isoindoline-1,3-dione and 6.0 mL of a 1:1:1 (v/v/v) solution of

water : glacial acetic acid : concentrated hydrochloric acid. The light yellow solution was stirred under nitrogen atmosphere at 100 °C for about 32 h and cooled to room temperature. TLC (9 MeOH:1 NH_4OH, Rf 0.2) was used to monitor reaction progress. The white precipitate was filtered off and remaining solvents were removed *in vacuo*. When a white precipitate formed in the flask during *in vacuo* evaporation, it was filtered off. The crude product was a tan to yellow-tinted oil. The crude was dissolved in 100 mL DI water and the solvents were removed *in vacuo*. This process was repeated twice more (three times total). The crude was dissolved in about 5 mL water, basified with 3M aqueous KOH (about 26 mL) and saturated with NaCl powder. The suspension was transferred to a separatory funnel and extracted 6 times with 100 mL (600 mL total) 10:1 (v/v) CH_3Cl:*n*-BuOH. The organic layers were combined and concentrated *in vacuo*. The resulting light yellow solid was dissolved in 10 mL DI water, frozen, and dried on a lyophilizer to yield 115 mg (41%) of the title compound as a light yellow solid that was used in subsequent reactions without further purification. For use in DCC screen, the product was further purified using preparative HPLC (gradient H_2O (0.1% TFA):MeCN (0.1% TFA) from 100:0 to 0:100 to obtain the title compound as a white solid. 99% pure by analytical HPLC (retention time 0.988 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.) ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.70 (s, 1H), 6.60 (s, 2H), 5.74 (s, 2H), 3.94 (t, J = 7.14, 2H), 2.54 (t, J = 6.93, 2H), 1.75 (p, J = 7.40, 2H), 1.30 (p, J = 7.31, 2H), 1.24 (bs, 1H) *from partially exchanged amine. ^{13}C NMR (500 MHz, DMSO-d_6) δ 160.19, 156.55, 152.23, 137.93, 113.71, 42.78, 41.62, 30.73, 27.44. HR-ESI MS (*m/z*) calcd for [M+H$^+$] 222.15; found 222.1460. Elemental composition calcd for 222.1460: $C_9H_{16}N_7$ (desired [M+H$^+$] $C_9H_{16}N_7$).

Synthesis of **N19**

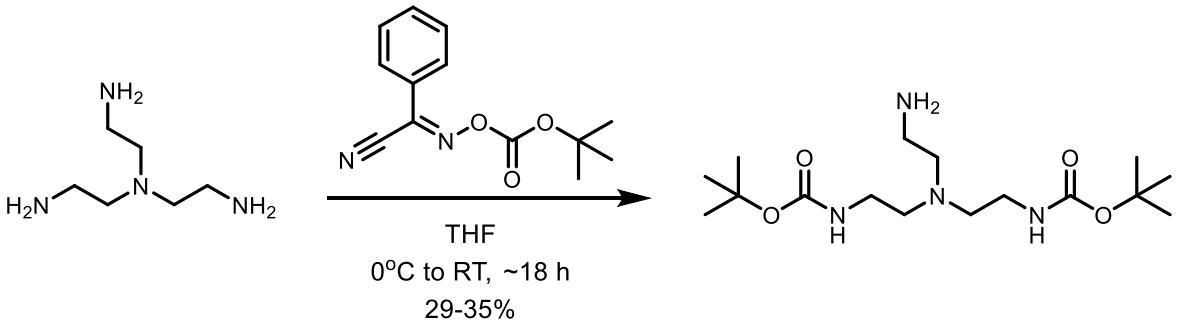


9-(2-Chloroethyl)-9H-purine-2,6-diamine.[21] To an oven-dried 250 mL 24/40 round-bottom flask was added 2.01 g (13.4 mmol) 9*H*-purine-2,6-diamine and 50 mL anhydrous DMF stored on 4Å molecular sieves. To the resulting tan suspension was added 6.50 g (20.0 mmol) cesium carbonate and the mixture was stirred at 50 °C under nitrogen atmosphere for 30 min. To the tan suspension was added 4.01 mL (48.6 mmol) ethylene bromochloride and the mixture was stirred at 50 °C for an additional 18 h. TLC (9 DCM:1 MeOH, Rf 0.4) was used to monitor reaction progress. The reaction was cooled to room temperature, 30 mL DCM was added, and the solution was sonicated and transferred to centrifuge tubes and spun at 5,000 rpm for 5 min. The supernatant was pipetted off and the resulting yellow solid was washed with 2 x 25 mL water, centrifuged, and the supernatant was removed. The resulting yellow solid was dried on a lyophilizer. The reaction yielded 2.13 g (75%) of the title compound as a tan solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.72 (s, 1H), 6.67 (bs, 2H), 5.80 (bs, 2H), 4.29 (t, J = 5.98, 2H), 3.98 (t, J = 6.00, 2H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 160.79, 156.61, 152.21, 138.07, 113.56, 44.71, 113.56. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 213.1; found 213.1.

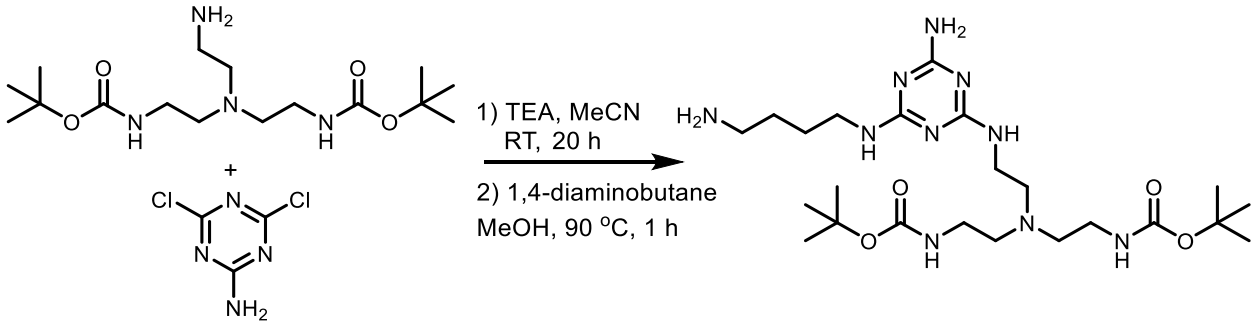


Compound N19. To a 50 mL 24-40 round-bottomed flask was added 1.0 g (4.7 mmol) 9-(2-chloroethyl)-9H-purine-2,6-diamine and 9.4 mL (141 mmol) of ethylenediamine. The light yellow solution was stirred under nitrogen atmosphere at room temperature for 6 h. TLC (8 MeOH:2 NH_4OH, Rf 0.6) with ninhydrin stain was used to monitor reaction progress. To the resulting solution was added 10 mL ethanol and the solution was evacuated *in vacuo* to obtain a thick yellow oil with crystalline precipitate that was triturated with 3x25 mL ethanol and 30 mL diethyl ether (centrifuging at 5000 rpm for 5 min to settle product), and freeze dried on a lyophilizer to obtain 0.88 g (80%) of the title compound as a tan powder that was used in subsequent reactions without further purification. For use in DCC screen, the product was further purified using preparative HPLC (gradient H_2O (0.1% TFA):MeCN (0.1% TFA) from 100:0 to 0:100 to obtain the title compound as a white solid. 99% pure by analytical HPLC (retention time 0.778 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.). ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.70 (bs, 1H), 6.61 (bs, 2H), 5.74 (bs, 2H), 3.99 (t, J = 6.22, 2H), 2.85 (t, J = 6.22, 2H), 2.65 (t, J = 5.69, 2H), 2.58 (t, J = 5.84, 2H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 160.64, 156.55, 152.25, 138.34, 113.60, 49.35, 48.78, 43.06, 41.82. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 237.16; found 237.2.

Synthesis of **N20**

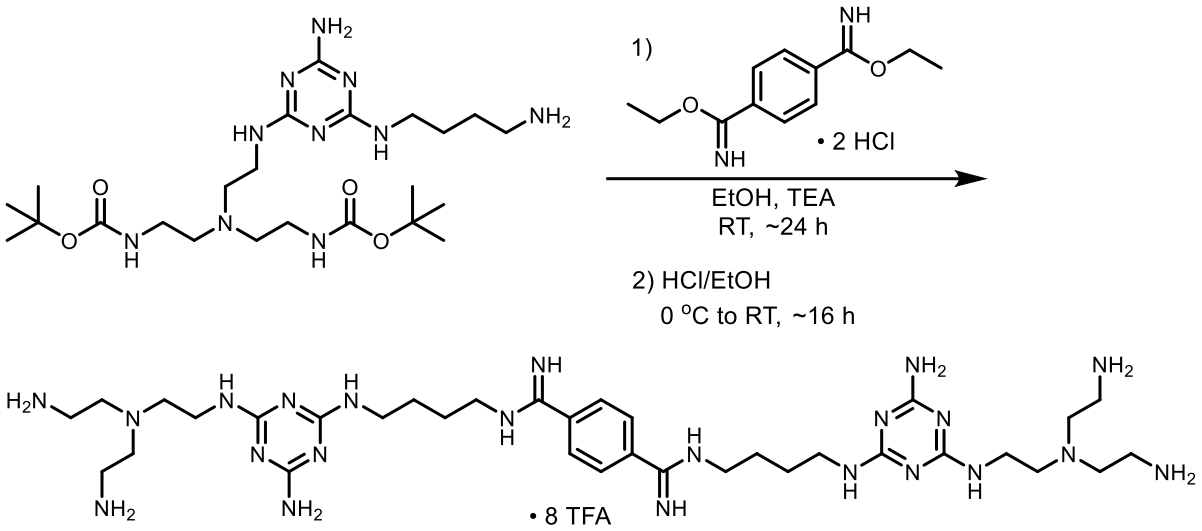


di-tert-butyl (((2-aminoethyl)azanediyl)bis(ethane-2,1-diyl)dicarbamate:[22] To a 500 mL round-bottom flask was added 5.0 mL (33 mmol) tris(2-aminoethyl)amine and 125 mL tetrahydrofuran. The resulting clear solution was cooled to 0 °C in an ice bath. In a 250 mL Erlenmeyer flask was mixed 16.09 g (65.3 mmol) 2-(Boc-oxyimino)-2-phenylacetonitrile (Boc-ON) and 175 mL tetrahydrofuran to form a clear solution that was added dropwise to the round-bottom flask over 4 h while stirring. The resulting yellow solution was warmed slowly to room temperature and stirred for an additional 14 h. Thin layer chromatography on silica gel plates (8 DCM:2 MeOH, rf 0.1, ninhydrin stain) was used to monitor reaction progress. The yellow solution was concentrated *in vacuo* to obtain a yellow oil that was dissolved in 200 mL ethyl acetate. The organic layer was transferred to a 1 L separatory funnel and washed with 2 x 200 mL 0.5 M NaOH (aq). The combined aqueous layer was diluted with 200 mL brine and extracted with 2 x 200 mL ethyl acetate. The organic phases were combined, dried over sodium sulfate, and concentrated *in vacuo* to obtain a yellow oil. The crude product was purified in two batches on a Combi-Flash RediSep column (silica, 40 g – 2 different columns) with elution gradient 100 DCM:0 MeOH to 70 DCM:30 MeOH. The fractions containing product were combined and dried *in vacuo* to obtain 4.07 g (35%) of the title compound as a light yellow oil. ^{1}H NMR (500 MHz, DMSO-d_6) δ 6.73 (bt, J = 5.70, 2H), 2.95 (q, J = 6.25 Hz, 4H), 2.50-2.48 (m, 2H), 2.45-2.36 (m, 6H), 1.38 (s, 18H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 156.17, 77.90, 58.11, 54.52, 49.06, 38.90, 28.73.



di-tert-butyl (((2-((4-amino-6-((4-aminobutyl)amino)-1,3,5-triazin-2-yl)amino)ethyl)azanediyl)bis(ethane-2,1-diyl))dicarbamate. To a 100 mL round-bottom flask was added 0.66 g (4.0 mmol) 4,6-dichloro-1,3,5-triazine-2-amine, 9.0 mL acetonitrile, and 0.70 mL (5.0 mmol) triethylamine. The resulting suspension was stirred at room temperature for 5 min to obtain a clear solution. To the RBF was added 1.66 g (4.8 mmol) di-tert-butyl (((2-aminoethyl)azanediyl)bis(ethane-2,1-diyl)dicarbamate dissolved in 1.0 mL acetonitrile to obtain a transparent yellow solution that was stirred at room temperature. Thin layer chromatography on silica gel plates (9 DCM:1 MeOH, rf 0.7, iodine chamber stain) was used to monitor reaction progress. After about 16 h, an additional 0.3 mL (2.2 mmol) triethylamine and 5.0 mL acetonitrile were added and the resulting milky white solution was stirred for an additional 24 h. The solution was concentrated *in vacuo* to obtain a thin yellow solid that was mixed with 30.0 mL MeOH and 2.0 mL 1,4-diaminobutane (2.0 mmol). The mixture was placed under nitrogen atmosphere and stirred at 90 °C for about 4 h. TLC (8 MeOH:2 NH_4OH, 1,4-diaminobutane rf 0.05, pdt rf 0.8). After no more starting material was observed, the solution as concentrated *in vacuo* to obtain a clear oil. The crude product was purified on a 1" silica gel column with elution gradient 90 DCM:10 MeOH to 70 DCM:30 MeOH:2 NH_4OH. The fractions containing product were combined and dried *in vacuo* to obtain 0.956 g (44%) of the title compound as a thin white solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 6.74 (bs, 2H), 6.60-6.20 (bm, 2H), 5.97 (bs, 1H), 5.79 (bs, 1H), 3.38-3.28 (bm, 2H), 3.24-3.12 (bm, 4H), 3.00-2.92 (bm, 4H), 2.55 (t, J = 6.98, 2H), 2.46 (t, J =

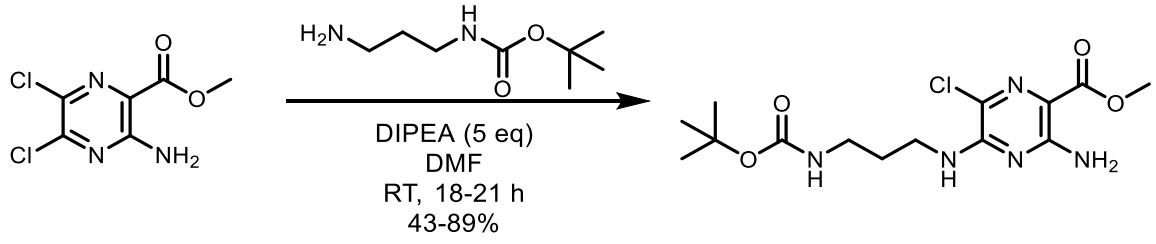
6.72, 4H), 1.52-1.42 (bm, 2H), 1.38 (s, 18H), 1.38-1.32 (bm, 2H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 166.55, 156.16, 77.99, 54.44, 41.72, 38.87, 38.49, 30.62, 28.73, 27.38. HR-ESI MS (*m/z*) calculated [M+H$^+$] 527.38, found 527.3781, calculated [M+H$^+$ - Boc] 427.33, found 427.3259, calculated [M+H$^+$ - 2 Boc] 327.27, found 327.2730. Elemental Composition 527.3782 $C_{23}H_{47}N_{10}O_4$ (desired [M+H$^+$] $C_{23}H_{47}N_{10}O_4$).



Compound N20. To an oven-dried 50 mL round-bottom flask was added 0.322 g (1.1 mmol) diethyl terephthalimidate hydrochloride, 14.0 mL ethanol, and 0.58 mL (3.3 mmol) freshly distilled triethylamine. The resulting suspension was stirred at room temperature for 5 min to obtain a clear solution. To the RBF was added 1.30 g (2.5 mmol) di-tert-butyl (((2-((4-amino-6-((4-aminobutyl)amino)-1,3,5-triazin-2-yl)amino)ethyl)azanediyl)bis(ethane-2,1-diyl))dicarbamate dissolved in 5 mL ethanol to obtain a tan solution that was stirred at room temperature for about 16 h. Thin layer chromatography on silica gel plates (8 MeOH:2 NH_4OH, rf 0.4, ninhydrin stain) was used to monitor reaction progress. The solution was concentrated *in vacuo* to obtain a thin tan solid that was dissolved in 24 mL 2 M ethanolic HCl added dropwise over ice. The mixture was placed under nitrogen atmosphere and stirred at room temperature for

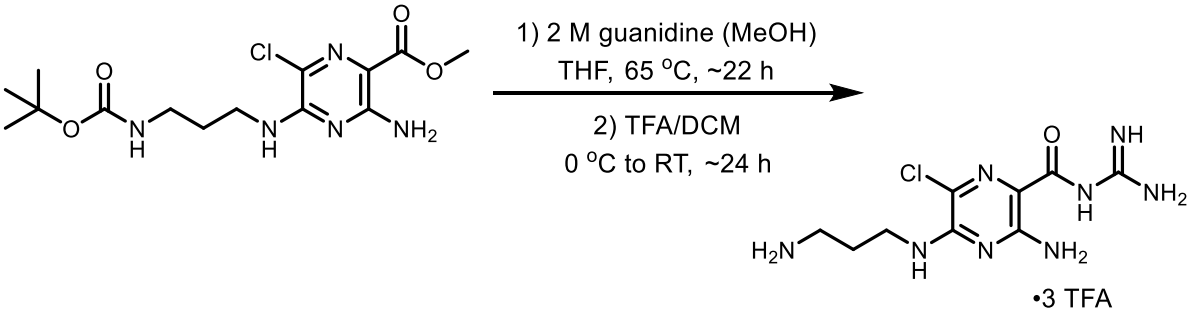
about 12 h. TLC (9 MeOH:1 NH_4OH, Rf 0.05, ninhydrin stain). After no more starting material was observed, the solution as concentrated *in vacuo* to obtain a tan oil. The compound was purified by preparatory HPLC (see general methods) using gradient from 0 to 100% acetonitrile in water, 0.1% TFA. The fractions containing product were combined and dried *in vacuo* to obtain 128.4 mg of the title compound as a thin white solid. No yield is reported because the compound was purified as needed, a small amount each time. 99.9% pure by analytical HPLC (retention time 4.038 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min / retention time 4.883 min, analytical HPLC gradient methanol in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min). The solution was stored as a master 500 mM solution in molecular biology grade water for biological studies. ^{1}H NMR (500 MHz, D_2O) δ 7.80 (s, 4H), 3.54-3.31 (bm 12H), 3.10-3.03 (bm, 8H), 2.90-2.82 (bm, 8H), 2.81-2.73 (bm, 4H), 1.79-1.71 (bm, 4H), 1.70-1.62 (bm, 4H). ^{13}C NMR (500 MHz, D_2O) δ 163.29, {163.27, 163.01, 162.73, 162.45,} 133.27, 128.52, {119.77, 117.45, 115.13, 112.81,} 51.15, 50.09, 42.68, 40.18, 36.74, 36.17, 25.64 , 24.06 .TFA peaks are included in {}. Note: melamine carbon peaks are thought to be overlapped. MALDI-MS (*m/z*) calculated [$M+H^+$] 782.05, found 782.072.

Synthesis of **N21**



methyl 3-amino-5-((3-tert-butoxycarbonyl)amino)propyl)amino)-6-chloropyrazine-2-carboxylate. To an oven-dried 25 mL round-bottom flask was added 262.6 mg (1.183 mmol) methyl 3-amino-5,6-dicholoropyrazine-2-carboxylate and 10.0 mL anhydrous DMF. The dark

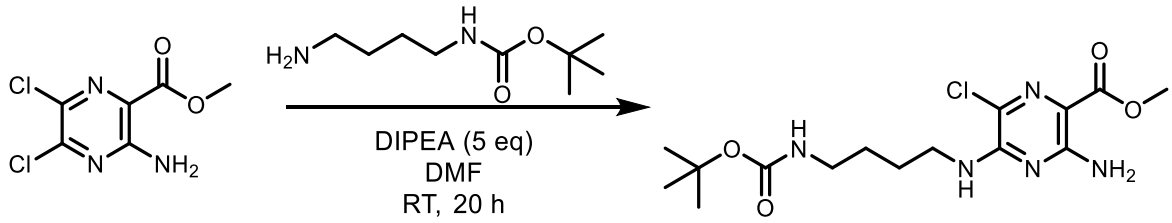
brown solution was placed under nitrogen atmosphere and 1.0 mL (5.7 mmol) *N*,*N*-diisopropylethylamine was added. The mixture was stirred for 5 min and 0.25 mL (1.4 mmol) tert-butyl (3-aminopropyl)carbamate was added dropwise to the round-bottom flask. The resulting dark brown solution was stirred at room temperature for about 18 h. TLC was used to monitor reaction progress (3 Hex:2 EtOAc, Rf = 0.4). The mixture was concentrated *in vacuo* to obtain a brown residue that was dissolved in 50 mL ethyl acetate and transferred to a separatory funnel with 50 mL DI water. The ethyl acetate layer was washed with DI water (3 x 50 mL) and saturated brine (aqueous, 2 x 50 mL). The organic layers were combined, dried over sodium sulfate to obtain a yellow liquid, and concentrated *in vacuo* to afford 361 mg (89%) of the title compound as a thin orange film. ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.51 (bt, *J* =5.81, 1H), 7.24 (bs, 2H), 6.84 (bt, *J* = 5.91, 1H), 3.74 (s, 3H), 3.37 (q, *J* = 6.32, 2H), 2.98 (q, *J* = 6.47, 2H), 1.68 (p, *J* = 6.72, 2H), 1.39 (s, 9H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 165.98, 155.71, 154.80, 151.37, 119.78, 108.30, 77.60, 51.19, 38.28, 37.09, 28.83, 28.30. LR-ESI MS (*m/z*) calcd for [M+H^{+}] 360.14; found 360.1.



Compound N21. To a 35 mL round-bottom flask was added 361 mg (1.00 mmol) methyl 3-amino-5-((3-((tert-butoxycarbonyl)amino)propyl)amino)-6-chloropyrazine-2-carboxylate and 7.5 mL THF. The dark orange solution was placed under nitrogen atmosphere and 2.5 mL 2M guanidine in methanol solution was added. The resulting orange solution was stirred at 65 °C for about 22 h. TLC was used to monitor reaction progress (8 DCM:2 MeOH, Rf = 0.4). The mixture was

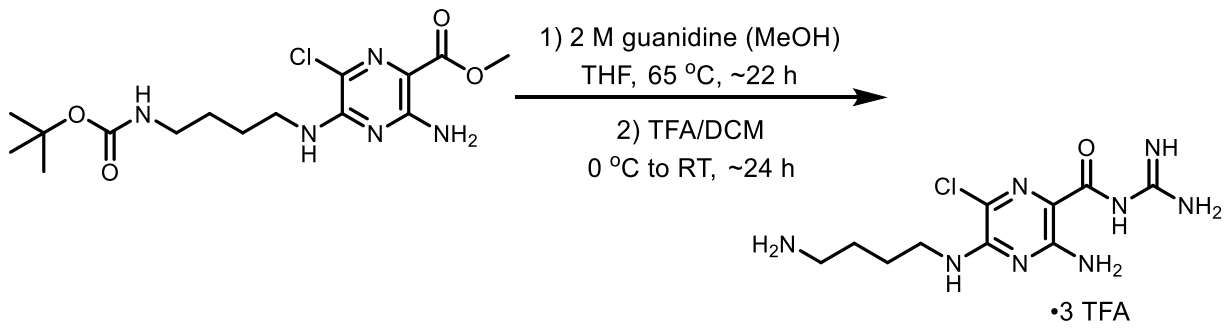
concentrated *in vacuo* to obtain an orange oil. To the orange oil was added 2.0 mL DCM. The resulting solution was cooled on ice, 2.0 mL TFA was added dropwise, and the solution was stirred at room temperature for about 24 h. TLC was used to monitor reaction progress (8 DCM:2 MeOH, Rf = 0.1). The deep orange solution was dried *in vacuo* to obtain a red oil that was purified using preparative HPLC (gradient H_2O (0.1% TFA):MeCN (0.1% TFA) from 100:0 to 0:100 to yield 140.89 mg (45% over two steps, only half of the deprotected product was purified by preparative HPLC) of the title compound as an orange solid. 99% pure by analytical HPLC (retention time 3.735 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.). ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.62 (s, 1H), 8.86 (bs, 2H), 8.58 (bs, 2H), 8.07 (bt, J = 5.94, 1H), 7.92 (bs, 3H), 7.58 (bs, 2H), 3.46 (q, J = 6.23, 2H), 2.84 (sextet, J = 6.26, 2H), 1.97 (p, J = 6.94, 2H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 165.79, 159.64*, 159.38*, 159.12*, 158.86*, 156.17, 155.61, 152.77, 120.79*, 120.64, 118.41*, 116.06*, 113.70*, 108.83, 38.27, 37.17, 26.96. *TFA LR-ESI MS (*m/z*) calcd for [M+H$^+$] 287.11; found 287.1.

Synthesis of **N22**



methyl 3-amino-5-((4-((tert-butoxycarbonyl)amino)butyl)amino)-6-chloropyrazine-2-carboxylate. To an oven-dried 35 mL round-bottom flask was added 263.8 mg (1.188 mmol) methyl 3-amino-5,6-dicholoropyrazine-2-carboxylate and 10.0 mL anhydrous DMF. The dark brown solution was placed under nitrogen atmosphere and 1.0 mL (5.7 mmol) *N*,*N*-diisopropylethylamine was added. The mixture was stirred for 5 min and 0.30 mL (1.6 mmol) tert-

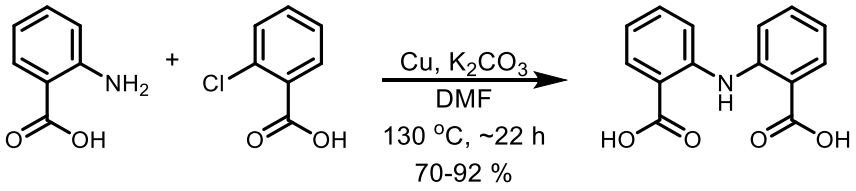
butyl (4-aminobutyl)carbamate was added dropwise to the round-bottom flask. The resulting dark brown solution was stirred at room temperature for about 20 h. TLC was used to monitor reaction progress (3 Hex:2 EtOAc, Rf = 0.5). The mixture was concentrated *in vacuo* to obtain a brown residue that was dissolved in 50 mL ethyl acetate and transferred to a separatory funnel with 50 mL DI water. The ethyl acetate layer was washed with DI water (3 x 50 mL) and saturated brine (aqueous, 2 x 50 mL). The organic layers were combined, dried over sodium sulfate to obtain a yellow liquid, and concentrated *in vacuo* to afford 454 mg of the title compound as an orange oil (contains ethyl acetate). ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.54 (bt, J = 5.81, 1H), 7.23 (bs, 2H), 6.80 (bt, J = 5.78, 1H), 3.74 (s, 3H), 2.97 (q, J = 6.66, 2H), 1.53 (bp, J = 7.80, 2H), 1.40 (bm, 4H), 1.37 (s, 9H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 166.43, 156.18, 156.07, 151.83, 120.18, 108.62, 77.87, 51.61, 28.75, 27.49, 26.37, 21.24. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 374.16; found 344.1 (M-OCH_3).



Compound N22. To a 15 mL round-bottom flask was added 421 mg (1.13 mmol) methyl 3-amino-5-((4-aminobenzyl)amino)-6-chloropyrazine-2-carboxylate and 8.4 mL THF. The orange solution was placed under nitrogen atmosphere and 2.8 mL 2M guanidine in methanol solution was added. The resulting orange solution was stirred at 65 °C for about 24 h. TLC was used to monitor reaction

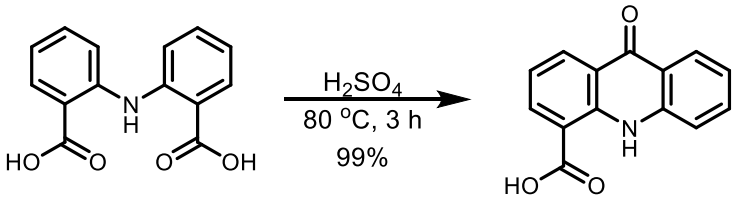
progress (8 DCM:2 MeOH, Rf = 0.2). The mixture was concentrated *in vacuo* to obtain an orange oil. To the orange oil was added 2.0 mL DCM. The resulting solution was cooled on ice, 2.0 mL TFA was added dropwise, and the solution was stirred at room temperature for about 24 h. TLC was used to monitor reaction progress (8 DCM:2 MeOH, Rf = 0.1). The deep orange solution was dried *in vacuo* to obtain a red oil that was purified using preparative HPLC (gradient H_2O (0.1% TFA):MeCN (0.1% TFA) from 100:0 to 0:100 to yield 75 mg (21% over two steps, only half of the deprotected product was purified by preparative HPLC) of the title compound as an orange solid. 99% pure by analytical HPLC (retention time 3.849 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.) ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.61 (s, 1H), 8.68 (s, 1H), 8.58 (s, 1H), 8.00 (bt, J = 5.85, 1H), 7.88 (bs, 2H), 7.55 (bs, 2H), 3.41 (q, J = 6.28, 2H), 2.82 (bp, J = 6.22, 2H), 2.55 (s, 3H), 1.60 (bm, 4H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 165.76, 159.73*, 159.45*, 159.18*, 158.91*, 156.27, 155.57, 152.70, 120.60, 120.21*, 117.87*, 115.53*, 113.19*, 108.68, 40.83, 39.02, 25.68, 24.86. *TFA. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 301.13; found 301.1.

Synthesis of **N23**



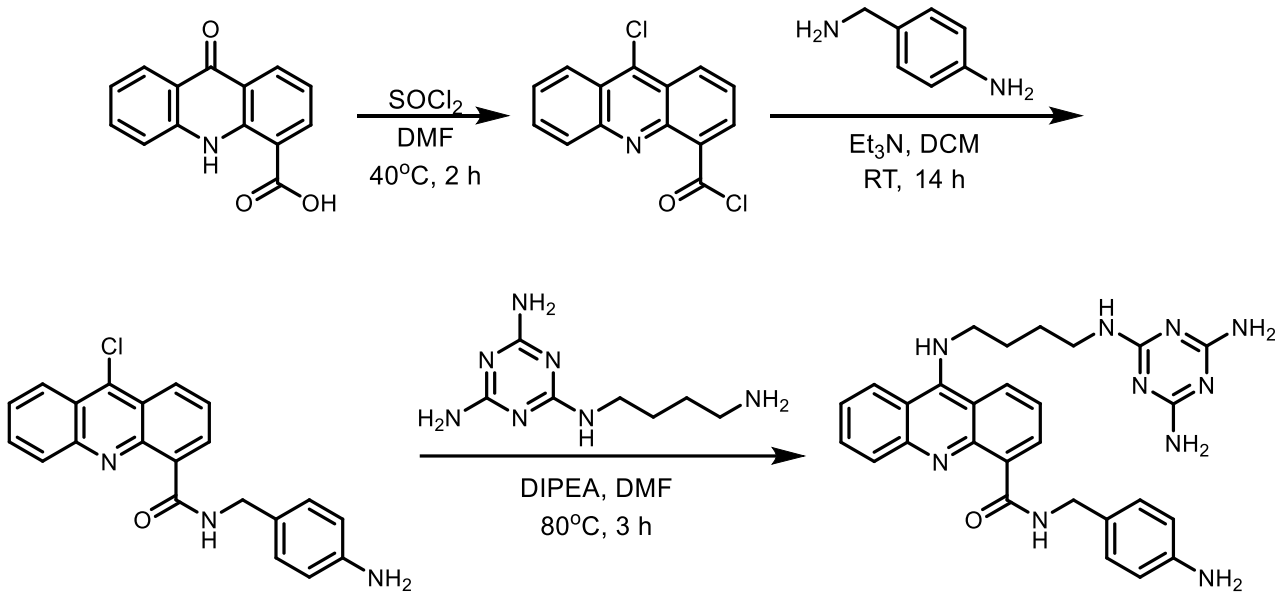
2,2'-azanediyldibenzoic acid.[4,23] To a 250 mL round bottom flask was added 7.69 g (56.1 mmol) anthranilic acid, 8.01 g (51.2 mmol) 2-chlorobenzoic acid, 21.29 g (154 mmol) potassium carbonate, and 68 mL DMF. To the resulting orange suspension was added 0.63 g (9.9 mmol) copper powder. The resulting dark orange suspension was heated to 130 °C and stirred overnight

(~18 h) under nitrogen. After cooling to room temperature, 150 mL ice water was added to the flask and it was sonicated to break up the clumps of precipitate. To the resulting brown suspension was added 4 scoopulas of activated carbon the black suspension was filtered through celite. The celite was washed with 1 L DI water and the filtrate was acidified with ~41 mL conc. HCl added dropwise while stirring. The suspension was filtered to obtain a tan solid and a dark brown filtrate. The tan solid was washed with 3x400 mL hot ethanol and dried *in vacuo* to yield 13.1 g (92%) of the title compound as a tan solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 13.08 (s, 2H), 10.88 (s, 1H), 7.97 (dd, *J* = 7.92, 1.56, 2H), 7.52 (m, 4H), 7.02 (ddd, *J* = 8.10, 6.51, 1.67, 2H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 168.83, 144.03, 133.81, 132.24, 120.43, 118.06, 118.02. LR-ESI MS (*m/z*) calcd for $C_{14}H_{12}NO_4^+$ [M+H] 258.25; found 258.1.



9-oxo-9,10-dihydroacridine-4-carboxylic acid.[23] To a 200 mL round bottom flask was added 2.24 g (8.71 mmol) 2,2'-azanediyldibenzoic acid and 45.0 mL conc. sulfuric acid was added dropwise. The brown solution was stirred at 80 °C for ~3 h under nitrogen. TLC was used to monitor reaction progress (8 DCM:2 MeOH, Rf 0.7). The fluorescent yellow solution was cooled to room temperature and dumped into200 mL ice. The beaker containing ice and product was cooled in an ice bath for 5 min and filtered with a Buchner funnel to obtain a yellow solid that was washed with 200 mL hot water, 200 mL ethanol, and 200 mL ether. The product was dried *in vacuo* to yield 2.05 g (99%) of the title compound as a yellow solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 11.96 (s, 1H), 8.54 (dd, *J* = 7.93, 1.68, 1H), 8.45 (dd, *J* = 7.52, 1.67, 1H), 8.24 (d, *J* = 8.18,

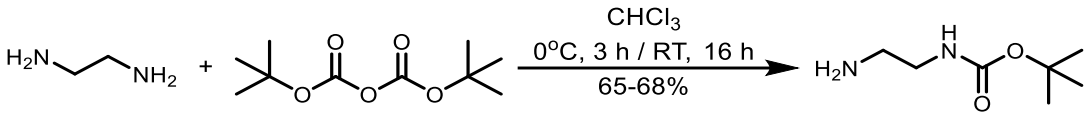
1H), 7.79 (dd, *J* = 5.88, 1.63, 2H), 7.35 (m, 2H). Note: acridone NH seems to be exchanged. ^{13}C NMR (500 MHz, DMSO-d_6) δ 177.00, 169.59, 141.66, 140.39, 137.37, 134.58, 132.88, 126.34, 122.79, 122.09, 121.06, 120.71, 119.10, 115.46. LR-ESI MS (*m/z*) calcd for $C_{14}H_{10}NO_3^+$ [M+H] 240.07; found 240.1.



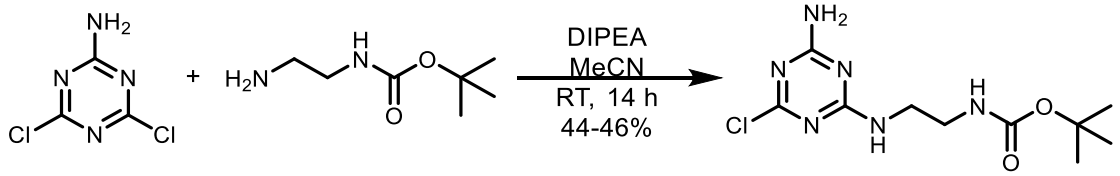
Compound N23. All reactions were run under dry nitrogen and with oven-dried glassware. To an oven-dried 65 mL round-bottom flask was added 1.18 g (4.9 mmol) 9-oxo-9,10-dihydroacridine-4-carboxylic acid (previously dried *in vacuo*) and 6.0 mL freshly distilled thionyl chloride. To the resulting yellow suspension was added 2 drops anhydrous DMF, and after about 10 s the solution became transparent orange in color. The reaction was heated at 70 °C for 25 min to obtain a brown solution and then an additional h. TLC was used to monitor reaction progress (9 DCM:1 MeOH, Rf 0.9). The resulting brown solution was cooled to room temperature and the thionyl chloride was distilled off. To the resulting yellow suspension was added 10 mL DCM that was distilled off, and this was repeated twice more (three times total) before the resulting yellow solid was dried *in vacuo* and used directly in the next reaction. The yellow solid was dissolved in 10 mL dry DCM and cooled to 0 °C. After stirring for 5 min, anhydrous triethylamine was added dropwise until the

pH was about 11 (about 280 drops). To the brown solution was added 0.63 mL (5.5 mmol) 4-(aminoethyl)aniline to obtain an orange suspension. Reaction progress was monitored by TLC (9 DCM:1 MeOH, Rf 0.6) and after 22 h the resulting brown suspension was dried *in vacuo* to yield an orangish brown solid. To the dry brown solid was added 1.00 g (5.1 mmol) N^2-(4-aminobutyl)-1,3,5-triazine-2,4,6-triamine and 91 mL DMF. To the resulting suspension was added 1.90 mL (11.0 mmol) *N,N*-diisopropylethylamine was added dropwise. The orange suspension was stirred at 70 °C for about 3 h. TLC was used to monitor reaction progress (8 DCM:2 MeOH, Rf 0.3). The resulting brown solution was dried *in vacuo* to obtain a brown oil that was purified using a Combi-Flash RediSep column (silica, 12 g) with elution gradient 100 DCM:0 MeOH to 80 DCM:20 MeOH. The fractions containing product were combined and concentrated *in vacuo*. The product was further purified further purified using preparative HPLC (gradient H2O (0.1% TFA):MeCN (0.1% TFA) from 100:0 to 0:100 to yield 1.81 mg (<1 % over three steps) of the title compound as an orange solid. 96% pure by analytical HPLC (retention time 5.043 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over 5 min.). ^{1}H NMR (500 MHz, DMSO-d_6) δ 15.02 (s, 1H)*, 8.74 (d, *J* = 7.21, 1H), 8.63 (d, *J* = 8.89, 1H), 8.47 (d, *J* = 8.88, 1H), 8.08 (d, *J* = 8.61, 1H), 7.86 (t, *J* = 8.13, 2H), 7.75 (bs, 1H), 7.48 (m, 2H), 7.38 (d, *J* = 7.99, 1H), 6.55 (m, 1H), 6.44 (d, *J* = 5.95, 1H), 6.29 (s, 1H), 6.01 (bs, 2H), 5.88 (bs, 2H), 4.54 (t, *J* = 7.14, 1H), 3.39 (d, *J* = 7.33, 2H), 3.18 (d, *J* = 6.66, 2H), 2.65 (bs, 1H), 1.82 (d, *J* = 8.00, 1H), 1.54 (p, *J* = 7.36, 1H), 1.25 (s, 4H), 0.86 (t, *J* = 6.69, 1H). *protonated acridine. ^{13}C NMR (500 MHz, DMSO-d_6) δ 204.74, 168.27, 168.00, 167.01, 161.31, 160.29, 155.42, 154.46, 154.18, 153.55, 152.73, 145.99, 140.96, 135.70, 129.13, 127.78, 126.47, 123.75, 121.18, 105.06, 55.16, 47.89, 43.95, 31.64, 30.10. HR-ESI MS (*m/z*) calcd for $C_{28}H_{31}N_{10}O^+$ [M+H$^+$] 523.2677; found 523.2670.

Synthesis of **N24**

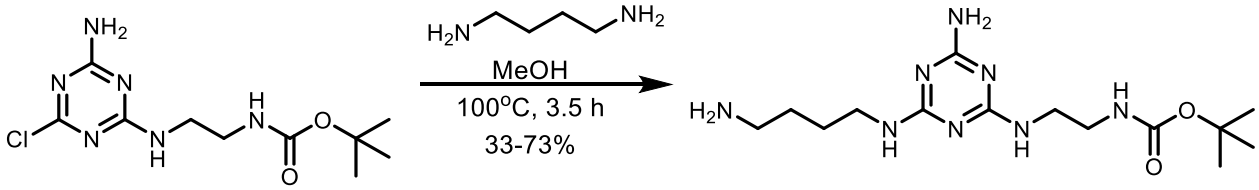


tert-butyl (2-aminoethyl)carbamate:[14] To a 1000 mL round-bottom flask was added 22.5 mL (337 mmol) 1,2-diaminoethane and 386 mL chloroform. The resulting clear solution was cooled to 0 °C in an ice bath. In a 500 mL Erlenmeyer flask was mixed 18.38 g (84.2 mmol) di-tertbutyl dicarbonate and 400 mL chloroform to form a clear solution that was added dropwise to the round-bottom flask while stirring using an addition funnel. The resulting milky white mixture was kept at 0 °C for 3 h, removed from the ice bath to slowly warm to room temperature and stirred for an additional ~16 h. Thin layer chromatography on silica gel plates (9 MeOH:1 NH_4OH, rf 0.8, ninhydrin stain) was used to monitor reaction progress. The white solution was decanted into a separatory funnel, washed with DI water (6 x 300 mL) and the organic layer was checked for remaining 1,2-diaminoethane by TLC (9 MeOH:1 NH_4OH, 1,2-diaminoethane Rf 0.1, pdt Rf 0.8). The organic layer was dried over sodium sulfate and concentrated *in vacuo* to obtain 11.30 g (84%) of the title compound as a colorless oil.^{1}H NMR (500 MHz, d_6-DMSO) δ 8.32 (s, 1H), 2.92 (bt, 2H), 2.52 (bs, 2H), 1.38 (s, 9H), 1.30 (bs, 2H).



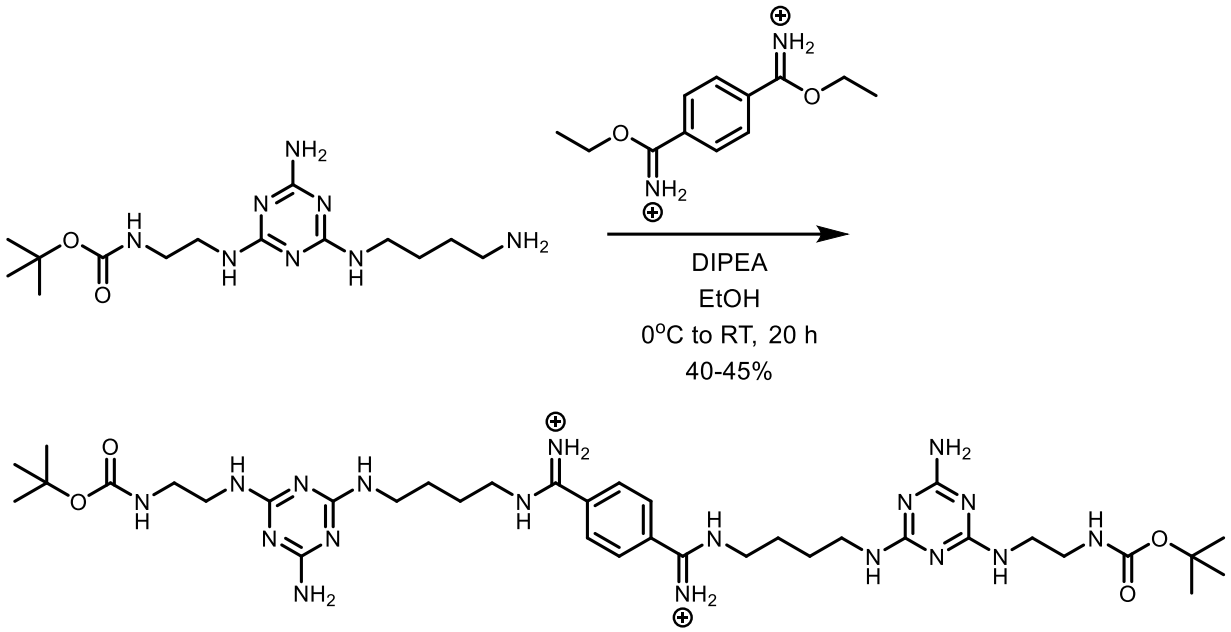
***tert*-butyl (2-((4-Amino-6-chloro-1,3,5-triazin-2-yl)amino)ethyl)carbamate.** To a 200 mL round-bottom flask was added 2.41 g (15 mmol) 4,6-dichloro-1,3,5-triazin-2-amine and 79 mL acetonitrile. To the resulting white suspension was added 7.0 mL (40 mmol) N, N-

diisopropylethylamine and the mixture was stirred. A mixture of 2.50 g (16 mmol) tert-butyl (2-aminoethyl)carbamate and 10 mL acetonitrile was added dropwise via syringe to obtain a faintly yellow suspension. The mixture was stirred for about 20 h, filtered with a Buchner funnel, washed with 3x100 mL water and 100 mL diethyl ether, and dried on a lyophilizer to obtain a white powder that was stored at 0 °C until further use. TLC (9 DCM:1 MeOH, Rf 0.7) was used to monitor reaction progress and the reaction yielded 1.88 g (46%) of the title compound as a white solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.66 (t, J = 5.76, 1H), 7.31 (bd, 1H), 7.16 (bd, 1H), 6.85 (dt, J = 11.42, 5.40, 1H), 3.24 (dq, J = 12.49, 6.28, 2H), 3.05 (dq, J = 12.92, 6.25, 2H), 1.37 (s, 9H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 168.53, 167.40, 166.15, 156.10, 78.16, 40.59, 40.42, 28.70. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 289.12; found 289.1.



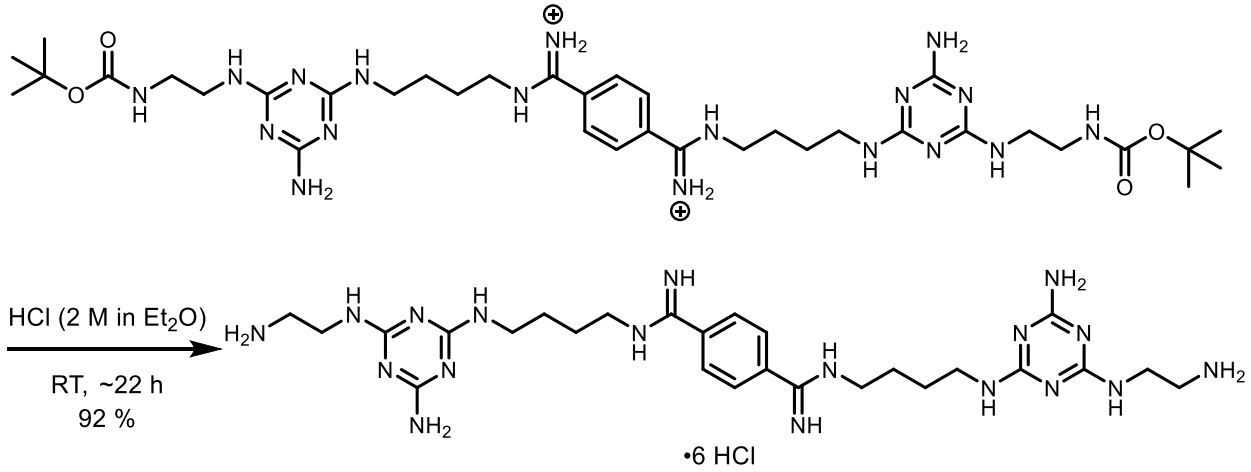
***tert*-butyl (2-((4-amino-6-((4-aminobutyl)amino)-1,3,5-triazin-2-yl)amino)ethyl)carbamate.** To a 25 mL round-bottom flask was added 1.4 mL (14 mmol) 1,4-diaminobutane and 5.0 mL methanol. The mixture was stirred at 100 °C to obtain clear solution and a mixture of 974.6 mg (3.4 mmol) *tert*-butyl (2-((4-Amino-6-chloro-1,3,5-triazin-2-yl)amino)ethyl)carbamate in 5.0 mL methanol was added slowly over 10 min. The reaction was stirred for an additional 3.5 h and dried *in vacuo* to obtain colorless oil. TLC was used to monitor progress with ninhydrin stain (8 MeOH:2 NH_4OH, Rf triazine 0.95, Rf pdt 0.7, Rf 1,4-DAB 0.1). The crude product was purified on Combi-Flash RediSep column (silica, 12 g) with elution gradient 95 DCM:5 MeOH to 70 DCM:30 MeOH. The fractions containing product were combined and dried *in vacuo* to obtain 873 mg (73%) of the

title compound as a fine white powder. ^{1}H NMR (500 MHz, DMSO-d_6) δ 6.81 (bs, 1H), 6.60-6.33 (bm, 1H), 6.31-5.80 (bm,3H), 3.25-3.11 (bm, 6H), 3.04 (bs, 2H), 1.45 (bm, 4H), 1.38 (s, 9H). ^{13}C NMR (500 MHz, DMSO-d_6) δ 166.49, 166.48, 166.39, 156.13, 78.05, 41.96, 31.20, 28.71, 27.42. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 341.24; found 341.1.



1,4-phenylenebis(((4-((4-amino-6-((2-((tert-butoxycarbonyl)amino)ethyl)amino)-1,3,5-triazin-2-yl)amino)butyl)amino)methaniminium). To an oven-dried 100 mL round-bottom flask was added 0.38 g (1.3 mmol) diethyl terephthalimidate dihydrochloride and 24 mL ethanol stored on 4Å molecular sieves. The resulting white suspension was cooled to 0 °C and kept under nitrogen atmosphere. To the white suspension was added 0.74 mL (3.8 mmol) *N*,*N*-diisopropylethylamine to obtain a colorless solution. A solution of 872.7 mg (2.56 mmol) *tert*-butyl (2-((4-amino-6-((4-aminobutyl)amino)-1,3,5-triazin-2-yl)amino)ethyl)carbamate in 10 mL ethanol stored on 4Å molecular sieves was added to the round-bottom flask dropwise with a syringe over 5 min. The transparent tan solution was warmed slowly to 25 °C and stirred for about 18 h. TLC was used to monitor progress with ninhydrin stain (9 MeOH:1 NH_4OH, Rf 0.1). The

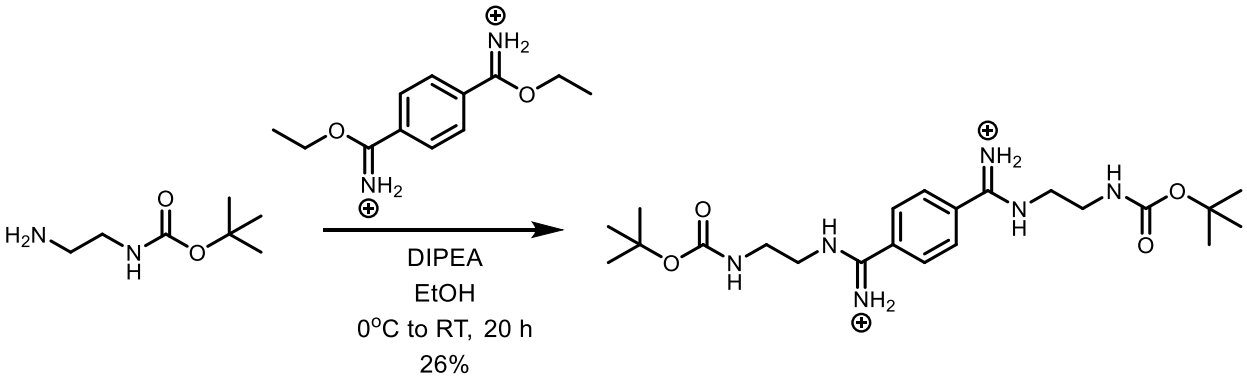
solvents were removed *in vacuo* to obtain a white solid. The crude product was purified on a 2" x 6" silica column with elution gradient 9 DCM:1 MeOH to 7 DCM:3 MeOH + 5% NH_4OH. The fractions containing product were combined and dried *in vacuo* to obtain 475 mg (45%) of the title compound as a white solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.08 (s, 2H), 9.71 (s, 2H), 9.28 (s, 2H), 7.95 (s, 4H), 7.22 (bm, 4H), 6.83 (bm, 2H), 6.71-6.41 (bm, 2H), 6.18-5.95 (bm, 2H), 3.47-3.43 (bt, *J* = , 6.34, 4H)*, 3.24 (bm, 8H), 3.05 (bs, 4H), 1.69 (bm, 4H), 1.59 (bm, 4H), 1.38 (s, 18 H). *overlapped with water peak. ^{13}C NMR (500 MHz, DMSO-d_6) δ 172.25, 166.51, 164.57, 162.29, 155.69, 133.48, 129.11, 77.71, 63.42, 42.80, 41.91, 39.50, 28.72, 25.27, 25.23. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 809.51; found 809.3.



Compound N24. In a 50 mL round bottom flask was mixed 475 mg (0.587 mmol) 1,4-phenylenebis(((4-((4-amino-6-((2-((tert-butoxycarbonyl)amino)ethyl)amino)-1,3,5-triazin-2-yl)amino)butyl)amino)methaniminium) and 20 mL of 2 M HCl in diethyl ether. The suspension was stirred at room temperature for about 22 h. TLC (8 MeOH:2 NH_4OH, Rf pdt = 0.1) was used to monitor reaction progress. The solvents were removed *in vacuo* to yield 446 mg (92%) of the title compound as a white solid HCl salt. 99% pure by analytical HPLC (retention time 3.765 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-

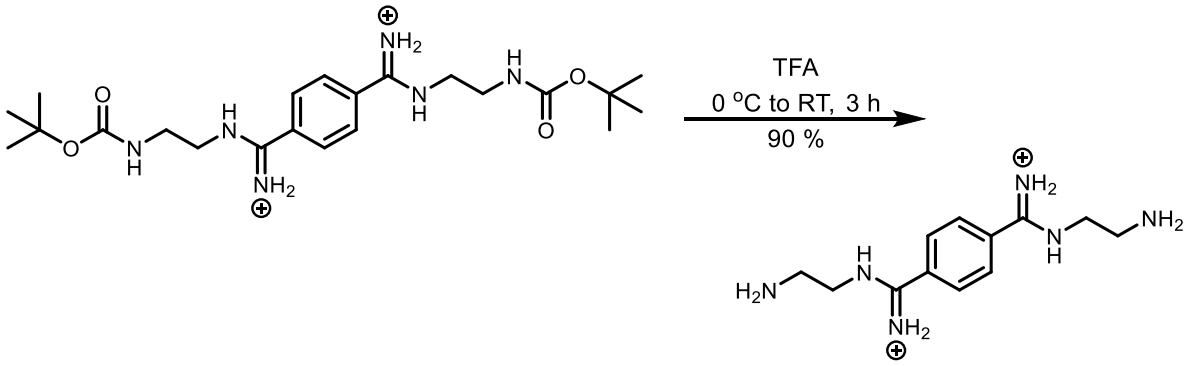
100% over 5 min.) ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.31-10.23 (bm, 2H), 9.80 (s, 2H), 9.52-9.38 (bm, 2H), 8.63-8.50 (bm, 2H), 8.30 (s, 2H), 8.20 (s, 2H), 8.00 (s, 4H), 7.95 (s, 2H), 3.61 (q, *J* = 6.12, 2H), 3.57-3.47 (bm, 6H), 3.43 (q, *J* = 6.11, 2H), 3.40-3.33 (bm, 2H), 2.99 (dq, *J* = 17.87, 6.53, 4H), 1.73 (p, *J* = 8.00, 4H), 1.65 (p, *J* = 7.21, 4H), 1.24 (s, 1H)*. *partially exchanged amine. ^{13}C NMR (500 MHz, DMSO-d_6) δ 162.98, 162.08, 161.74, 156.34, 133.34, 129.15, 49.06, 42.99, 38.65, 38.37, 26.49, 25.02. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 609.41; found 609.2. MALDI-MS (*m/z*) calc for [M+H$^+$] 609.4069; found 609.406.

Synthesis of **N25**



1,4-phenylenebis(((2-((tert-butoxycarbonyl)amino)ethyl)amino)methaniminium). To an oven-dried 100 mL round-bottom flask was added 1.04 g (3.5 mmol) diethyl terephthalimidate dihydrochloride and 70 mL ethanol stored on 4Å molecular sieves. The resulting white suspension was cooled to 0 °C and kept under nitrogen atmosphere. To the white suspension was added 1.80 mL (10 mmol) *N*,*N*-diisopropylethylamine to obtain a colorless solution. A solution of 1.19 g (7.4 mmol) tert-butyl (2-aminoethyl)carbamate in 10 mL ethanol stored on 4Å molecular sieves was added to the round-bottom flask dropwise with a syringe over ~5 min. The colorless solution was warmed slowly to 25 °C and stirred for about 18 h. TLC was used to monitor progress with ninhydrin stain (8 MeOH:2 NH_4OH, Rf 0.2). The solvents were removed *in vacuo* to obtain a white

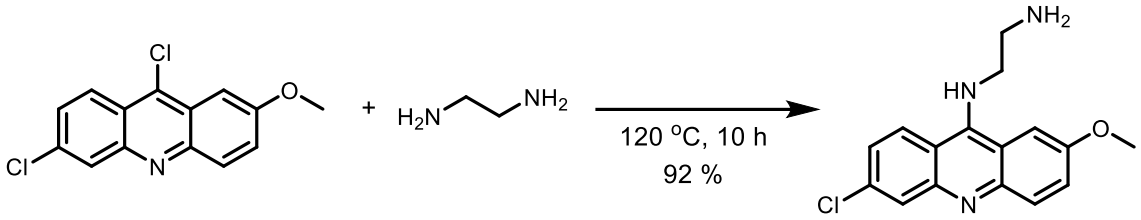
solid. The crude product was purified on a CombiFlash column (silica, 24 g) with gradient 95 DCM:5 MeOH to 70 DCM:30 MeOH. The fractions containing product were combined and dried *in vacuo* to obtain 474 mg (26%) of the title compound as a white solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.97 (s, 4H), 7.12 (t, J = 5.99, 2H), 4.13 (q, J = 5.26, 2H), 3.53 (t, J = 5.69, 4H), 3.30 (d, J = 5.97, 4H), 3.17 (d, J = 4.87, 4H), 1.39 (s, 18H). *Amine protons seem to have exchanged. ^{13}C NMR (500 MHz, DMSO-d_6) δ 162.59, 156.35, 133.49, 129.07, 78.54, 43.77, 38.38, 28.71. HR-ESI MS (*m/z*) calcd for [M+H$^+$] 449.29; found 449.3.



Compound N25. In a 20 mL vial on ice was dissolved 473.94 mg (0.907 mmol) 1,4-phenylenebis(((2-((tert-butoxycarbonyl)amino)ethyl)amino)methaniminium) in 5 mL trifluoroacetic acid (added dropwise). After addition, the solution was warmed to room temperature and stirred for about 3 h. TLC was used to monitor progress with ninhydrin stain (9 MeOH:1 NH_4OH, Rf 0.1). The solvents were removed *in vacuo* to obtain a tan to pink oil. The crude product was dissolved in 100 mL acetone to obtain a cloudy solution that was let sit overnight at room temperature. The white precipitate was filtered off using a Buchner funnel and washed with 100 mL acetone. The tan filtrate was rotovapped to yield 572 mg (90%) of the title compound as a white solid. 99% pure by analytical HPLC (retention time 0.684 min, analytical HPLC gradient acetonitrile in water, 0.1% TFA 0-50% over 5 min, 50% for 5 min, 50-100% over

5 min.) ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.17 (bs, 2H), 9.97 (s, 2H), 9.48 (s, 2H), 8.04 (s, 4H), 3.72 (q, *J* = 5.80, 4H), 3.20 (q, *J* = 5.80, 4H), 2.48 (bs, 3H)*. *Terminal amine protons seem partially exchanged. ^{13}C NMR (500 MHz, DMSO-d_6) δ 163.25, 159.35*, 159.09*, 158.82*, 158.56*, 133.56, 129.21, 120.62*, 118.26*, 115.91*, 113.55*, 49.06, 41.06. *TFA peaks. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 249.18; found 249.2.

Synthesis of **N26**



Compound N26. To a 300 mL round bottomed flask was added 6.03 g (22 mmol) 6,9-dicholoro-2-methoxyacridine and 117 mL (1.75 mol) 1,2-diaminoethane. The yellow suspension was stirred in an oil bath at 120 °C under nitrogen. The yellow suspension transitioned to an orange solution and then a brown solution that was cooled to room temperature after about 10 h. To the flask was added 280 mL DI water to obtain a yellow precipitate in an orange solution. Filtered with a Buchner funnel and washed with 20x6 mL DI water. The solid was suspended in 400 mL diethyl ether, sonicated, filtered again, and washed with ~1200 mL diethyl ether. The solid was dried *in vacuo* to yield 6.00 g (92%) of the title compound as an orange solid. ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.37 (d, *J* = 9.35, 1H), 7.91-7.76 (bm, 2H), 7.64 (d, *J* = 2.71, 1H), 7.42 (dd, *J* = 9.30, 2.68, 1H), 7.33 (d, *J* = 9.40, 1H), 3.94 (s, 3H), 3.72 (t, *J* = 6.36, 2H), 2.88 (t, *J* = 6.35, 2H). Amine protons seem to be exchanged. ^{13}C NMR (500 MHz, DMSO-d_6) δ 155.50, 151.13, 133.89, 131.10, 127.56, 126.97, 124.70, 123.12, 122.52, 117.68, 115.25, 101.37, 101.17, 56.07, 53.11, 42.82. LR-ESI MS (*m/z*) calcd for [M+H$^+$] 302.78; found 302.0.

A.5. REFERNCES

(1) Vichai, V.; Kirtikara, K. Sulforhodamine B Colorimetric Assay for Cytotoxicity Screening. *Nat. Protoc.* **2006**, *1* (3), 1112–1116.

(2) Nguyen, L.; Luu, L. M.; Peng, S.; Serrano, J. F.; Chan, H. Y. E.; Zimmerman, S. C. Rationally Designed Small Molecules That Target Both the DNA and RNA Causing Myotonic Dystrophy Type 1. *J. Am. Chem. Soc.* **2015**, *137* (44), 14180–14189.

(3) Hagler, L. D.; Krueger, S. B.; Luu, L. M.; Lanzendorf, A. N.; Mitchell, N. L.; Vergara, J. I.; Curet, L. D.; Zimmerman, S. C. Versatile Target-Guided Screen for Discovering Bidirectional Transcription Inhibitors of a Trinucleotide Repeat Disease. *ACS Med. Chem. Lett.* **2021**, *12* (6), 935–940.

(4) Jahromi, A. H.; Nguyen, L.; Fu, Y.; Miller, K. A.; Baranger, A. M.; Zimmerman, S. C. A Novel CUGexp·MBNL1 Inhibitor with Therapeutic Potential for Myotonic Dystrophy Type 1. *ACS Chem. Biol.* **2013**, *8* (5), 1037–1043.

(5) Shaikh, M. S.; Rana, J.; Gaikwad, D.; Leartsakulpanich, U.; Ambre, P. K.; Pissurlenkar, R. R. S.; Coutinho, E. C. Antifolate Agents Against Wild and Mutant Strains of Plasmodium Falciparum. *Indian J. Pharm. Sci.* **2014**, *76* (2), 116–124.

(6) Baliani, A.; Jimenez Bueno, G.; Stewart, M. L.; Yardley, V.; Brun, R.; Barrett, M. P.; Gilbert, I. H. Design and Synthesis of a Series of Melamine-Based Nitroheterocycles with Activity against Trypanosomatid Parasites. *J. Med. Chem.* **2005**, *48*, 5570–5579.

(7) Marcheschi, R. J.; Tonelli, M.; Kumar, A.; Butcher, S. E. Structure of the HIV-1 Frameshift Site RNA Bound to a Small Molecule Inhibitor of Viral Replication. *ACS*

Chem. Biol. **2011**, *6* (8), 857–864.

(8) NAGAO, Y.; TAKASU, A. ‘“Click Polyester”’: Synthesis of Polyesters Containing Triazole Units in the Main Chain via Safe and Rapid ‘“Click”’ Chemistry and Their Properties. *J. Polym. Sci. Part A Polym. Chem.* **2010**, *48*, 4207–4218.

(9) Srinivasan, R.; Tan, L. P.; Wu, H.; Yang, P. Y.; Kalesh, K. A.; Yao, S. Q. High-Throughput Synthesis of Azide Libraries Suitable for Direct “Click” Chemistry and in Situ Screening. *Org. Biomol. Chem.* **2009**, *7* (9), 1821–1828.

(10) Staudinger, H.; Meyer, J. Synthesis of Phosphazo Compounds by the Reaction of Tertiary Phosphines with Organic Azides. *Helv. Chim. Acta* **1919**, *2*, 635–646.

(11) Prasad, B.; Jamroskovic, J.; Bhowmik, S.; Kumar, R.; Romell, T.; Sabouri, N.; Chorell, E. Flexible Versus Rigid G-Quadruplex DNA Ligands: Synthesis of Two Series of Bis-Indole Derivatives and Comparison of Their Interactions with G-Quadruplex DNA. *Chem. - A Eur. J.* **2018**, *24* (31), 7926–7938.

(12) Chaturvedi, D.; Chaturvedi, A. K.; Mishra, N.; Mishra, V. An Efficient and Novel Approach for the Synthesis of Substituted N-Aryl Lactams. *Org. Biomol. Chem.* **2012**, *10* (46), 9148–9151.

(13) Helin, J.; Saarinen, J.; Ekholm, F. S.; Leppanen, A. Payload-Polymer-Protein Conjugates, 2017.

(14) Muller, D.; Zeltser, I.; Bitan, G.; Gilon, C. Building Units for N-Backbone Cyclic Peptides. 3. Synthesis of Protected N α -(ω-Aminoalkyl)Amino Acids and N α -(ω-Carboxyalkyl)Amino Acids. *J. Org. Chem.* **1997**, *62* (2), 411–416.

(15) Wong, C. H.; Nguyen, L.; Peh, J.; Luu, L. M.; Sanchez, J. S.; Richardson, S. L.; Tuccinardi, T.; Tsoi, H.; Chan, W. Y.; Chan, H. Y. E.; Baranger, A. M.; Hergenrother, P. J.; Zimmerman, S. C. Targeting Toxic RNAs That Cause Myotonic Dystrophy Type 1 (DM1) with a Bisamidinium Inhibitor. *J. Am. Chem. Soc.* **2014**, *136* (17), 6355–6361.

(16) Bouillon, C.; Paolantoni, D.; Rote, J. C.; Bessin, Y.; Peterson, L. W.; Dumy, P.; Ulrich, S. Degradable Hybrid Materials Based on Cationic Acylhydrazone Dynamic Covalent Polymers Promote Dna Complexation through Multivalent Interactions. *Chem. - A Eur. J.* **2014**, *20* (45), 14705–14714.

(17) Hagler, L. D.; Luu, L. M.; Tonelli, M.; Lee, J.; Hayes, S. M.; Bonson, S. E.; Vergara, J. I.; Butcher, S. E.; Zimmerman, S. C. Expanded DNA and RNA Trinucleotide Repeats in Myotonic Dystrophy Type 1 Select Their Own Multitarget, Sequence-Selective Inhibitors. *Biochemistry* **2020**, *59* (37), 3463–3472.

(18) Ferlin, M. G.; Marzano, C.; Chiarelotto, G.; Baccichetti, F.; Bordin, F. Synthesis and Antiproliferative Activity of Some Variously Substituted Acridine and Azacridine Derivatives. *Eur. J. Med. Chem.* **2000**, *35* (9), 827–837.

(19) Mao, J.; Desantis, C.; Bong, D. Small Molecule Recognition Triggers Secondary and Tertiary Interactions in DNA Folding and Hammerhead Ribozyme Catalysis. *J. Am. Chem. Soc.* **2017**, *139* (29), 9815–9818.

(20) Arambula, J. F.; Ramisetty, S. R.; Baranger, A. M.; Zimmerman, S. C. A Simple Ligand That Selectively Targets CUG Trinucleotide Repeats and Inhibits MBNL Protein Binding. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (38), 16068–16073.

(21) Galindo, M. A.; Amantia, D.; Martinez, A. M.; Clegg, W.; Harrington, R. W.; Martinez,

V. M.; Houlton, A. Probing Metal-Ion Purine Interactions at DNA Minor-Groove Sites. *Inorg. Chem.* **2009**, *48* (21), 10295–10303.

(22) Scheffer, U.; Strick, A.; Ludwig, V.; Peter, S.; Kalden, E.; Göbel, M. W. Metal-Free Catalysts for the Hydrolysis of RNA Derived from Guanidines, 2-Aminopyridines, and 2-Aminobenzimidazoles. *J. Am. Chem. Soc.* **2005**, *127* (7), 2211–2217.

(23) Cuenca, F.; Moore, M. J. B.; Johnson, K.; Guyen, B.; De Cian, A.; Neidle, S. Design, Synthesis and Evaluation of 4,5-Di-Substituted Acridone Ligands with High G-Quadruplex Affinity and Selectivity, Together with Low Toxicity to Normal Cells. *Bioorganic Med. Chem. Lett.* **2009**, *19* (17), 5109–5113.

APPENDIX B: PUBLICATIONS

CONTENTS (as supplemental files)

Serrano, J.F., Lee, J.Y., Curet, L.D., Hagler, L.D., **Bonson, S.E.**, Schuster, E.J., Zimmerman, S.C. Development of novel macrocyclic small molecules that target CTG trinucleotide repeats. *Bioorganic & Medicinal Chemistry* **2019**, *27*, 2978-2984. https://doi.org/10.1016/j.bmc.2019.05.022

Hagler, L.D., **Bonson S.E.**, Kocheril, P.K., Zimmerman S.C. Assessing the Feasibility and Stability of U-base Flipping in RNA-Small Molecule Complexes Using Molecular Dynamics Simulations. *Canadian Journal of Chemistry* **2020**, *98*, 261-269. https://doi.org/10.1139/cjc-2019-0421

Chen, J., Li, K., **Bonson S.E.**, Zimmerman, S.C. A Bioorthogonal Small Molecule Selective Polymeric "Clickase". *J. Am. Chem. Soc.* **2020**, *142*, 13966-13973. https://doi.org/10.1021/jacs.0c06553

Hagler, L.D., Luu, L.M., Tonelli, M., Lee, J., Hayes, S., **Bonson, S.E.**, Vergara, J.I., Butcher, S.E., Zimmerman, S.C. Expanded DNA and RNA Trinucleotide Repeats in Myotonic Dystrophy Type 1 Select Their Own Multitarget, Sequence-Selective Inhibitors. *Biochemistry* **2020**, *59*, 3463-3472. https://doi.org/10.1021/acs.biochem.0c00472

Hagler, L.D., **Krueger, S.B.**, Luu, L.M., Lanzendorf, A.N., Mitchell, N.L., Vergara, J.L., Curet, L.D., Zimmerman, S.C. Versatile Target-Guided Screen for Discovering Bidirectional Transcription Inhibitors of a Trinucleotide Repeat Disease. *ACS Med. Chem. Lett.* **2021**, *12*, 935-940. https://doi.org/10.1021/acsmedchemlett.1c00064

Li, K., **Krueger, S.B.**, Zimmerman, S.C. A Novel Minor Groove Binder as a Potential Therapeutic Agent for Myotonic Dystrophy Type 1. *ChemMedChem* **2021**, *16*, 2638-2644. – selected as a VIP. https://doi.org/10.1002/cmdc.202100243

MANUSCRIPTS IN PREPARATION

Li, K., Chembazhi, U.V., **Krueger, S.B.**, Chen, J., Dewald, Z., Bai, Y., Kim, D., Kocheril, P., Chen, J., Kalsotra, A., Zimmerman S.C. Versatile Discrete Oligomer for Myotonic Dystrophy Type 1. Manuscript in revision for resubmission to *Nature Communications*

Krueger, S.B., Lanzendorf, A.N., Jeon, H., Zimmerman, S.C. Toward Template-Selected, Reversible Assembly of Therapeutic Agents for Myotonic Dystrophy Type 1. Manuscript in preparation for submission to *ChemBioChem*

Krueger, S.B., Zimmerman, S.C. Dynamic Covalent Target-Guided Screen for Nucleic Acid-Targeting Agents.