

“RNA KISSING LOOP LIBRARY TO EXPAND RNA FIBER FUNCTIONALITY
AND CONTROL CARGO LOAD
FOR POTENTIAL INTRACELLULAR siRNA DELIVERY”

by

Oleg Anatolyevich Shevchenko

A thesis submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Master of Science in
Chemistry

Charlotte

2020

Approved by:

Dr. Kirill Afonin

Dr. Swarnapali De Silva Indrasekara

Dr. Kausik Chakrabarti

Dr. Caryn D. Striplin

©2020
Oleg Anatolyevich Shevchenko
ALL RIGHTS RESERVED

ABSTRACT

OLEG ANATOLYEVICH SHEVCHENKO. "RNA kissing loop library to expand RNA fiber functionality and control cargo load for potential intracellular siRNA delivery".
(Under the direction of DR. KIRILL AFONIN)

A kissing loop (KL) is an RNA motif that among others may have a 180° geometry when derived from a viral genome. The KL interactions of this geometry consist of nine base pairs, six of which are interacting in complementary binding while three are stacking for support of the geometry. In RNA nanotechnology, the rational design of nanoparticles is dependent on the canonical complementarity of nucleobases, where adenines bind uracils and guanines bind cytosines. The sequences of various components play a critical role in their intercompatibility with one another (or intracompatibility within a single self-folding strand) of the intermediate or final nanostructures, as their crosstalk or undesired interactions may hinder the assembly of components or the final nanostructure. This KL motif is important for its incorporation into larger RNA nanostructures as it functions as a "linker" to attach other RNA functionalities together. This work investigates the expansion of a single KL sequence to a KL library that can be intercompatible for use in a single nanostructure. This KL library would be crucial to assembling RNA vehicles with higher heterogeneous carrying capacity for therapeutic siRNAs or any other cargo into target cells.

ACKNOWLEDGEMENTS

I would like to thank Dr. Kirill Afonin for giving me the opportunity to work in his research lab, for helping me develop ideas for the kissing loop library project, for exposing me to outside literature during Journal Clubs, and practicing data presentation to my lab peers. I would also like to thank lab members Justin Halman and Morgan Chandler for helping answer my numerous questions while learning lab techniques.

I would like to thank Dr. Jerry Troutman and his lab members, Amanda Reid and Beth Scarbrough, for providing and allowing to use the equipment in the Troutman Lab, and helping answer any questions I had. Amanda Reid was extremely helpful in guiding and troubleshooting during expression of T7 RNA polymerase. Tatiana Lyalina (a visiting scientist in Dr. Afonin's lab from St. Petersburg, Russia) helped understand T7 RNA polymerase gene expression and helped get the experiment started.

Lastly, I am thankful to my God, the God of all creation, that gave me life, for giving me wisdom and strength on this journey. I would also like to thank to my family, who encouraged me and provided emotional support during this time of my life.

TABLE OF CONTENTS

LIST OF TABLES	ix
LIST OF FIGURES	xi
LIST OF ABBREVIATIONS	xii
PREFACE	1
CHAPTER 1: INTRODUCTION TO NUCLEIC ACIDS AND KISSING LOOP PROJECT	4
1.1 Structural properties of nucleic acids	4
1.2 RNA in Nanotechnology	11
1.3 Kissing Loop Motif	12
1.4 RNA fibers	14
1.5 Problem Statement (and solution)	15
1.6 Goal of Project	17
1.7 Hypothesis	18
CHAPTER 2: METHODS OF T7 RNA POLYMERASE EXPRESSION FROM <i>E. COLI</i>	19
2.1 LB (Luria-Bertani) broth preparation	20
2.2 Preculture preparation	20
2.3 Bacterial cell culture growth	20
2.4 Induction of bacterial cells with IPTG	22
2.5 Bacterial cell collection (with possible stopping point)	23
2.6 SDS-PAGE preparation	25
2.7 Induction confirmation of T7 RNA polymerase gene expression via SDS-PAGE	27

2.8	Staining SDS-PAGE with Coomassie blue dye	27
2.9	Preparation of lysed bacterial cells	28
2.10	Purification of T7 RNA polymerase	30
2.11	Evaluation of protein purification stages via SDS-PAGE	33
2.12	SDS-PAGE analysis of T7 RNA polymerase expression	35
2.13	Protein Dialysis	37
2.14	Determining protein concentration and preparation for storage	37
2.15	Recipes for T7 RNA polymerase purification buffers	39
CHAPTER 3:	WORKING WITH NUCLEIC ACIDS	40
3.1	RNA strand design	40
3.2	Preparing DNA template and primers for PCR	42
3.3	DNA amplification via PCR	43
3.4	DNA purification after PCR	45
3.5	PCR-product (p)DNA verification via agarose gel	47
3.6	In vitro RNA synthesis using T7 RNA polymerase (“Transcription”)	50
3.6.1	Transcription Initiation	52
3.6.2	Transcription Termination	53
3.6.3	Transcription Purification via Urea-PAGE	53
3.7	RNA precipitation and Recovery (Transcription Purification)	57
3.8	Measuring nucleic acid concentrations on NanoDrop 2000 spectrophotometer	59
3.9	General Assembly Protocol	60
3.10	Assembled structure verification via native-PAGE	61

CHAPTER 4:	KISSING LOOP PROJECT DEVELOPMENT	64
4.1	Active KL region sequence selection	64
4.2	KL nomenclature	66
4.3	Hairpin design	66
4.4	Kissing Loop crossTalk Results Matrix design	68
4.5	KL strand crosstalk experiments	71
4.6	Methods for crossTalk experiments	72
4.7	Kissing loop T _m determination	74
CHAPTER 5:	RESULTS AND DISCUSSION	77
5.1	Results from in vitro transcription optimization	77
5.2	Results from T7 RNA Polymerase Optimization for in vitro transcription	79
5.3	Color code decryption of crossTalk Results Matrix	81
5.4	Kissing Loop crossTalk Results	82
5.5	CrossTalk Results filtering	87
5.6	T _m Results of Kissing Loop Complexes from TGGE analysis	88
CHAPTER 6:	APPLICATIONS OF KISSING LOOPS	93
6.1	New KLs crosstalk with nanoring KLs	93
6.2	RNA and RNA/DNA hybrid Fiber strands	95
6.3	Modifications based on hybrid strands	96
6.4	SuperFibers / Fiber Bundles	98
CHAPTER 7:	CONCLUSIONS	100
REFERENCES		101

APPENDIX: SUPPLEMENTARY DETAILS	111
A.1 Hairpin KL full sequences	111
A.2 Recipes for general laboratory buffer preparation	112
A.3 Alternative view of crossTalk results matrices	113
A.4 Sample gels from crossTalk #1 experiment	115
A.5 Sample gels from crossTalk #2 experiment	117
A.6 Sample gels from Tm #1 experiment	118

LIST OF TABLES

TABLE 2.1:	Reagent recipe for making a 10% SDS-PAG, SDS-PAGE running buffer, and destain solution.	26
TABLE 2.2:	Reagent recipes for making T7 RNA polymerase elution buffers.	39
TABLE 3.1:	PCR reaction recipe and Thermocycler program protocol.	44
TABLE 3.2:	Reagent recipe for making a 1.5 % Agarose gel.	48
TABLE 3.3:	Reagent recipe for setting up <i>in vitro</i> transcription reaction.	51
TABLE 3.4:	Recipes for preparing reagents for in vitro transcription reaction.	51
TABLE 3.5:	Reagent recipe for making an 8M urea-PAG (polyacrylamide gel).	54
TABLE 3.6:	Recipe for making Crush and Soak buffer.	56
TABLE 3.7:	Assembly Reagents and their concentrations.	61
TABLE 3.8:	Reagent recipe for making an 8 % native-PAG (polyacrylamide gel).	62
TABLE 4.1:	16 sets of mutated KL sequence regions selected for testing.	66
TABLE 4.2:	Initial KL crossTalk combination matrix.	70
TABLE 4.3:	Reduced KL crossTalk combination matrix.	71
TABLE 5.1:	<i>In vitro</i> transcription optimization recipes.	78
TABLE 5.2:	T7 RNA polymerase optimization recipe for <i>in vitro</i> transcription.	80
TABLE 5.3:	Results from crossTalk #1 experiment.	85
TABLE 5.4:	Results from crossTalk #2 experiment.	86
TABLE 5.5:	Kissing loops complex sets kept as a result of 4 different results filters.	88
TABLE 5.6:	Melting temperature of all tested KLCs.	92
TABLE 6.1:	CrossTalk between active fiber and nanoring KL fragment sequences.	94

TABLE 6.2:	Differences in helix rotation between various HIV strands.	96
TABLE A.1:	Full-length sequences of KL hairpins and displacement strands.	111
TABLE A.2:	Recipes for various laboratory buffers and solutions.	112
TABLE A.3.1:	Results from crossTalk #1 experiment; alternative view.	113
TABLE A.3.2:	Results from crossTalk #2 experiment; alternative view.	114

LIST OF FIGURES

FIGURE 1:	Lab logo developed for Dr. Afonin's laboratory.	3
FIGURE 1.1:	C3'- <i>endo</i> and C2'- <i>endo</i> sugar pucker conformations.	7
FIGURE 1.2:	Forms of nucleic acids, and their structures.	7
FIGURE 1.3:	RNA kissing loop 3D and 2D representations.	13
FIGURE 1.4:	A-B type simple RNA fiber 2D representation.	14
FIGURE 1.5:	A-B type functionalized RNA fiber 2D representation.	15
FIGURE 1.6:	ABCDE-type functionalized RNA fiber 2D representation.	17
FIGURE 2.1:	Sample SDS-PAG from T7 RNA polymerase expression.	35
FIGURE 3.1:	Sequence samples for RNA strand design.	41
FIGURE 4.1:	Secondary structure of a functionalized RNA fiber and KL hairpin.	67
FIGURE 4.2:	Template for hairpin strand sequence and its development.	68
FIGURE 4.3:	Temperature-gradient gel electrophoresis (TGGE) apparatus.	75
FIGURE 4.4:	TGGE thermal block and a sample native-PAG.	76
FIGURE 5.1:	<i>In vitro</i> optimization RNA product yields at different conditions.	79
FIGURE 5.2:	RNA product yields from T7 RNA polymerase optimization.	80
FIGURE 5.3:	Decryption of color codes representing degree of crosstalk.	82
FIGURE 5.4:	Melting temperature determination approach.	91
FIGURE 6.1:	RNA fiber depicted with computer graphics.	97
FIGURE 6.2:	Counterclockwise rotation of the kissing loop.	98

LIST OF ABBREVIATIONS

A	adenosine nucleoside
Ade	adenine nitrogenous base
AA or A:bA	acrylamide bisacrylamide
AB	Assembly buffer
AFM	atomic force microscopy
APS	ammonium persulfate
ATP	adenosine triphosphate
bp	base pair
C	cytidine nucleoside
Cyt	cytosine nitrogenous base
ddiH ₂ O	double deionized water, Ultra-Pure water, typically 17.5-17.8 MΩ·cm
DIS	dimerization initiation site
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
dsRNA	double-stranded RNA
DTT	dithiothreitol
E1, E2, ...	Elution 1, 2, ...
EDTA	ethylenediaminetetraacetic acid
EtBr	ethidium bromide
FT	flow through
FWD	forward

G	guanosine nucleoside
Gua	guanine nitrogenous base
HIV	human immunodeficiency virus
IDT	Integrated DNA Technologies
IPTG	isopropyl- β -D-thiogalactopyranoside
K ⁺	potassium ion, monovalent
KL	kissing loop
KLC	kissing loop complex
LB	loading buffer
LB broth	Luria-Bertani broth
Mg ²⁺ or Mg ⁺⁺	magnesium ion, divalent
NA	nucleic acid
NMR	nuclear magnetic resonance
nt	nucleotide
OD	optical density
PAG	polyacrylamide gel
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	protein data bank
pDNA	purified DNA
RB	running buffer
REV	reverse

RISC	RNA-induced silencing complex
RNA	ribonucleic acid
RNAi	RNA interference
rNTP	ribonucleotide triphosphates
SCL	Soluble Cell Lysate
SDS	sodium dodecyl sulfate
siRNA	short interfering
T	thymidine nucleoside
TB	transcription buffer (i.e. 5x TB)
TB	tris-borate
TBE	tris-borate-EDTA
TEMED	tetramethylethylenediamine
Thy	thymine
U	uridine nucleoside
Ura	uracil nitrogenous base
W	Wash
WCL	Whole Cell Lysate

PREFACE

Life is fascinating when looking at all the living creatures and wondering how everything is unique in its own way. Every living organism propagates its kind through transfer of the seed, which contains certain instructions, or a blueprint of life for that organism. People look at their children and see traits of their parents, of themselves, and sometimes elements that they do not recognize, but learn that they occurred in some of their relatives.

In mid 19th century (in 1866), an Austrian monk, Gregor Mendel,¹ experimented with pea plants and saw that there were certain units of inheritance that were transferred from parent to progeny. In 1869, a Swiss doctor Friedrich Miescher, isolated a novel substance from nuclei of leukocytes that resisted digestion to proteases, and based on analyses was not protein or lipid.² Miescher called this substance “nuclein”. In 1910, Thomas Morgan proposed the chromosomal theory of heredity,³ where he talked about chromosomes having unit characters responsible for development of different parts of the organism. Around the same time, in 1909, a Danish botanist Wilhelm Johannsen,⁴ called those Mendelian units of inheritance genes.

Later, in 1944, Oswald Avery, Colin MacLeod, and Maclyn McCarty reported work⁵ on virulent strain of *Streptococcus pneumoniae*, where after heat treatment the filtrate was treated with enzymes to remove proteins, lipids, and carbohydrates. What was left of the filtrate was analyzed and found that a white fibrous substance with chemical properties to DNA was in the filtrate. This filtrate was further treated with enzymes against protein and RNA. This filtrate was then used to transform the non-

virulent strain of bacteria into a virulent strain.⁵ They found that when treating the filtrate with DNA-digesting enzyme that the filtrate lost the capacity to transform bacteria from non-virulent into virulent strain, concluding that DNA was the carrier of genetic inheritance.

In 1950, Erwin Chargaff reported⁶ about nucleic acid composition of 4 nucleic bases (adenine, guanine, cytosine, and thymine) of different organismal origin, finding that in nucleic acids of different species purine and pyrimidine content was equal. It is based on this research that Chargaff's rules are based, where adenine interacts with thymine and cytosine interacts with guanine in a double-stranded DNA.⁶ In 1952, Raymond Gosling, a student who worked with Rosalind Franklin, took the famous "photo 51" of the DNA X-ray diffraction. In 1953, Maurice Wilkins showed photo 51 to James Watson, which with helped him and Francis Crick deduce that the structure of DNA was double-helical,⁷ a work that they published that same year. Until this day the conventional hydrogen-bonding interaction between nucleic bases is termed Watson-Crick interaction. Leontis / Westhof classification⁸ describes the non-traditional, non-Watson-Crick base-base interactions that can be found in nucleic acids, especially in RNA.

In the early 1980's, a new field of DNA nanotechnology was born with the publication by Nadrian Seeman⁹ on "Nucleic acid junctions and lattices", who realized that three-dimensional structures made from DNA could be constructed in a programmable fashion. Since then, nucleic acids (DNA and RNA) are evermore increasingly used as scaffolds to construct nanoparticles.¹⁰⁻¹⁵

Dr Kirill Afonin operates a research laboratory at the University of North Carolina at Charlotte. The nature of the research in the lab is on topics of RNA nanotechnology, although work with DNA is also very common. While working in Dr Afonin's Lab, there came a need to develop a laboratory logo to represent the lab. I volunteered for the task. Many company and organization logos are simple and abstract. Since the lab works mostly with RNA, which is a single-stranded nucleic acid, an idea came to attempt to make the logo somehow represent a nucleic acid. Double-stranded DNA has a very characteristic major and minor grooves, and nucleic bases forming base-pairs between the two strands making the appearance of a ladder, although RNA does not have as pronounced major and minor grooves as DNA. The logo (see Figure 1, below) has vertical lines of different length, representing nucleic bases on a backbone (an imaginary horizontal line through the logo). The shape of the larger curve formed by the 4 lines below the letters "nano" represent the major groove, and the shape of the curve formed by the 3 lines above the letters "LAB" represent the minor groove of the nucleic acids. The work to implement the design has been done in Adobe Illustrator suite.



Figure 1: Lab logo developed for Dr. Afonin's laboratory.

CHAPTER 1: INTRODUCTION TO NUCLEIC ACIDS AND KISSING LOOP PROJECT

1.1 Structural properties of nucleic acids

Ribonucleic and deoxyribonucleic acid molecules (RNA and DNA, respectively) are biological polymers, generally referred to as nucleic acids (NAs), and composed of repeating units, called nucleotides, which in turn are each composed of a phosphate, a sugar, and a nucleic base.¹⁶ Nucleotides are bound together in a chain, alternating phosphate and sugar monomers throughout the length of a polynucleotide. This alternating sugar-phosphate chain forms the backbone of the polymer, where nucleic bases are attached to the sugars, and sugars are attached between each other with phosphates.¹⁶ As sugar-phosphate double-helical backbones wind around the axis in antiparallel orientation, nucleic bases form base-pairs towards the axis of the double.¹⁷

¹⁸ Ribose is a simple sugar that can either be in a linear or a furanose form (five-membered ring with four carbons and an oxygen, and the fifth carbon not part of a ring but attached to it. Ribose carbons are named $C_{1'}$, $C_{2'}$, $C_{3'}$, $C_{4'}$, and $C_{5'}$, to distinguish them from carbons of nucleic bases.

In nucleic acid chains, nucleotides are linked together by phosphodiester bonds. Every individual nucleotide is attached within the chain at $C_{5'}$ to the previous nucleotide and at the $C_{3'}$ to the next nucleotide in the polynucleotide chain. The resultant NA chain is formed from alternating phosphates and sugars: $(\text{phosphate} \rightarrow C_{5'} \rightarrow C_{4'} \rightarrow C_{3'} \rightarrow)_n$, where n is the number of nucleotides in a chain. The understanding of the 5' to 3' directionality of the nucleic acids derives from the fact that in living systems replication

and synthesis of new DNA, and synthesis of RNA happen exclusively in the 5' to 3' direction.¹⁹

A nucleotide is composed of a ribose sugar with a phosphate group at the C_{5'} of the ribose sugar and a nucleic base at the C_{1'}. The unit composed of only a ribose sugar and the nucleic base at the C_{1'} is called a nucleoside. A nucleotide can be a mono-, di-, and tri-phosphate depending on the consecutive number of phosphates at the C_{5'} of the ribose. There are 5 nucleic bases, of which there are purines and pyrimidines, which can be substituted at the C_{1'} of the ribose and attached by a β -glycosyl C–N bond. Purine bases are adenine (Ade) and guanine (Gua), and pyrimidine bases are cytosine (Cyt), uracil (Ura), and thymine (Thy). Purine bases attach to the N₉-positions and pyrimidines attach to the N₁-positions of the C_{1'} of the ribose sugar. Similar to nucleic bases, purine nucleotides are called adenosine (A) and guanosine (G), and pyrimidine nucleotides are called cytidine (C), uridine (U), and thymidine (T).¹⁶ RNA nucleotides have hydroxyl (–OH) groups at C_{2'}, where DNA nucleotides only have a hydrogen (–H).¹⁶ DNA by its definition is missing an oxygen somewhere, which happens to be at the 2'-carbon.

Ribose ring can assume different conformations and is almost never planar, where torsion changes geometry of each carbon in the ring relative to the plane of the molecule.^{16, 17} These geometry changes make some carbons “bulge” in and out of plane of the sugar. Bending or bulging of endocyclic carbons (within the ring) out of plane of the sugar molecule denotes their *endo*- (toward the C_{5'}, in a “cis” position) or *exo*- (away from the C_{5'}, in a “trans” position) orientation.^{16, 17} Because a ribose sugar is unsymmetrically substituted (or has groups attached to it) within a nucleotide, this

limits the number of pucker conformations that are favorable.^{16, 17} Among many conformations, *C*_{2'}-*endo* and *C*_{3'}-*endo* pucker conformations are preferred over others due to minimal nonbinding interactions between the furanose ring and its substituents.^{17, 20} Based on NMR spectroscopy data,²⁰⁻²⁵ substituents at the *C*_{2'} and *C*_{3'} influence the sugar pucker mode; the more electronegative substituent at one of the furanose carbons assumes the axial (rather than equatorial) orientation, which alters the furanose ring conformation.¹⁷ Since DNA nucleotides have only a slightly electronegative hydrogen at the *C*_{2'} (which assumes equatorial position), DNA conforms to a *C*_{2'}-*endo* pucker, because the 3'-oxygen (phosphate-oxygen bound to *C*_{3'}) takes the axial position. On the contrary, RNA nucleotides with a 2'-hydroxyl (which assumes the axial position) prefer the *C*_{3'}-*endo* pucker (a more equatorial substituent position).^{17, 20}

Due to the change in pucker conformation, the distance between phosphate of one nucleotide to the next changes from 5.9 Å of *C*_{3'}-*endo* (A-form, RNA) to 7.0 Å of *C*_{2'}-*endo* (B-form, DNA)^{17, 26, 27} (see Figure 1.1, below). These changes affect the pitch and width of the nucleic acid helix. Under physiological conditions, DNA double-helix adopts the B-form DNA (see Figure 1.2, below) that has a helical pitch (number of nucleotides/base-pairs needed to complete one helical rotation) of ~ 10.4 base-pairs, with a rise per base residue in the range of 3.0 – 3.4 Å, and ~ 2 nm helical diameter.^{27, 28} RNA double-helix assumes the A-form (DNA) structure (see Figure 1.2, below), which has a helical pitch of ~ 11 base-pairs per turn, with a rise per base residue in the range of 2.6 – 3.3 Å, and ~2.6 nm helical diameter.^{17, 27}

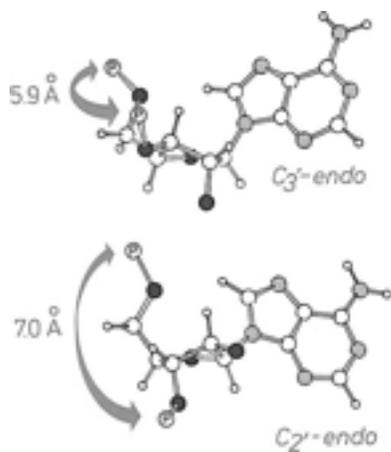


Figure 1.1: C_3' -endo and C_2' -endo sugar pucker conformations. Figure adapted from “Principles of Nucleic Acid Structure” book.¹⁷

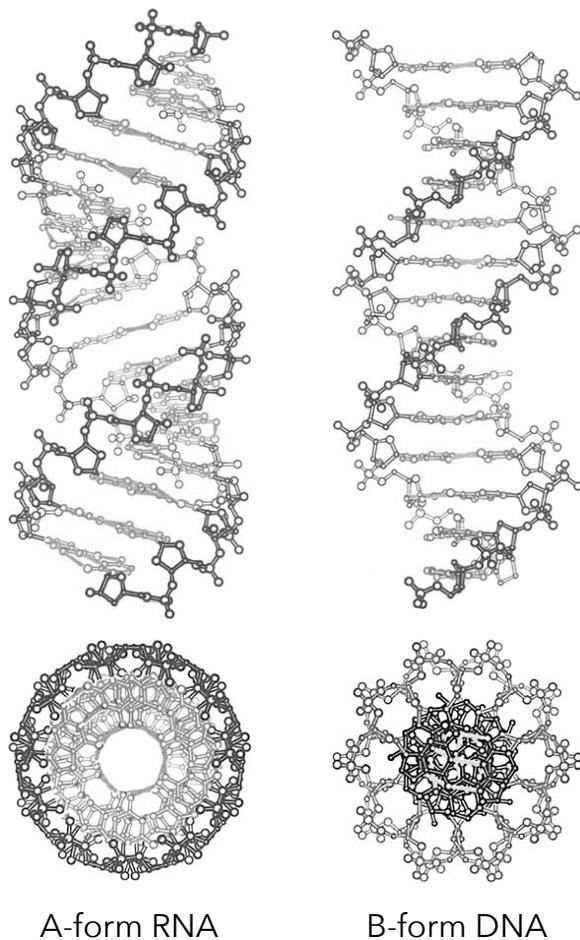


Figure 1.2: Forms of nucleic acids, and their structures. Figure adapted from “Principles of Nucleic Acid Structure” book.¹⁷

At different pH conditions nucleotides can lose or gain protons on the base, sugar, or the phosphate. Nucleic bases can be protonated and deprotonated within 3.5 pH units of neutral conditions (pH 7).^{17, 29} N₃-hydrogens of uracil/thymine and N₁-hydrogen of guanine are removed in alkaline conditions. Adenine N₁- and cytosine N₃-positions are protonated at slightly acidic pH of > 3. N₇- of adenine and guanine, and O₄- of uracil are also protonated when pH < 2.²⁹⁻⁴⁰ The pK_a values of nucleosides and nucleotides with 5'-phosphate at their protonation/deprotonation sites are as follows: 3.52 and 3.88 for adenosine N₁-position, 4.17 and 4.54 for cytidine N₃-position, 9.42 and 10.00 for guanosine N₁-position, 9.38 and 10.06 for uridine N₃-position, and, 9.93 and 10.47 for 2'-deoxythymidine N₃-position, respectively.^{17, 29, 30} Two main pK values can be observed for phosphate groups of nucleoside mono-, di-, and tri-phosphates. In the first (primary) ionization state (at pH of 1-2), one hydrogen is removed from each phosphate, yielding one negative charge per phosphate group.^{29, 41, 42} This primary ionization state succeeds up to pH of ~ 7, when the last hydrogen is removed from the terminal phosphate of the nucleotide.^{29, 41, 42} Secondary hydroxyls of riboses are deprotonated above pH of ~ 12.5^{29, 43, 44}.

Hydrogen bonds are a weak type of interactions often found in biological systems, which stabilizes interactions between proteins and nucleic acids. Hydrogen bonds, among other types, are mainly electrostatic.⁴⁵⁻⁴⁹ In nucleic acids, hydrogen bonds (shown as dotted lines) form between N-H...O or N-H...N atoms of the bases,^{17, 28} and their strength depends on the partial charges on each of the three interacting atoms "in-line" of the hydrogen bond. When compared to covalent bonds, hydrogen bonds are

about 20 – 30 times weaker and have less defined length, strength, and orientation, making them more prone to bending and stretching.^{17, 28}

Just as nucleic bases can form base-pairs within a plane horizontally (with some torsion), nucleic bases can have interactions between base-pair planes vertically. Nucleic bases have aromatic systems that can interact through π -bonding, allowing bases to stack.^{17, 28} Within such a stack, polarized regions ($-\text{NH}_2$, $=\text{N}-$, $=\text{O}$, or a halogen⁵⁰ such as bromine) of one base superimpose over the aromatic system of the adjacent base, which promotes stacking. The strength of such stacking is stronger than that of hydrogen bonds.^{17, 28} Though permanent dipoles of nucleic bases (predominately $\text{C}=\text{O}$ or $\text{C}-\text{NH}_2$ groups) superimposed over π -electronic systems of adjacent bases play a minor role, dipole-induced dipole interactions appear to play a major role in base-stacking stabilization.^{17, 28, 50} Furthermore, NMR studies^{17, 28, 51, 52} suggest that six-membered pyrimidine rings of bases preferentially participate in base-stacking rather than their five-membered imidazole fragments, and that base orientation within the stack depends on its' amino- or keto- substituents.^{17, 28, 53} Solubility experiments and NMR data^{28, 54-58} showed that base interaction in stacks was most stable between purine-purine, followed by purine-pyrimidine, and lastly between pyrimidine-pyrimidine bases.^{17, 28}

To help in vertical base stacking in nucleic acid chains there are forces that play an important role, including dipole, π -electron systems, and dipole-induced dipole moments.^{17, 28} Properties of base stacking, where purines stack better than pyrimidines and methylation improves stacking, can be explained by electronic systems of bases; they are due to London dispersion forces and interactions between dipoles of bases in

the stack.^{17, 28} The induced dipoles in one group of atoms (or a nucleic base) polarize the electronic system(s) of neighboring atoms or molecules (or nucleic bases), which induce parallel dipoles between nucleic bases that attract them together.^{17, 28} These attractive forces that aid in base stacking are additive, although they decrease with $1/6^{\text{th}}$ power of the distance.^{17, 28}

A nucleic acid double helix can be separated into two distinct single strands, and this process depends on the melting temperature of that particular nucleic acid sequence. The two strands come together and break apart near the melting temperature, where hydrogen bonds form and disrupt, respectively. For the two separate strands to associate with each other, one base-pair needs to form *via* hydrogen bonding. Due to proximity of other bases on the sugar-phosphate backbone, a second base-pair can form, but the first base-pair can also fall apart.^{17, 28} If there are three consecutive base pairs that form between two separate NA strands, a nucleus (or a “seed”) is formed, taking into account not only hydrogen bonding but also base stacking; the seed helps propagate duplex formation.^{28, 59} This is referred to as a cooperative zipper mechanism of helix formation, which requires a three-base-pair nucleus as the “seed”.^{17, 28} Nucleus formation needs to overcome the unfavorable positive free energy contribution, after which free energy becomes negative for additional steps, where the double helix grows spontaneously.^{28, 59} In double-helical formation, reduction in UV absorption (hypochromicity or hypochromic effect^{60, 61}) can be observed due to base stacking.^{17, 28} On the contrary, hyperchromic effect can be observed through the increase of absorbance at or around 260 nm, which would happen

upon double-stranded DNA denaturation.⁶¹ Therefore, one way to measure the melting temperature of double-stranded fragment is to measure the UV absorbance at different temperatures, which would witness the formation and breakdown of double helices, since the melting temperature (T_m) is determined at the midpoint of transition from double-stranded to single-stranded species.^{17, 28} Since A:T(U) base-pairs have only two hydrogen bonds rather than three in G:C base-pairs, it is expected that A:T(U)-rich regions should melt at lower temperatures than the more stable G:C-rich regions.^{17, 28} This results in local single-stranded regions within the still double-stranded DNA or RNA random sequence, and is referred to as “breathing modes”.^{17, 62, 63}

1.2 RNA in Nanotechnology

Since the elucidation of DNA double helix structure in 1953, much work has been done, especially in sequencing human DNA in the Human Genome Project.^{64, 65} DNA is said to be the blueprint that codes for life in all living organisms.⁶⁶ Cells have DNA that codes for every aspect of their existence, although in cells of eukaryotic organisms only certain DNA regions may be available for transcription dictated by the differentiation of specific tissues based on their needs.⁶⁷ Information stored in DNA has gene-coding and non-coding regions, first of which is translated into RNA for its further processing and translation into proteins that play various roles within the cell.⁶⁸ Since gene coding regions make up about 2 percent of total DNA,⁶⁹⁻⁷² the rest of DNA was considered “junk” DNA⁷³ for some time.

In more recent times non-coding DNA was also found to be transcribed but not translated into proteins.⁶⁹ RNA was found to play an important role in gene regulation.⁷⁴

Additionally, with the discovery of RNA interference (RNAi) mechanism by Professors Andrew Fire and Craig Mello ⁷⁵ showed that cells can detect double-stranded RNA (dsRNA) using RNA-induced silencing complex (RISC) and via homology-driven mechanism control expression of various genes.⁷⁴ This novel approach has already proven useful in determining the function of various genes,⁷⁶ but it can also be tweaked to target certain genes to be silenced.⁷⁷⁻⁸¹

Nucleic acids (NAs) have proven to be an important biopolymer material in biotechnological applications. NAs are native to biological systems, which makes them great as material due to biocompatibility and high programmability.^{78, 82} DNA is great for its high stability, and RNA is especially important for its ability to form secondary and tertiary structures with itself.⁷⁸

1.3 Kissing Loop Motif

RNA, in contrast to DNA, can form a plethora of secondary and tertiary structures with itself intra- or inter-molecularly, many of which can be traced to specific biological RNA motifs. Some of the motifs aside from duplexes (dsRNA) that RNA forms are kinks (such as 90° kinks), 3-way loops, 3-way junctions, 4-way junctions, hairpins, and kissing loops (KLs).⁸³⁻⁸⁵ These RNA motifs are dependent on the number of interacting nucleotides and sometimes are sequence specific. They can be utilized in larger nanoconstructs by designing nucleic acid (often RNA) to include motif-specific sequences in the monomer total sequence.

A structural RNA motif was originally derived from genomic RNA of HIV (human immunodeficiency) type-1 virus. The motif is a part of the dimerization initiation site

(DIS) of the viral genome.⁸⁶ The motif consists of two loops, each of which forms a hairpin structure, and two hairpin structures form intermolecular complementary binding *via* six nucleotides (6nts) on the hairpin loop. Two and one flanking adenines support the geometry of the assembled complex *via* base stacking. This long-range interacting motif is called a kissing loop (KL) complex (KLC), which is characteristic of a 180° orientation or geometry. The original, wild-type sequence reported for the kissing loop is 5'-...AGGUGCACA...-3', where both hairpins have the same hairpin-loop sequence. A protein data bank (PDB) structure of the KL complex⁸⁷ was uploaded, code-naming it 2FCX (see Figure 1.3, below).

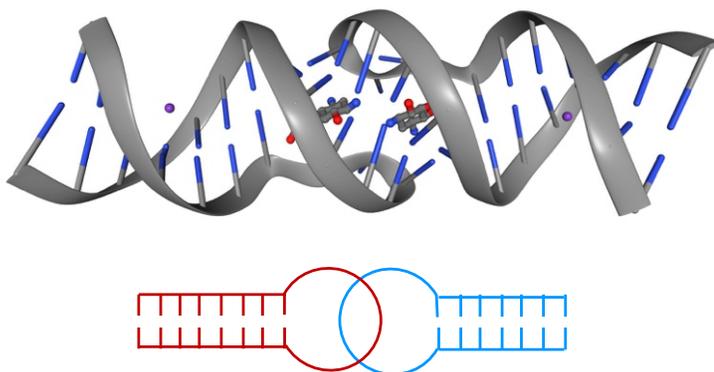


Figure 1.3: RNA kissing loop 3D and 2D representations. Kissing loop interaction (PDB: 2FCX⁸⁷) as shown on PDB website (above). PDB structure title: “HIV-1 DIS kissing-loop in complex with neamine”. Cartoon showing simple 2D representation of a kissing loop interaction, consisting of two kissing loops, shown in red and blue color (below).

The original homodimeric sequence^{87, 88} of the KLCs (5'-...AGGUGCACA...-3') has two and one purines (not adenines specifically) flanking the six intermolecular-interacting nucleobases (underlined in bold). This KL complex was modified and integrated into “linear assemblies” or RNA fibers.^{10, 77} The kissing loop sequences used in

RNA fibers are 5'-...AAGGAGGCA...-3' and 5'-...AAGCCUCCA...-3', having two and one adenines to retain the approximately 180° geometry of the KL complex.



Figure 1.4: A-B type simple RNA fiber 2D representation. Cartoon showing red and blue monomers.

1.4 RNA fibers

A “linear assembly”,¹⁰ or an RNA fiber,⁷⁷ can be made by incorporating the two kissing loop sequences into two different RNA monomer strands. Each RNA monomer folds on itself and forms a “dumbbell” shape due to designed sequence complementarity, where adenines bind uracils and guanines bind cytosines. RNA monomers each have two kissing loops at their termini, where one RNA monomer has Fiber sequence with HIV1 KL on both sides and the other RNA monomer has Fiber sequence HIV2 KL on both sides (see Figure 1.4, above). This RNA fiber system only has two monomers and is an A-B-type fiber system, where monomer A and monomer B alternate to form the longer fiber structure.

This RNA Fiber can be used to deliver cargo (genetic “messages”) into the cells,⁷⁷ where the Fiber monomers can be functionalized (see Figure 1.5, below) with specific sequences (need to be dsRNA) that can be cleaved or “diced” by Dicer endonuclease into shorter RNA fragments (usually 21-23nts long), called short interfering RNA (siRNA).⁸⁹ Dicer then helps load siRNA guide strand into RISC (RNA-induced silencing

complex), which in turn cleaves any homologous mRNAs (messenger RNAs), preventing their translation, that is - expression of those genes.⁸⁹ Work done by Rackley et al⁷⁷ showed that it is possible to form these functionalized RNA fibers, and more importantly that they showed low immunostimulation. The benefits of this platform are that it is modular, highly programmable, and has low immunostimulation due to its' linear rather than bulky structure.

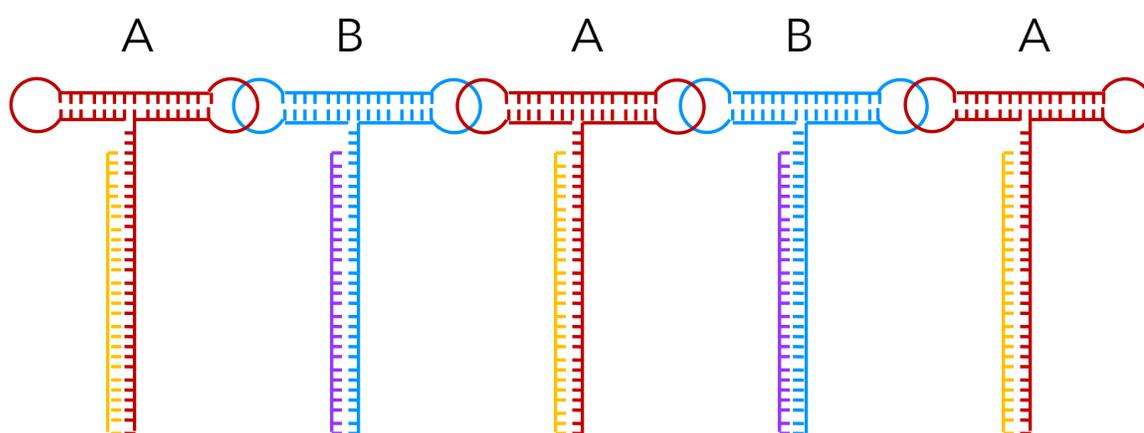


Figure 1.5: A-B type functionalized RNA fiber 2D representation. Cartoon showing red and blue monomers with two different toehold extensions and two different Dicer substrate complementary-bound short strands.

1.5 Problem Statement

Although RNA fibers are already a promising platform due to its ability to carry siRNA into cells and for their low immunostimulation,⁷⁷ being an A-B system, they are still limited to carry just two different siRNA. Additionally, to make RNA fibers with only two monomers, they need to be able to form “indefinitely” for them to be long linear assemblies. Each fiber forms a different length fiber and is likely controlled or defined by

structural rigidity or flexibility, beyond which, fiber either may break kissing loop interaction or be limited by monomer depletion (locally or systemically).

More importantly in an A-B-type system RNA fiber are comprised only of two different monomers, each able to carry one siRNA into the cells for downregulation of gene expression. Expanding RNA fibers to accommodate higher capacity of heterogeneous siRNA cargos could be used to deliver more numerous siRNA cargos for targeting and downregulation the expression of multiple genes simultaneously.^{78, 81, 90-95}

(Solution)

The limits imposed by the A-B system can be overcome by expanding the utilized KLC 1 + 2 into a library of KLC sets that can coexist or be utilized in larger one pot assemblies without cross-reacting with each other, being unique and inter-compatible. If there were 8 unique and inter-compatible KLC sets, 8 different siRNA cargo would be available for intracellular delivery. Additionally, it would be possible to control RNA fiber length, since a unique KLC would be utilized between each monomer in a string of 9 monomers. This would be a(n) (poly-heterogeneous) ABCDE-type system, in contrast to the A-B system currently available. This poly-substituted RNA fiber would be controlled of controlled length and be more heterogeneous in its intracellular siRNA cargo delivery. Finally, it would be possible to incorporate the KL library into larger nanoconstructs for quaternary assemblies formed from multiple tertiary components.

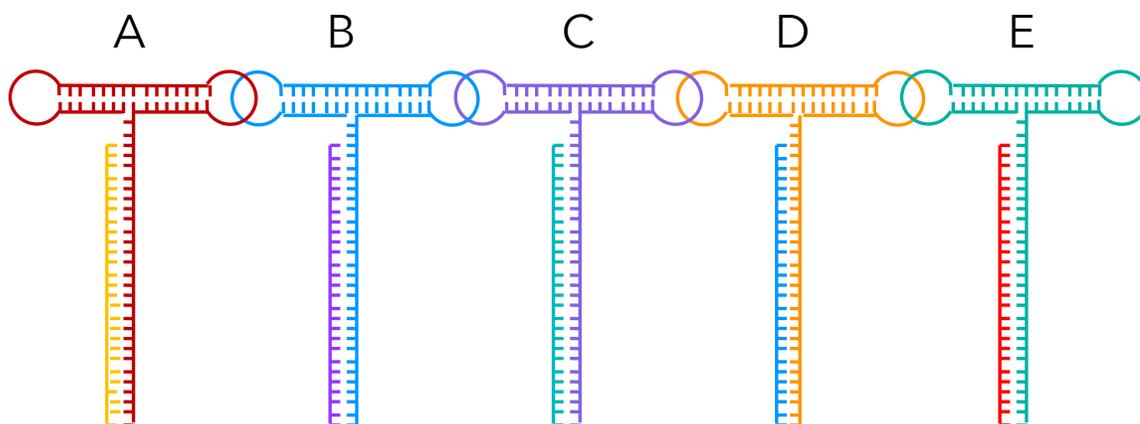


Figure 1.6: ABCDE-type functionalized RNA fiber 2D representation. Cartoon showing differently colored monomers with five different toehold extensions and five different Dicer substrate complementary-bound short strands.

1.6 Goal of Project

A limited A-B system currently in use showed a need to develop more kissing loop complex (KLC) sets to expand into a KL library that can be used simultaneously for a larger structured nanoparticle. The near 180° KL set currently in use is the HIV1/HIV2, consisting of two loops, HIV1 and HIV2, which form the “1 + 2” interaction complex. Sequences of both HIV1 and HIV2 have the same KL motif, consisting of 5'-CAAXXXXXXAG-3' nucleobases, where the flanking guanine and cytosine bases interact, bringing the hairpin formation together. The six interior bases are important to the KL interaction and flanking two and one adenines are part of the KL interacting motif, which support the KL interaction by allowing bases to stack, contributing to the near 180° geometry. To expand the KLC into a library, loop sequences of strands 1 and 2 were taken as a basis. To test the new KL sets, short hairpin strands need to be tested before incorporating them into more complex structures.

Working with this RNA kissing loop motif, nucleotides of 6nt active region of the KL were mutated and 32 different combinations were selected for a total of 16 interacting KL complex (KLC) sets. To determine that each KL only interacts with its complementary KL to form a KLC in one pot application, a crosstalk experiment would be designed and performed to check the cross-reactivity (interference) of all KLCs. Melting temperature (T_m) of each KLC would be assessed using Temperature Gradient Gel Electrophoresis (TGGE) to know the thermal properties of KLCs to be able to utilize them in larger nanoconstructs.

1.7 Hypothesis:

Based on the designed 16 sets of KLCs, there should be about 8 KLC sets with high certainty and about 8 KLC sets with low certainty (due to sequence similarity); about 8 KLC sets are expected to be unique as a library to be used simultaneously. Further, substituting new 6nt sequences into the KL loop region for KL library expansion should still retain their structural/functional integrity and preserve the linear, near 180° geometry of the KL motif.

CHAPTER 2: METHODS OF T7 RNA POLYMERASE EXPRESSION FROM *E. COLI*

T7 RNA polymerase is typically expressed by the T7 bacteriophage, but *E. coli* cells have been engineered (transformed) to express the plasmid containing the gene for expressing T7 RNA polymerase. In general, bacterial cells are allowed to grow to reach a specific optical density (mid-log phase), then they are induced to overexpress this protein. Cells are then lysed and T7 RNA polymerase is separated from the cell lysate and purified for performing *in vitro* transcription reactions (in Eppendorf tubes) to synthesize RNA strands without the need for cells to produce the material. Below are the steps detailing this process.

This process is extensive and takes 3 to 4 days of work to complete. The first day of work includes the preparation of LB (Luria-Bertani) broth media (steps 1 – 3), and preparation of preculture (steps 4 – 6). The second day includes antibiotic solution preparation (steps 7 – 8), bacterial cell culture growth (steps 9 – 13), induction of bacterial cells with IPTG (steps 14 – 18), and bacterial cell collection (steps 19 – 20). At this point there is an optional stopping point, and confirmation of bacterial cell induction with IPTG via SDS-PAGE and staining (steps 21 – 28). The third day of work includes preparation of lysed bacterial cells (steps 29 – 32), purification of T7 RNA polymerase (steps 33 – 38), and evaluation of T7 RNA polymerase elution stages via SDS-PAGE (steps 39 – 47). The end of the third day and fourth day include overnight dialysis (steps 48 – 50).

2.1 LB (Luria-Bertani) broth preparation

1. Prepare LB (Luria-Bertani) broth media by weighing 10.0 g of LB broth and bring up to 500 mL with ddiH₂O (double deionized water, 17.5-17.8 MΩ·cm).
 - Pick a glass bottle that will accommodate LB broth on the shaker for proper aeration. If the glass bottle is almost full, it will not allow proper aeration.***
2. Autoclave the glass bottle with LB broth for 15 min at 121 °C.
 - Place the cap on the glass bottle, but do NOT close the cap completely, as the pressure buildup of the heated liquid can cause the glassware to break.
Loosely cover the loose cap of the bottle with aluminum foil.
3. After bottle with LB broth is autoclaved, wait for it to cool down.

2.2 Preculture preparation

4. Prepare a preculture (starter culture) by pouring 25 mL of LB broth into a 50 mL Falcon tube “over flames” to prevent contamination of the preculture.
5. Add a micropipette tip poke (about a drop, on a pipette tip) of frozen glycerol stock bacteria (*E. coli*) from a 2 mL tube to the falcon tube with LB broth.
6. Place 25 mL Falcon tube onto shaker overnight.
 - Set the incubator / shaker to 37°C and ~150 (or 130) rpm.
 - Let precultures (starter culture(s)) grow for 8 – 18 hours.

2.3 Bacterial cell culture growth

7. Set aside ~10 mL of control aliquot of LB broth into 15 mL Falcon tube for OD (optical density) measurements.

8. Prepare antibiotic solution using ampicillin or carbenicillin (interchangeable) at 50-100 $\mu\text{g}/\text{mL}$ (*this step can be optimized*).
 - carbenicillin used with 250 μg in 5 mL, yielding 50 $\mu\text{g}/\text{mL}$.
9. Add 500 μL of carbenicillin (antibiotic) into 500mL LB.
10. Add 12.5 mL of preculture into 500 mL LB with antibiotic “over flames” to prevent contamination of culture growth.
 - Place on shaker at 37 °C and set the shaker at ~ 150 (or 130) rpm (*this step can be optimized*).
11. Wait until OD is between 0.4 – 0.6.
 - Use LB broth from LB control in 15 mL falcon tube as baseline for measuring OD. (Sample from step 7).
 - It takes *E. coli* about 30 minutes to double its population, based on which bacterial culture growth will reach the appropriate OD.
 - At OD between 0.4 – 0.6, bacterial culture growth is about mid-log phase, which usually takes about 4 hours, after which bacteria will enter the death phase. ***
12. Check OD of bacterial culture growth at 1.5 – 2 hours.
 - (total time was 3.5 hours)
 - $\text{OD}_{600} = 0.564$
13. At max growth, drop the temperature in the incubator to 25 °C for 30 minutes to stop bacterial growth.

2.4 Induction of bacterial cells with IPTG

14. At 0 hours induction with IPTG, take out 1 mL of bacterial culture into a 2 mL Eppendorf tube.

- Centrifuge at 5,000 rpm for 2 min.
 - Eppendorf Centrifuge 5415R used for this step.
- Remove 900 μ L of supernatant.
- Add 25 μ L of Laemmli (5x SDS-PAGE loading buffer) for a total of 125 μ L.
- Pipette up and down.
- Place at 95 °C for 5 min.
- After heating / boiling, place on ice.

15. Add IPTG (isopropyl- β -D-thiogalactopyranoside) at 0.5 – 1 mM (*this step can be optimized*), which will induce T7 polymerase gene expression in *E. Coli*.

- Prepare 1 mL of 0.5 M IPTG
- Add 1mL 0.5 M IPTG to flask (500 mL bacterial culture)

16. At 4 hours after induction with IPTG, take out 1 mL of bacterial culture into a 2 mL Eppendorf tube.

- Centrifuge at 5,000 rpm for 2 min.
 - Eppendorf Centrifuge 5415R used for this step.
- Remove 900 μ L of supernatant.
- Add 25 μ L of Laemmli (5x SDS-PAGE loading buffer) for a total of 125 μ L.
- Pipette up and down.
- Place at 95 °C for 5 min.

- After heating / boiling, place on ice.
17. Keep both 2 mL Eppendorf tubes on ice or in -20 °C to run a gel and confirm induction of T7 RNA polymerase with IPTG.
18. Allow bacterial cell culture to incubate at 25 °C for 4 – 16 hours (*this step can be optimized*).
- In one instance, bacteria were incubated for 4 hours after induction with IPTG.
 - In another instance, bacteria were incubated for about 18 hours.

2.5 Bacterial cell collection

19. To collect bacterial cells, remove excess LB media.
- Set centrifuge to 4 °C and precool before use.
 - Use two 15mL Falcon tubes to centrifuge bacterial culture solution to remove LB broth.
 - Add 10mL of LB broth to each tube.
 - Using 2 tubes and loading same volume in each keeps the tubes balanced.
 - Centrifuge at 5000 rpm for 5min (repeat cycle until solution is removed).
 - Eppendorf Centrifuge 5810R used for this step.
 - Remove tubes from centrifuge,
 - remove LB broth from each tube,
 - place into empty beaker,
 - add next 10mL amount to tubes,

- and place tubes back into centrifuge.
- After all LB broth has been removed and cell pellets are remaining, cells can be frozen at -20°C or -80°C .
 - Add 10% bleach to the LB broth to be discarded, to kill any lingering bacterial cells.
 - Decant LB broth from the tubes with cell pellets into the container with the rest of LB broth.
 - Allow LB broth with bleach to sit / decontaminate for about 10 minutes (LB broth liquid should change color) before pouring down the drain.
- Alternatively, bucket centrifuge can be used.
- Set the centrifuge at $15,000 \times g$ for 10 min at 4°C . Make sure that the centrifuge tubes / flasks are well balanced before loading them into the rotor.

20. This can be a **Stopping Point**

- Plan time wisely, as **this is the only stopping point.**
- Bacterial cells can be resuspended in 0.9% NaCl (do NOT vortex the cells), centrifuged at $15,000 \times g$ for 10 min at 4°C , and then stored at -80°C for several months.
 - To soften the pellet, cells may be placed in larger bucket centrifuge tubes / flasks, then placed on larger shaker at low temperature (below 25°C) for ~ 10 min.

- Be sure to decant 0.9% NaCl after centrifugation and decontaminate in bleach for 5 minutes, after which pour down the drain.

(Stopping Point)

2.6 SDS-PAG preparation

21. For confirmation of proper T7 RNA polymerase expression pour a two-layer SDS-PAG (sodium dodecyl sulfate – polyacrylamide gel). Make two gels, one for confirming proper induction of T7 RNA polymerase with IPTG, and another for protein expression to evaluate stages of purification and elutions. Follow the gel recipe below (see Table 2.1, below).

- First prepare the gel stand with clean glass plates.
 - Gel was prepared on a Bio-Rad Mini-PROTEAN® gel system.
- Prepare resolving (bottom) layer of the gel first.
 - Be sure to wear PPE when preparing gel solutions and casting gels, since acrylamide is a neurotoxin.⁹⁶
- Pour the resolving layer of the gel almost to the top, but so as to leave room (~0.5 cm) below the gel comb, as if it was placed. Then, pour some isopropyl alcohol into the gel space above the resolving gel solution. This helps the top of the gel to polymerize flat.
- After resolving layer has polymerized, empty any remaining isopropyl alcohol on top of the resolving gel layer, and prepare stacking (top) layer of the gel next.

- Pour stacking layer on top of the resolving layer and insert the desired gel comb.
- After gels are polymerized, remove the gels, wrap them in a wet paper towel, and wrap them with aluminum foil around the wet towels. This helps keep the gels from drying out until they are used (hopefully within few days).

Table 2.1: Reagent recipe for making a 10% SDS-PAGE, SDS-PAGE running buffer, and destain solution.

10% SDS-PAGE(E)					
		resolving gel layer (Tris-HCl pH 8.8)		stacking gel layer (Tris-HCl pH 6.8)	
ddiH ₂ O	mL	2.5	5	1.9	3.8
buffer	mL	1.325	2.65	0.75	1.5
10% SDS	μL	60	120	30	60
10% APS	μL	30	60	15	30
40% AA	mL	1.325	2.65	30	60
TEMED	μL	10	20	5	10
Total Volume	mL	5.25	10.5	3	6

Buffer for resolving layer	1.5 M Tris-HCl pH 8.8
----------------------------	-----------------------

Buffer for stacking layer	1 M Tris-HCl pH 6.8
---------------------------	---------------------

SDS-PAGE running buffer (RB)				
------------------------------	--	--	--	--

10x RB				
	MW (g/mol)	amt (g)	C2 (M)	V2 (mL)
Glycine	75.07	432	1.92	3000
Tris base	121.14	90	0.25	
SDS	288.38	30	0.03	
ddiH ₂ O up to final volume				

pH should be ~8.3 without adjustment
10x RB diluted to 1x should be pH ~8.5

1x RB

192 mM
25 mM
0.10%

Destain solution		
------------------	--	--

	amount (mL)	V2 (mL)
Methanol	200	1000
Acetic Acid	100	
ddiH ₂ O	700	

2.7 Induction confirmation of T7 RNA polymerase gene expression via SDS-PAGE

22. Prepare bacterial cultures (if have not done already) taken at 0 hours and at 4 hours of induction of *E. coli* with IPTG and run on SDS-PAGE to confirm that expression of T7 RNA polymerase was successful.

- Prepare bacterial cultures prepared earlier for gel electrophoresis (from steps 14, 16, and 17).
- Mix 10 μL of each sample (1:1) with 10 μL SDS loading dye.
- Place each Eppendorf tube at 95 $^{\circ}\text{C}$ for 5 min to denature the proteins.
- Use Dual Xtra ladder (Precision Plus Protein Dual Xtra Standards, by Bio-Rad).

23. Prepare Bio-Rad gel rig to run the gel and fill with 1x SDS-PAGE running buffer (see Table 2.1, above).

24. Load 1 μL of Ladder, 5 μL of prepared cell lysate from 0 h (at induction with IPTG), and 5 μL of prepared cell lysate from 4 h (after induction with IPTG) onto the SDS-PAGE.

25. Run the gel at 180 V for 40 – 50 minutes at room temperature, until the dye reaches near the bottom of the gel.

- Run for 40 minutes first, then add more time as desired.

26. After gel run is complete, remove the gel and perform fast stain with Coomassie blue stain.

2.8 Staining SDS-PAGE with Coomassie blue dye

27. Perform a “Fast Stain” and destain of the induction confirmation gel.

- Place the gel flat into an empty micropipette tip box.

- Pour enough Coomassie blue stain to cover the gel in the tip box.
- For Fast Stain, stain for only 5 minutes. Fast Stain is good for a quick assessment, to be able to proceed to protein purification and dialysis.
 - Place the tip box with gel and dye on the rocker for 5 minutes.
 - After 5 min stain, pour the dye back into its container for reuse, and rinse the gel with deionized water (can be discarded down the drain).
 - Add destain solution (see Table 2.1, above) into the tip box with gel and a slightly crumbled Kim Wipe on top to soak up excess Coomassie blue dye.
 - Place the box onto the rocker and allow to destain for 10 – 15 minutes, or until bands are visible on the gel with the help of white lamp.
 - After destain is complete, discard destain down the drain.
 - Visualize the gel on white lamp.

28. For a thoroughly stained gel, stain the gel for 1 hour with Coomassie blue dye.

- Destain overnight.
- Visualize the gel on white lamp. Take a picture for record.

2.9 Preparation of lysed bacterial cells

29. Prepare bacterial cells to be lysed.

- Thaw bacterial cells on ice for ~30 min (if **Stopping Point** was made).
- Resuspend bacterial cells in 20 mL of T7 Lysis buffer **B**.
 - At this point cell pellets can be vortexed.

30. Keep vortexed cells in 50 mL falcon tube on ice in an ice bucket. Cells can be sonicated directly in the falcon tube, on ice.

- Position the sonicator tip into the falcon tube with cells, but make sure that the tip is not touching the bottom or the sides of the falcon tube, otherwise this will damage the sonicator tip.
- Sonicate cells for 3 min at 25% amplitude with 1 sec pulse on and 2 sec pulse off.
 - Total sonication time is ~ 9 min (3 min pulse on, and 6 min pulse off).
 - Clean the sonicator tip with ethanol before and after use. Respect the equipment.
 - Take a **10 μ L aliquot** of this whole cell lysate (**WCL**) for SDS-PAGE.

31. Transfer the now lysed cells (WCL, or lysate) into ultracentrifuge tubes with red caps while keeping them on ice.

- Take two ultracentrifuge tubes for balance. **Make sure tubes are balanced!**
- If only using one tube for WCL, fill with WCL and weigh on balance. Fill the other tube with water while matching the weight of the first tube.
- If using both tubes for WCL, split the contents between the two tubes. To ensure balance, add T7 Lysis buffer **B** with a pipette. **Match the weight of both tubes!**

32. Spin the cells (WCL) in the TI-70 rotor at 30,000 rpm at 4 °C for 1 hour.

- Make sure that the ultracentrifuge tubes are placed opposite each other to **ensure that the rotor is balanced!**

2.10 Purification of T7 RNA polymerase

33. Prepare the nickel column by pouring **1 – 2 mL** (*this step can be optimized*) of clean nickel resin slurry into a plastic column.

- Add small amounts of nickel resin to adjust to the desired resin volume.
 - *** Nickel resin is stored in ethanol to prevent bacterial growth.
- Alternatively, a commercial column can be used.

34. Rinse nickel column with 100 mL of T7 Lysis buffer **B** before adding the lysate to it.

- Lysis buffer rinse can be discarded down the drain immediately.
 - Nickel resin can sit in the column without liquid for <1 hour.
 - If more time is needed, add more T7 Lysis buffer **B** to keep the nickel resin hydrated. Be very cautious, nickel resin is expensive!
- Place a plastic beaker under the column to collect flow through (FT).

35. Prepare Eppendorf tubes and plastic beakers:

- FT – plastic beaker/cup for Flow Through
- W – plastic beaker/cup for Wash
- E1, E2, E3, E4, E5, E6 – one Eppendorf tube for each Elution
- For SDS-PAGE: Eppendorf tubes
 - WCL – for Whole Cell Lysate
 - SCL – for Soluble Cell Lysate
 - FT – for Flow Through
 - W – for Wash
 - E1, E2, E3, E4, E5, E6 – one Eppendorf tube for each Elution sample

36. Pour SCL supernatant through the column to chelate T7 RNA polymerase to the nickel resin while taking 10 μ L aliquots during following steps:

- Recall that a 10 μ L **WCL aliquot** was taken during preparation of lysed cells step.
- Take a 10 μ L **SCL aliquot** from SCL into SCL Eppendorf tube for SDS-PAGE.
- Pour SCL supernatant through the Nickel-Resin column and allow supernatant to drip.
 - Collect this fluid in the plastic beaker / cup labelled FT (flow through).
- After all fluid has run through, pour it through the column again (2x).
- Take a 10 μ L **FT aliquot** of 2x flow through into FT Eppendorf tube for SDS-PAGE.
 - Store / save the remaining flow through (FT) at 4 °C, in case there is considerable amount of protein remaining in the flow through.
 - This T7 RNA polymerase has a polyhistidine tag, which other proteins in SCL do not have. Nonspecific binding is deterred from the presence of imidazole in the buffers. T7 RNA polymerase chelates to the nickel resin as it passes through the resin bed, while other proteins pass through and stay in the flow through.
 - Passing flow through twice (2x) through the nickel column ensures more efficient chelation of T7 RNA polymerase to the nickel resin.
 - It is not uncommon to see T7 RNA polymerase in the flow through for especially efficient protein expressions. This can be seen on the SDS-

PAG. If this happens, more nickel resin may need to be used for future protein expressions.

- If more than needed nickel resin is used (to make sure that all protein is collected the first time), then the elution volumes being 0.5x proportional to the resin volume will have the expressed protein more diluted than what may be desired.

37. Wash the nickel column with T7 Wash buffer **C** with the volume 12x the nickel resin volume.

- Ex: if resin volume is 1 mL, use 12 mL of T7 Wash buffer **C**.
- This step will help remove traces of undesired proteins which may have non-specifically bound to the nickel resin, while keeping T7 RNA polymerase bound in the resin bed.
- Collect Wash into plastic beaker labelled W (wash). (Figure above, panel D)
- Take a 10 μ L **W aliquot** of Wash into W Eppendorf tube for SDS-PAGE.

38. Elute T7 RNA polymerase from nickel resin with T7 Elution Buffer **D** with the volume 0.5x the nickel resin volume.

- Ex: if resin volume is 1 mL, use 0.5 mL of T7 Elution Buffer **D** for each Elution (E1, E2, ...)
- 0.41mL (410 μ L) of T7 Elution Buffer **D** was used for each Elution in this experiment.

2.11 Evaluation of protein purification stages via SDS-PAGE

39. Pour a two-layer SDS-PAGE if not already for evaluating stages of purification and T7 RNA polymerase allocation in different Elutions. Resolving layer is on the bottom and stacking layer is on the top of the gel.

40. Get samples in Eppendorf tubes for SDS-PAGE:

- WCL – 10 μ L Whole Cell Lysate
- SCL – 10 μ L Soluble Cell Lysate
- FT – 10 μ L Flow Through
- W – 10 μ L Wash
- E1, E2, E3, E4, E5, E6 – 10 μ L of each Elution sample

41. Prepare each sample by adding 10 μ L of SDS loading dye (1:1) to each of the 10 Eppendorf tubes listed above. (similar to step 22)

- Place each Eppendorf tube at 95 °C for 5 min to denature the proteins.
- Use Dual Xtra ladder (Precision Plus Protein Dual Xtra Standards, by Bio-Rad).

42. Prepare Bio-Rad gel rig to run the gel and fill with 1x SDS buffer. (same as step 23).

43. Load 1 μ L of Ladder and 5 μ L of each prepared sample onto the SDS-PAGE. (same as step 24).

44. Run the gel at 180 V for 40 – 50 minutes, until the dye reaches near the bottom of the gel. (same as step 25).

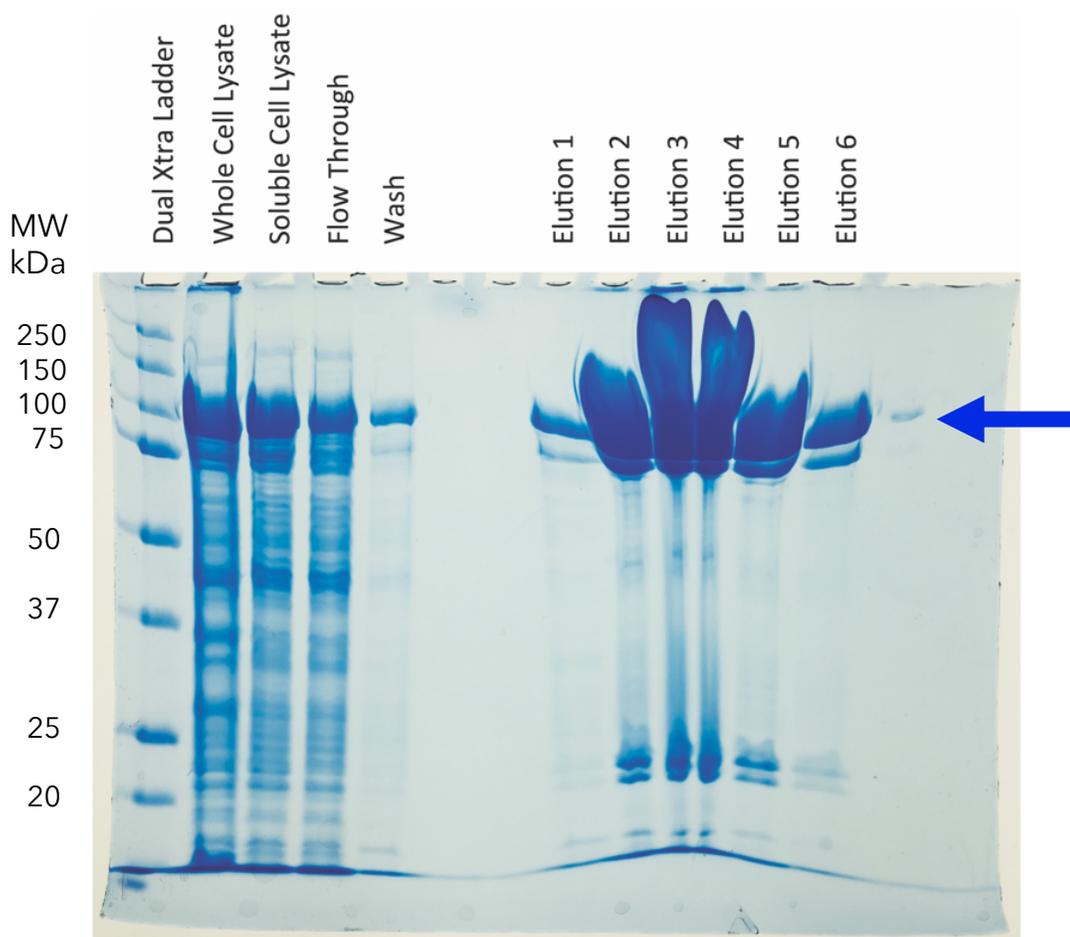
- Run for 40 minutes first, then add more time as desired.

45. After gel run is complete, remove the gel and perform fast stain with Coomassie blue stain. (same as step 27)

- Place the gel flat into an empty micropipette tip box.
- Pour enough Coomassie blue stain to cover the gel in the tip box.
- For Fast Stain, stain for only 5 minutes.
- Destain for 10 – 15 minutes.
- Visualize the gel on white lamp.

46. For a thoroughly stained gel, stain the gel for 1 hour with Coomassie blue dye. (same as in step 28).

- Destain overnight.
- Visualize the gel on white lamp. Take a picture for record (see Figure 2.1, below).



Gel 164, RNA_T7polymerase_elutionTest_OAS,
(37.5:1 A:bA) 10% SDS_PAG,
bioRad 180V, 40min,
total Coomassie Blue staining 1hr, overnight (~22hrs) destaining

File: Gel164_T7p_elutionTest_OAS_11.06.18_IMG_1411_bioRad180V40min_coomassieBlue

Figure 2.1: Sample SDS-PAGE from T7 RNA polymerase expression. See next section (2.12) for analysis of the gel. Molecular weight of T7 RNA polymerase is 99 kDa^{97,98}. Precision Plus Protein Dual Xtra Standards by Bio-Rad is the protein ladder loaded in the very left lane of the gel. The blue arrow is pointing to the overexpressed T7 RNA polymerase protein in different lanes.

2.12 SDS-PAGE analysis of T7 RNA polymerase expression

47. Interpret the gel by analyzing the bands in different wells / lanes of the gel.

- Ladder (well 1) gives details on relative mobility of various proteins across the gel.
- Next 4 wells / lanes serve as controls to monitor the protein expression profile during every step of the experiment.
 - WCL (well 2) includes soluble and insoluble (membrane-bound) proteins, which will show up all along the well product migration and looks fairly “dirty”.
 - A dark band of T7 RNA polymerase protein should be visible.
 - SCL (well 3) may look cleaner since membrane-bound proteins have been removed, but all soluble proteins will still be visible.
 - A dark band of T7 RNA polymerase protein should be visible.
 - FT (well 4) should still contain soluble proteins, except T7 RNA polymerase (which should be bound to the nickel resin), unless insufficient nickel resin was used, or protein expression was particularly good.
 - W (well 5) should have least amount of protein but may still have some T7 RNA polymerase if insufficient nickel resin was used, or if protein expression was particularly good.
- Last 6 wells / lanes show expression profiles of T7 RNA polymerase in each Elution step.
 - Typically, Elutions 2 – 5 show the most target expressed protein.

- Elutions containing largest bands of T7 RNA polymerase should be collected for dialysis to get rid of imidazole, which was present in buffers **A**, **B**, **C**, and **D** during the purification steps.

2.13 Protein Dialysis

48. If performing dialysis to remove imidazole:

- Prepare dialysis buffer (1x T7 Purification Buffer **A**) in a 1 L beaker with a stir bar. (see Table 2.2, below)
- Pre-cool the dialysis buffer in a beaker at 4 °C.
- Pre-wet ~5 inches of dialysis tube in the dialysis buffer.
- Close an empty 1.5 mL Eppendorf tube onto a portion an end portion of dialysis tube. Eppendorf tube serves as a flotation device to prevent the dialysis tube from hitting the spinning stir bar on the bottom of the beaker, potentially damaging the dialysis tube.
- Fold dialysis tube on one side a few times and place a clip on the folded region.
- Add desired Elutions (E2 – E5) into the open side of the dialysis tube.
- Fold the open side of dialysis tube closed a few times and place a clip on the folded region, ensuring that the seals on both sides of the tubing is tight so that the protein does not leak out.
- Place the beaker on the stir plate at 4 °C and set the rate to ~100 rpm.

49. Change dialysis buffer 3 times total. Discard dialysis buffer down the sink. Rapid dialysis buffer changes (less than 3 – 4 hours) are not recommended.

- First dialysis buffer change is recommended at ~4 hours of dialysis. Most imidazole will be removed during this step.
- One dialysis is recommended to run 12 – 16 hours (can be done overnight).

2.14 Determining protein concentration and preparation for storage

50. When dialysis is complete, carefully pipette out the contents of the dialysis tube (containing T7 RNA polymerase) and place into an Eppendorf tube.

- Concentration of T7 RNA polymerase can be measured on NanoDrop 2000 at 280 nm at this time.
- Once the concentration of T7 RNA polymerase has been determined, protein can be mixed with glycerol (50% v/v). **Do NOT vortex the protein!**
 - Add same volume of glycerol to the Eppendorf tube with T7 RNA polymerase.
 - Pipette up and down slowly, many times, until and after the mixture appears homogenous.
- *** Concentration of T7 RNA polymerase at this point is half of the protein concentration measured just prior to mixing it with glycerol (50% v/v).
- It is recommended to separate the newly prepared T7 RNA polymerase into multiple Eppendorf tubes to minimize the freeze-thaw cycles and protect the integrity of the protein.
- Store in -20 °C or in -80 °C.

2.15 Recipes for T7 RNA polymerase purification buffers

Multiple buffers are prepared for lysing bacterial cells and purifying T7 RNA polymerase from the cell lysate (see Table 2.3, below). The main importance is the 5x T7 purification buffer **A**, because all other buffers, including those for dialysis, are made by diluting buffer **A** to 1x concentration and either incrementally increasing imidazole concentration or taking imidazole out of the buffer completely. Below are the recipes to prepare all required buffers for T7 RNA polymerase expression, lysis, purification, and dialysis.

Table 2.2: Reagent recipes for making various buffers for T7 RNA polymerase elution.

5x T7 purification Buffer A		
Reagent	amount	final conc.
NaH ₂ PO ₄ •H ₂ O	30g	250mM
NaCl	146.1g	2.5M
ddiH ₂ O	fill to 900mL	
"titrate" with NaOH to pH 7.50 fill to 1L		

T7 Lysis Buffer B		
Reagent	amount	final conc.
5x [A]	40mL	1x [A]
1M Imidazole stock pH 7.5	4mL	20mM
fill to 200mL adjust to pH 7.50		

T7 Wash Buffer C		
Reagent	amount	final conc.
5x [A]	10mL	1x [A]
1M Imidazole stock pH 7.5	2.5mL	50mM
fill to 50mL adjust to pH 7.50		

T7 Elution Buffer D		
Reagent	amount	final conc.
5x [A]	10mL	1x [A]
1M Imidazole stock pH 7.5	25mL	500mM
fill to 50mL adjust to pH 7.50		

CHAPTER 3: WORKING WITH NUCLEIC ACIDS

For most applications in the lab, the following steps are taken. DNA templates and PCR primers are ordered from Integrated DNA Technologies (IDT). Transcription templates are prepared by setting up PCR reactions and amplifying DNA templates for *in vitro* transcription. Reverse complement RNA strands are produced by T7 RNA polymerase via *in vitro* transcription process in Eppendorf tubes. RNA strands are then purified by running them through a denaturing 8M Urea-PAGE (polyacrylamide gel electrophoresis), bands containing presumed RNA product are cut out, and RNA material is eluted with “crush & soak” buffer. Supernatant is collected and mixed with 100 % Ethanol, after which, Eppendorf tubes are placed in -20 °C freezer for RNA precipitation. RNA strands are collected by further precipitating them on an ultracentrifuge, followed by evaporation any lingering ethanol on SpeedVac apparatus (used Labconco CentriVap micro IR Vacuum Concentrator). RNA strands are reconstituted with Ultra-Pure water and stored at 4 °C, or at -20 °C. Concentrations of RNA strands are determined by measuring the absorbance values collected by NanoDrop 2000 spectrophotometer and calculated using Beer-Lambert law.

3.1 RNA strand design

As in any project that employs nucleic acids (NAs), experiments are designed first, which includes oligonucleotides to be used. An already used sequence of an RNA fiber monomer strand 1 was taken as the basis to develop a simple hairpin to test various sequence mutations of the KL binding region (see Figure 3.1, below, panel A). The monomer, once it folds, has two KL regions (outlined in red and blue nucleotides in

panels A and B), and a GFP antisense sequence starting with an orange U (uracil). The sequence within and including the purple G (guanine) and C (cytosine) is removed for the hairpin sequence (shown in Figure 3.1, below, panel B). The GFP antisense (as) sequence region (extensively used in our previous works^{79, 99-105}) can be used as a toehold (sticky end) to attach oligonucleotide fragment to larger nanoconstructs or to bind a complementary (RNA) strand for siRNA delivery into cells. A displacement strand was designed to protect the GFPAs toehold during testing the hairpin structures due to self-reactivity of the sequence, where nucleotides (shown in Figure 3.1, below, in bold, for GFP sequence regions) interact between each other. Similar situation is observed (in NUPACK¹⁰⁶ simulations) for the strand sequences shown in panels A, B, and C/D. This should not be a problem as long those GFPAs sequence regions are not displaced and are protected by the GFP displacement strands.

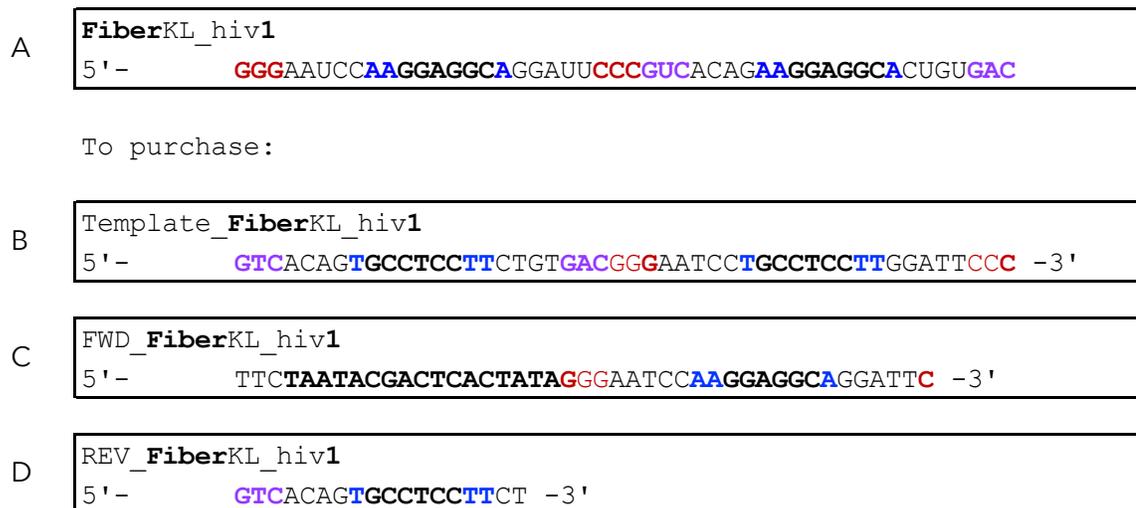


Figure 3.1: Sequence samples for RNA strand design. In panel **A**, original RNA fiber monomer strand 1, featuring KL1 marked in bold letters. Panel **B** shows the DNA template strand that would be ordered from IDT for synthesis. Notice that it is complementary to the sequence in Panel **A**, where underlined, bold, and color-coded

nucleotides are complementary. Panel **C** shows the Forward (FWD) primer sequence of the FiberKL_hiv1 strand (in panel A), which is needed for synthesis of the double-stranded (ds) DNA template. Region of the Forward sequence at the 5'-end, shown in bold, is the T7 RNA polymerase promoter region, which promotes the transcription of the strand by T7 RNA polymerase. Panel **D** shows the sequence of the Reverse (REV) primer needed for the synthesis of the dsDNA template that is used to synthesize RNA during *in vitro* transcription.

3.2 Preparing DNA template and primers for PCR

When RNA sequences are designed for a particular application, a DNA reverse complement of the target sequence is determined and appropriate FWD and REV primers are also determined. An order is placed to a company specializing in producing synthetic oligonucleotides, such as Integrated DNA Technologies (IDT). Upon arrival of the ordered strands they need to be reconstituted in Ultra-Pure water as synthetic oligonucleotides arrive dry in the pellet form. The process of reconstituting pelleted oligonucleotides is outlined below.

1. Centrifuge microcentrifuge tubes with dry DNA strands for 5 seconds before opening.
2. Add 10 times the amount of Ultra-Pure water (in μL) than the amount of DNA strand (in nmol).
 - a. * amount of DNA strand is usually listed on the microcentrifuge tube label.
 - b. ** ex. for 72 nmoles, add 720 μL of Ultra-Pure water to get $\sim 100 \mu\text{M}$ strand.
3. Vortex the microcentrifuge tube and centrifuge briefly.
4. Prepare $0.2 \mu\text{M}$ DNA template by adding 1 μL of $100 \mu\text{M}$ template into 500 μL of Ultra-Pure water. Label 1.5 mL Eppendorf tube(s) accordingly.

- a. *** for PCR use, DNA template is at 0.2 μM , while FWD and REV primers are kept at 100 μM .
5. Keep on ice (4 $^{\circ}\text{C}$) if using supplies immediately, otherwise microcentrifuge tubes can be stored in the freezer (-20 $^{\circ}\text{C}$).

3.3 DNA amplification via PCR

As outline in the section above, DNA templates needed for *in vitro* transcription need to be double-stranded and need to have a T7 RNA polymerase promoter region to promote the transcription of the desired strands. Forward primer is the one that has the T7 promoter region incorporated in it. When DNA template, FWD, and REV primers are added to the PCR mixture (see Table 3.1, below, for the recipe to set up PCR), the mixture is brought to 95 $^{\circ}\text{C}$ to melt any secondary structure that oligonucleotides may assume, then the temperature slightly lowers to 94 $^{\circ}\text{C}$, followed by a drop of temperature to 55 $^{\circ}\text{C}$ of the annealing step (see Table 3.1, below, for the PCR protocol). During annealing step, primers bind the DNA template. The temperature then is raised to 72 $^{\circ}\text{C}$ for elongation step, where optimal conditions are set for Taq polymerase to synthesize DNA. Temperature is then raised back to 94 $^{\circ}\text{C}$ for denature the double stranded DNA, where the newly synthesized DNA strand is no longer bound to the DNA template it was synthesized on. This cycle repeats for 30 cycles, where the number of DNA strands is doubled every cycle of annealing, elongation, and denaturation. After 30 cycles of amplifying the DNA the thermocycler lowers the temperature to 4 $^{\circ}\text{C}$ for storage. At this point the amplified DNA has the appropriate T7 RNA polymerase promoter on the template strands for the transcription into RNA strands.

Table 3.1: Reagent recipe for setting up PCR reaction (left), and protocol for setting up Thermocycler program for PCR (right).

PCR Recipe		Thermocycler Protocol		
amount	reagent	# cycles	temperature	step duration
25 μ L	2X MyTaq Mix (from Biorline)	1	95 $^{\circ}$ C	2 min
22 μ L	Ultra-Pure H ₂ O (17.5-17.8 M Ω ·cm)	30	94 $^{\circ}$ C	1.5 min
1 μ L	0.2 μ M DNA template		55 $^{\circ}$ C	1.5 min
1 μ L	FWD primer		72 $^{\circ}$ C	1.5 min
1 μ L	REV primer		4 $^{\circ}$ C	infinitely

1. Thaw all reagents and then place on ice (4 $^{\circ}$ C).
2. Vortex and briefly centrifuge 0.2 μ M DNA template(s), FWD and REV primers before use, but **do NOT vortex** MyTaq™ Mix (from Biorline).
3. Label PCR tube(s) accordingly. Writing on top and side of PCR tubes is advised.
4. Add 25 μ L of MyTaq Mix to each PCR tube.
5. Add 22 μ L of Ultra-Pure water to each PCR tube.
6. Add 1 μ L of 0.2 μ M DNA template, 1 μ L of FWD primer, and 1 μ L of REV primer to each respective PCR tube.
7. Pipette up and down to mix the contents of each PCR tube. (**do NOT vortex**, enzymes present!)
8. Place PCR tubes into the thermocycler, choose appropriate protocol, and start the protocol run. The protocol takes about 3.5 hours. Samples can be held in the thermocycler at 4 $^{\circ}$ C or can be placed at 4 $^{\circ}$ C elsewhere until DNA purification.

3.4 DNA purification after PCR

After DNA has been amplified via PCR, it needs to be purified from excess dNTPs, enzyme, and buffers. This can be done using DNA Clean & Concentrator™-5 Kit by Zymo Research. The kit consists of a purification column, DNA binding buffer, and a DNA wash buffer. DNA product from PCR is mixed with DNA binding buffer and then passed through the purification column, where DNA is bound to the filter on the column. This mixture can be passed through the column twice and the filtrate can be discarded. The column is then washed with the DNA wash buffer to ensure that the filter of the column is clean of the unwanted PCR reagents. Finally, DNA is released from the purification column filter into a new Eppendorf tube by adding Ultra-Pure water. This step can also be performed twice but keep the filtrate as it now contains the purified DNA (pDNA) that can be used for *in vitro* transcription. This process is outline in the steps listed below.

1. Label filtration column, collection column (or a 2.0 mL Eppendorf without the lid), and two 1.5 mL Eppendorf tubes (with “pDNA” for purified DNA, sample name, etc.) for each DNA sample to be purified.
2. Transfer (~50 µL) PCR-amplified DNA (for each sample) into corresponding 1.5 mL Eppendorf tube(s).
3. Add 650 µL of DNA binding buffer to each Eppendorf tube with PCR-amplified DNA.
4. Vortex and centrifuge each Eppendorf tube briefly.

5. Place each labelled filtration column into respective collection column, and transfer each (~700 μL) liquid from Eppendorf tubes onto respective filtration column. Do not touch the filter with micropipette tip.
 - a. Alternative to steps 2 – 4, place filtration column into a 2.0 mL Eppendorf tube, pipette 650 μL of DNA binding buffer into the filtration column, then pipette ~50 μL of PCR-amplified DNA into the DNA binding buffer and pipette up and down multiple times to well mix the contents.
 - b. Also, if 2x PCR (~100 μL) was prepared for the same sample, filtration column can accommodate that volume also. May need to reduce DNA binding buffer to 600 μL (+ 100 μL of DNA product), due to filtration column volume capacity.
6. Place each filtration/collection column in a tabletop centrifuge and spin at 10,000 rpm for 30 seconds, carefully, as there are no lids and contents may spill over.
7. Empty each filtrate liquid from collection column into liquid waste container. DNA should be on the filter during this step.
 - a. Alternatively, filtrate can be passed through the filtration column again, steps 5 – 6.
8. Add 200 μL of DNA wash buffer to each filtration/collection column.
9. Centrifuge columns at 10,000 rpm for 30 seconds.
10. Repeat steps 7 – 9.
11. Remove filtration columns, throw away collection columns, and place filtration columns into corresponding newly labelled “pDNA” 1.5 mL Eppendorf tubes.

12. Add 50 μL of Ultra-Pure water onto filter of each filtration column. Do NOT touch the filter with micropipette tip.
 - a. If $\sim 100 \mu\text{L}$ PCR product was used, then reconstitute with 100 μL of Ultra-Pure water.
13. Centrifuge each filtration column / Eppendorf tube at 10,000 rpm for 30 sec. Turn each Eppendorf tube cap towards the center of the rotor to reduce the chances of caps being torn off by centrifugal force (tends away from rotor center; opposite of centripetal force, which tends toward the rotor center).
 - a. To try increase filtered pDNA product, reconstituted pDNA product can be passed through the filtration column again, steps 12 – 13.
14. Dispose of filtration column(s), close “pDNA” labelled 1.5 mL Eppendorf tube(s), and place on ice (4 $^{\circ}\text{C}$) for immediate use or store in the freezer (-20 $^{\circ}\text{C}$).

3.5 PCR-product (p)DNA verification via agarose gel

After PCR reaction is complete, it is good to verify that it was successful.

Transcribing RNA *in vitro* is a time-consuming process and setting it up with questionable material will only disappoint. It is possible that one was distracted and did not add all reagents into the PCR mix, or that the Taq polymerase was left out of freezer and degraded. Verifying pDNA presence after PCR is a quick enough analysis and is recommended. Recipe for making an agarose gel is listed below in Table 3.2 and the steps to perform the analysis are outline below the table.

Table 3.2: Reagent recipe for making a 1.5 % Agarose gel.

1.5% Agarose Gel			
1X TBE	mL	50	100
Agarose	g	0.75	2
** microwave to boil → bring to RT			
Ethidium Bromide	μL	1.25	1.45
Total Volume	mL	50	100

1. Prepare a 1.5% Agarose gel (see Table 3.2, above).
 - a. Set up a clean gel casting stand with an appropriate well comb.
 - b. Weigh 0.75 g of agarose in a weigh boat and place agarose into a 250 mL Erlenmeyer flask.
 - c. Measure 50 mL of 1x TBE in a graduated cylinder and pour into the Erlenmeyer flask. (A portion of 1x TBE can be poured down the weigh boat (serving as a funnel) and into the Erlenmeyer flask to ensure that all agarose was used. (see Table A.2 in Appendix A.2 for 10x TBE buffer preparation).
 - d. Microwave the mixture for ~45 seconds to boil the mixture and make sure that the solution turns clear. Do NOT overheat the gel mixture as it will splash out of the flask and much cleaning will be needed in the microwave.
 - e. Handle hot Erlenmeyer flask with heat-protective gloves.
 - f. Let the solution cool down on the benchtop until the heat is bearable to touch by bare hand. (Do NOT get burned.)
 - g. Add 2.5 μL of ethidium bromide. Swirl the Erlenmeyer flask to ensure even distribution of ethidium bromide throughout the gel mixture.

- h. Pout the still warm and liquid gel mixture into the gel stand / mold. Ensure that the gel is bubble-free by poking the air bubbles with a 200 μL pipette tip, or by gently moving the air bubble toward the surface or to the side until they disappear.
- i. Once the gel is polymerized, remove the well comb and position the gel in the gel rig appropriate for running the gel.

2. Agarose gel analysis

- a. Once the agarose gel polymerizes (it will become opaque), position the gel in the gel rig appropriate for running and pour 1x TBE buffer to completely fill the agarose gel in the gel rig.
 - If the gel rig is adjustable, make sure that the gel rig is level with the benchtop.
- b. Prepare each pDNA (or other NAs) by mixing (1:3) with agarose loading buffer (bromophenol blue, xylene cyanol, and glycerol)
 - On a piece of parafilm, pipette 3 μL of agarose loading buffer and into it pipette 1 μL of pDNA sample. Pipette up and down a few times to mix the sample(s).
 - If multiple samples are prepared, label each sample drop on the parafilm with a sharpie and keep track where on the gel each sample gets loaded.
- c. Load pDNA samples on the agarose gel. Close the lid and attach positive and negative electrodes into their respective receptacles.

- If gel wells are on the side of the gel and electrode polarity is reversed, pDNA samples will exit the gel and there will be no results to collect.
- d. Set the power supply on 120 V for 7 minutes.
- Make sure to set the timer, since some power supplies do not automatically turn off, which will result in pDNA samples also leaving the gel.
- e. Once the gel run is complete, turn the power supply off and detach the electrodes to avoid the possibility of getting shocked.
- f. Place the agarose gel onto a paper towel and carry to the UV transilluminator to visualize. Take a picture for records.

3.6 *In vitro* RNA synthesis using T7 RNA polymerase (“Transcription”)

In all living organisms, cells that contain DNA transcribe it to RNA, or at least certain genes from DNA, depending on the type of cells and their stage in molecular life. Viruses, although arguably considered alive, cannot replicate on their own, but rather hijack cellular machinery to complete their replication cycle. Bacteriophages have a specific T7 RNA polymerase which transcribes RNA well enough to catch interest. This enzyme is used to perform *in vitro* transcriptions (in an Eppendorf tube) at 37 °C. Table 3.3, below, shows the recipe of what reagents are needed for initiating *in vitro* transcription. To prepare the reagents needed for *in vitro* transcription use Table 3.4, below. The process for *in vitro* transcription is outlined below the Tables.

Table 3.3: Reagent recipe for setting up *in vitro* transcription reaction.

Transcription Master Mix		
reagent:	concentration:	amount:
Ultra-Pure water		10 μ L
Transcription Buffer	5x	10 μ L
rNTPs	25 mM each	10 μ L
DTT (dithiothreitol)	100 mM	5 μ L
T7 RNA polymerase	3 μ M	5 μ L
pDNA template		10 μ L
Total amount		50 μ L

Table 3.4: Recipes for preparing reagents for *in vitro* transcription reaction.

5x TB (Transcription Buffer)				
	MW (g/mol)	amount (g)	C2 (mM)	V2 (mL)
HEPES	238.3	0.9532	400	10
Spermidine	145.25	0.0145	10	
DTT	154.25	0.3085	200	
MgCl ₂	203.3	0.2440	120	

pH adjust to 7.5 with 5 M KOH or 5 M HCl

DTT (dithiothreitol)				
	MW (g/mol)	amount (g)	C2 (mM)	V2 (mL)
DTT	154.25	0.15425	100	10

pH adjust to 7.5 with 5 M KOH or 5 M HCl

rNTPs (ribonucleotide triphosphates)				
	MW (g/mol)	amount (g)	C2 (mM)	V2 (mL)
ATP	551.14	0.13779	25	10
CTP	527.12	0.13178	25	
GTP	523.18	0.13080	25	
UTP	586.12	0.14653	25	

3.6.1 Transcription Initiation

1. Turn on water bath (or heat block) to 37 °C.
2. Thaw all reagents and then place on ice (4°C). Leave T7 RNA polymerase in the enzyme box until ready to use.
3. Label all 1.5 mL Eppendorf tubes with their respective sample names.
4. Add Ultra-Pure water, transcription buffer, rNTPs, and DTT to each Eppendorf tube.
Vortex and briefly centrifuge, while T7 RNA polymerase is not added to the mix.
 - a. Alternatively, prepare a Transcription Master Mix.
 - b. # of samples (50 µL volumes) + 1 volume = # transcription volumes
 - c. Add each reagent (see table/figure above) to an Eppendorf tube in multiples of transcription volumes per reagent. Vortex and briefly centrifuge, while T7 RNA polymerase is not added to the mix, otherwise only pipette up and down.
5. Add T7 RNA polymerase to the mix and pipette up and down **only**.
 - a. Alternatively, if Transcription Master Mix was prepared, add # of transcription volumes amount of T7 RNA polymerase to the mix and pipette up and down **only**.
 - b. Aliquot # of transcription volumes per Eppendorf tube sample.
6. Add pDNA template to the transcription mix and pipette up and down **only**.
 - If multiple transcription volumes per Eppendorf tube were added, add same # of pDNA volumes per tube.

7. Place Eppendorf tubes on floater and into the 37 °C water bath or onto a 37 °C heatblock that accommodates that size Eppendorf tubes for 3.5 hours.

- Alternatively, transcription reactions can be left at 37 °C overnight. Good results were reported from teammates at 16 hours of transcription reactions.**

3.6.2 Transcription Termination

1. Add 3 µL of RQ1 DNase (RNase-free) per each transcription volume per Eppendorf tube, then pipette up and down.
2. Incubate transcription samples at 37 °C for another 30 minutes.

3.6.3 Transcription Purification via urea-PAGE

After *in vitro* transcription has been terminated, transcribed RNA strands need to be separated from the mixture, which has excess ribonucleotides, degraded DNA strands, T7 RNA polymerase, RQ1 DNase, and other reagents needed to provide the appropriate environment for successful transcription reaction. For RNA strands to be clean from all these contaminants (at this point), they need to be separated from the rest of the components. This can be done via electrophoretic mobility on a denaturing 8M urea-PAG (polyacrylamide gel). All components can be separated by size, at which point the RNA product can be cut out from the gel and eluted from it. Table 3.5 immediately below shows the recipe for reagents needed to make an 8M urea-PAG. Steps outlining the process of performing transcription purification are listed below the table.

Table 3.5: Reagent recipe for making an 8M urea-PAG (polyacrylamide gel).

8M Urea PAG(E)					
10X TBE	mL		3.5	5	10
20% AA / 8M urea	mL		14	20	40
8M urea	mL		17.5	25	50
*** mix well ***					
10% APS	μL		350	500	1000
TEMED	μL		17.5	25	50
Total Volume	mL		35	50	100

1. Prepare 8M urea-PAG (polyacrylamide gel) using the recipe shown in Table 3.5.
 - a. Gather appropriate supplies and equipment to cast the gel: two glass plates, rubber gasket, two gel spacers (depth of spacers must match the depth of the well comb to be used), well comb, four gel clamps, and a beaker to accommodate the appropriate volume of gel solution.
 - b. Clean the glass plates and put together the desired gel plate assembly.
 - c. Prepare the appropriate amount of urea-PAG solution.
 - d. Pour gel solution into the gel assembly while holding the gel assembly in ~60° tilted position as soon as all the reagents have been well mixed and place the desired gel well comb into the gel assembly. Lay the gel assembly flat onto an empty micropipette tip box to polymerize.

2. RNA purification via Urea-PAGE
 - a. Set up gel rig by removing the clamps and gaskets from the glass plate assembly.
 - b. Secure the medium or large C.B.S. Scientific gel assembly to the gel stand with clamps.
 - c. Fill top and bottom chambers with 1x TBE buffer.

- d. Rinse all gel wells with a 60 mL syringe and a (blunt) needle.
 - e. Add 8M urea loading buffer (1:1) to each sample and mix by pipetting up and down.
 - f. Load wells with transcribed samples, pipetting carefully not to cross-contaminate the wells.
 - g. Cover buffer chambers with designated covers and plug gel rig into the power source, matching positive and negative electrodes with their appropriate receptacles.
 - If electrodes are mismatched, RNA (or DNA) material will come out of the wells into the top chamber and mix. This mishap will be irreversible.
 - h. Set the power supply on 25 – 50 watts and run the gel until the dyes (bromophenol blue and xylene cyanol) have separated 4 – 5 inches.
3. Cut the RNA samples out from the gel
- a. Once urea-PAGE is complete, turn the power supply off, unplug the electrodes from the power supply, and remove buffer chamber covers.
 - b. Take the whole gel stand to the sink and pour out the running buffer, as it will spill if the gel is removed from the stand before emptying the buffer chambers.
 - c. Open the gel assembly and remove the gel. Place the gel onto saran wrap, covering front and back of the gel.

- Be sure to orient the gel in a way that samples are easily identified upon gel assembly breakdown. (bottom left or bottom right corner of the gel can be cut off to designate gel directionality.
- d. Place the saran-wrapped urea gel onto the TLC plate.
 - e. Visualize the gel in the dark with a handheld UV lamp. Write the details of the gel directly on the saran wrap using a thin permanent marker. Sample names, date, initials. A picture can be taken for records.
 - f. Trace around the presumed RNA product with a thin permanent marker. Another image can be taken for records.
 - g. Appropriately label 2.0 mL Eppendorf tubes for each sample on the gel.
 - h. Place the labelled / marked gel on a plastic board and carefully cut out each sample band with a scalpel or a precision blade. Be careful with sharp blades!
 - i. Place each cut out gel band in its respective Eppendorf tube and add 500 μ L of Crush and Soak buffer (see Table 3.6, below). *** Make sure that the gel bands are covered in Crush and Soak buffer.
 - j. Place the Eppendorf tubes in the cold room (4 $^{\circ}$ C) overnight.

Table 3.6: Recipe for making Crush and Soak buffer. See Table A.2 in Appendix A.2 for recipe to make needed reagents.

Crush and Soak buffer (300 mM Na ⁺)				
	C1	amount	C2	V2 (mL)
5 M NaCl	5M	12 mL	300 mM	200
10x TBE	10x	20 mL	1x	
ddiH ₂ O				

3.7 RNA precipitation and Recovery (Transcription Purification)

RNA product from cut out gel bands diffuses into the Crush and Soak buffer. It is important to understand that if gel bands are too big or occupy a larger ratio of the total (Crush and Soak + gel bands) volume, then proportionally to the total volume less RNA product will have diffused into the buffer. If the ratio of gel bands to the total volume of buffer is smaller, then the diffusion of RNA product is more efficient, experimentally. To further separate RNA product from the buffer, ethanol is added to help it precipitate while in cold conditions and further under centrifugal force. Steps outlining this process are listed below.

1. Pre-cool 100% and 90% ethanol in -20 °C.
2. Transfer liquid from each sample (RNA sample in Crush and Soak buffer) from 2.0 mL to newly labelled 1.5 mL Eppendorf tube.
3. Add 750 µL of cold 100% ethanol to each liquid sample in 1.5 mL Eppendorf tubes. Vortex briefly. If sample is not vortexed, liquids will not mix and will freeze in -20 °C freezer.
4. Place in the freezer (-20 °C) for 1 – 3 hours.
 - a. Alternatively, place vortexed samples on dry ice (~ -78.5 °C) or in -80 °C for ~5 minutes. Sample liquid will get very viscous (can be seen by tilting the Eppendorf tube). Be sure not to freeze samples!
 - b. Combination of -20 °C and -80 °C can also be used. Keep samples in in -20 °C for ~1 hour and -80 °C for ~ 5minutes. **Recommended.**
5. Pre-cool the large centrifuge, by turning on and setting to 4 °C.

6. Transfer Eppendorf tubes into the cold centrifuge (set to 4 °C) and spin at 14,000 rpm for 30 minutes. Orient all Eppendorf tubes on the centrifuge rotor the same way (rotationally, writing/label on cap right side up) to ensure that pellets form on the same side of the tube in all the tubes.
 - **Make sure to keep the centrifuge rotor balanced!**
7. After 30-minute spin, remove supernatant from each Eppendorf tube (leaving behind 100 – 150 µL) with a P1000 micropipette being careful not to disturb the pellet on the bottom of the tube.
8. Wash RNA product in each Eppendorf tube by carefully adding 900 µL of cold 90% ethanol to each Eppendorf tube.
9. Spin the Eppendorf tubes in the cold centrifuge (set to 4 °C) at 14,000 rpm for 10 minutes.
10. Repeat steps 8 – 9 again.
11. After two washing RNA products with 90% twice, remove supernatant, leaving behind 50 – 75 µL (or can be less, as long as the pellet is not disturbed).
12. Pre-heat Speed-Vac to 55 °C.
13. Place Eppendorf tubes with samples into the Speed-Vac with opened caps.
 - Make sure to keep the Speed-Vac rotor balanced!
14. When samples are dried, remove from Speed-Vac.
 - Samples can be kept dry in the cold room (4 °C) for some time.
15. Reconstitute each RNA sample in 30 µL Ultra-Pure water. Vortex well and briefly centrifuge.

16. Place RNA samples temporarily on ice (4 °C) or store in the freezer (-20 °C).

3.8 Measuring nucleic acid concentrations on NanoDrop 2000 spectrophotometer

To use nucleic acid samples for various applications, determine their concentration. Measure their absorbance on NanoDrop 2000 spectrophotometer. Get respective extinction coefficients of each of the oligonucleotides. Use the Beer-Lambert law ¹⁰⁷ ($A = \epsilon c l$) to calculate the concentration of each sample, where **A** is the absorbance value at 260 nm, ϵ is the extinction coefficient ($L/(\text{mol} \cdot \text{cm})$), **c** is the concentration of the sample (M), and **l** is the pathlength (cm). Process of measuring the nucleic acid absorbance is outlined below.

1. Turn on and login onto the computer that has NanoDrop 2000 software interface.
Open the software interface. Make sure that NanoDrop 2000 spectrophotometer is on and responding.
2. Select the option to measure absorbance of “Nucleic Acid”. Allow the instrument to perform a routine wavelength verification. Make sure that the sampling arm is lowered. Select the type of nucleic acid to be sampled (RNA or DNA).
3. Make sure that the stage or pedestal is clean by wiping it with Ultra-Pure water and a Kim Wipe.
4. Perform a “Blank” by pipetting 1.0 – 1.5 μL of Ultra-Pure water (or can be the solution in which a sample is dissolved) onto the lower pedestal and carefully lower the sampling arm. Click “Blank” in the upper left corner of the software interface. A blank measurement will be performed to set the reference absorbance of the solution of the sample.

5. Wipe of the top (on the sampling arm) and bottom pedestals with a Kim Wipe. Make sure to keep the stage / pedestals clean after every sample measurement.
6. Measure the absorbance of the sample by pipetting 1.0 – 1.5 μL of sample onto the lower pedestal. Carefully lower the sampling arm, input the sample name, and click “Measure” in the software interface.
7. Calculate the concentration of the sample. Collect the absorbance values from the software interface at 260 nm, as nucleic acids absorb light at that wavelength, and the extinction coefficient of the specific oligonucleotide. Use Beer-Lambert’s law to determine the concentration of oligonucleotides.

3.9 General Assembly Protocol

An experiment is typically designed, and layout prepared in Microsoft’s Excel spreadsheet. RNA strands are thawed and/or placed on ice (4 °C) from their storage. RNA strands are mixed together according to the prepared spreadsheet in appropriate stoichiometric ratio in water. Eppendorf tube with mixed RNA strands is placed on a heatblock at 95 °C for 2 minutes. The purpose of heating up the mixture is to break all intra- and intermolecular hydrogen bonds of the nucleic acid strands. After that, the tube is placed immediately on ice for 2 minutes, a step called “snap-cooling”. The purpose of this step is to induce individual strands to “self-hybridize” and assume their intended secondary (2°) and/or tertiary (3°) structure. After 2 minutes on ice, 5X AB (assembly buffer) is added to the Eppendorf tube in the amount of 20 % of final volume to achieve a final 1X concentration of AB. The purpose of this step is to introduce metal ions, Mg^{2+} in particular, to stabilize RNA molecules in areas of interactions, such as in

kissing loops and complimentary sticky end interactions. Depending on nanoparticles designed, final mixtures may need to be incubated at 30 °C, 45 °C or another temperature, to further allow the structures to assume their final intended conformations and be stabilized, after which the mixture can be stored at 4 °C or be frozen at -20 °C or even -80 °C for shipping. Typical concentration of assembled product is 1 μM. Volume of assembled product is more variable due to applications but can be between 5 and 50 μL. Salt concentrations of different buffers and of the assembly are shown in the Table 3.7 below.

Table 3.7: Assembly Reagents and their concentrations. See Table A.2 in Appendix A.2 for preparation of needed reagents.

native Assembly, Loading, and Running Buffers				
	5x AB	Assembly	1x LB	1x RB
Tris-Borate	(89 mM) * 5	89 mM	89 mM	89 mM
Mg ²⁺	(2 mM) * 5	2 mM	2 mM	2 mM
K ⁺	(50 mM) * 5	50 mM	50 mM	-
Glycerol	-	-	30% v/v	-

3.10 Assembled structure verification via native-PAGE

After assembling various nanoconstructs (that have secondary and/or tertiary structure) out of nucleic acids, the assembly of those structures needs to be confirmed. Assembled samples can often be seen via atomic force microscopy (AFM), which takes time to ship to a laboratory that can perform the analysis and can be costly. In a more widely accepted, practical and cheaper approach, assembled NA structures can be visualized via native-PAGE (polyacrylamide gel electrophoresis), where native conditions

(especially salt conditions) of the assembly are preserved throughout the process. See Table xx above to trace the salts and metal ion concentrations starting at the assembly buffer (5x AB), to the assembly, next to the loading buffer (1x LB), to the running buffer (1x RB). Recipe for making an 8% native-PAG can be seen in Table 3.8 below.

Table 3.8: Reagent recipe for making an 8 % native-PAG (polyacrylamide gel).

8% native-PAG(E), 2mM MgCl ₂					
ddiH ₂ O	mL	7	10.5	24.5	35
40% AA**	mL	2	3	7	10
10X TB	mL	1	1.5	3.5	5
1M MgCl ₂	μL	20	30	70	100
*** mix well ***					
10% APS	μL	100	150	350	500
TEMED	μL	6	9	21	30
Total Volume	mL	10	15	35	50

After nanoparticles are assembled, their assembly integrity is verified by running the samples on gel electrophoresis apparatus. Native-PAGE (8% acrylamide : bisacrylamide) is often used for this application, with 2 mM divalent (Mg²⁺) magnesium ions. In native-PAGE analysis, native conditions for assembled nanoparticles is preserved to retain their assumed conformation(s). Sample of assembled nanoparticles is mixed (often 1:1 ratio) with 1x LB. The purpose of adding LB to the sample is for the sample to better sink into the well when it is loaded onto the gel.

When the gel apparatus is assembled, it is pre-run for 5 minutes at lower power to ensure a more even salt/ion distribution throughout the gel. The sample is then loaded into the well of a gel and apparatus turned on. To produce a crisper gel band

upon visualization, slow gel run start can be performed to allow the sample to “compress” in the well before turning up the power to the intended setting.

Alternatively, to produce sharper gel bands, a lower amount but higher concentration sample can be used. A typical native gel is run at 300 V, 150 mA and an appropriate time at 4 °C (in the cold room), depending on the sample molecular weight. The rate of sample band migration through the gel can be observed with the help of loading buffer dyes. These dyes can interfere with visualization of the gel, in which case 1x LB can be prepared dye-free.

Once the gel run is complete, it is removed from the glass gel sandwich and stained with either ethidium bromide (EtBr) or SYBR green dyes. Typically, a native gel is stained with ethidium bromide by placing the gel in a glass basin and pouring dilute EtBr solution over the gel to totally submerge the gel for about a minute. After staining, pour EtBr into appropriate EtBr disposal container (do NOT pour down the drain). Rinse the native gel with ddiH₂O and also dispose of this rinse water into the EtBr waste disposal. These intercalating dyes help expose the sample bands when visualizing the gel on the UV transilluminator, such as a ChemiDoc. Be careful with ethidium bromide not to breathe its fumes or get on skin as it intercalates DNA and is a mutagen.¹⁰⁸

CHAPTER 4: KISSING LOOP PROJECT DEVELOPMENT

4.1 Active KL region sequence selection

To expand the library of KL sets, HIV1 and HIV2 strands from RNA fibers⁷⁷ was taken as the basis. There are 9 nucleotides on each KL that participate in a KL interaction forming a complex, where there are 6 nucleotides on each KL that interact via hydrogen bonding and two and one flanking adenines support by stacking interactions. These 6 active nucleotides were mutated to expand the KL complex sets. Since each of the 6 nucleotide (nt) positions in the sequence can be one of 4 bases, there are 4,096 ($4^6 = 4*4*4*4*4*4$) different ways they can be arranged. That does not mean that all 4,096 possibilities are unique, since some of those are palindromic. For complementary sequence selection, the amount of combinations is even less, since for every 6nt sequence there exists its reverse complement in the same “pool” of possible 6nt sequences. It is important to remember that the 6nt sequences are directional in terms of 5' to 3' NA sequence directionality, which expands the combination possibilities.

There are websites that can generate sequences with given criteria, such as using 4 different letters (A, C, U, and G, for nucleotides) for each position (or placeholder) and specifying the number of positions to mutate. Alternatively, this pool of sequence combinations can be produced manually, provided that no sequence duplicates were made in the process. This generated 6nt sequence pool can then be used to select potential 6nt sequences from it, given that certain selection restrictions can be applied. The restrictions applied on 6nt sequence selection were:

- sequences should be GC-rich, where 5 of 6 nucleotides are either G or C,

- keep G or C nucleotides on the “edges” of the 6nt sequence (1st or 6th position), due to G ≡ C base pair having three hydrogen bonds, compared to A = U having only two,
 - (only held true for half of the selected sequences)
- there should be no more than 2 consecutive identical bases in a 6nt sequence,
 - (only held true for half of the selected sequences)
- each mutated 6nt sequence should differ from another potential 6nt sequence by at least 3 consecutive base-pair forming nucleotides.
 - That is, if a potential 6nt sequence when analyzed to see if it can complementary bind to the target sequence (the sequence checked against), that it does not potentially form at least 3 consecutive base pairs.

The KL sequences, HIV1 (5'-CAAGGAGGCAG) and HIV2 (5'-CAAGCCUCCAG), used in RNA fibers⁷⁷ were taken as the basis, where the 6 hydrogen bonding base positions (shown in black and bold) were mutated. All 6nt sequences were selected and checked for inter-compatibility manually. Total of 32 6nt sequences were selected, for a total of 16 sets (see Figure 4.1, below), and one control strand with all adenine nucleotides on the 6nt KL region (5'-CAA**AAAAAA**AG). Selected 6nt sequences (Table 4.1, below) were then inserted into modified KL_1_GFPs (sense) sequence, derived from RNA fibers (see Figure 4.1, on p65). DNA templates and their PCR primers were ordered from IDT. RNA strands were produced via *in vitro* transcription.

Table 4.1: 16 sets of mutated KL sequence regions selected for testing. Name of each KL was shortened from KL1_GFPs to **1**, respectively, since this figure only shows the mutated sequence regions of the KL hairpin strands. See Table A.1 in Appendix A.1 for full hairpin KL sequences.

5'- CAAGGAGGCAG	1 + 2	5'- CAAGCCUCCAG	5'- CAAUGCGGGAG	9 + 10	5'- CAACCCGCAAG
5'- CAACGGAGGAG	1' + 2'	5'- CAACCUCCGAG	5'- CAAGGGCGUAG	9' + 10'	5'- CAAACGCCCAAG
5'- CAAGCGGACAG	3 + 4	5'- CAAGUCCGCAG	5'- CAAUCGGGGAG	11 + 12	5'- CAACCCGAAAG
5'- CAACAGGCGAG	3' + 4'	5'- CAACGCCUGAG	5'- CAAGGGGCUAG	11' + 12'	5'- CAAAGCCCCAAG
5'- CAAGCGCUGAG	5 + 6	5'- CAACAGCGCAG	5'- CAAUGCCCCAG	13 + 14	5'- CAAGGGGCAAG
5'- CAAGUCGCGAG	5' + 6'	5'- CAACGGGACAG	5'- CAACCCCGUAG	13' + 14'	5'- CAAACGGGGAG
5'- CAAGCCAGGAG	7 + 8	5'- CAACCUGGCAG	5'- CAAUCGCCAG	15 + 16	5'- CAAGGGGAAAG
5'- CAAGGACCGAG	7' + 8'	5'- CAACGGUCCAG	5'- CAACCCGCUAG	15' + 16'	5'- CAAAGCGGGAG

4.2 KL nomenclature

The basis for KL project were FiberKL_hiv1 and FiberKL_hiv2 strands from RNA fibers, which contain KL1 and KL2, yielding one KL set. For KLs of the same number (i.e. KL1 and KL1', ...) interior 6nt sequences were reversed in 5' → 3' directionality; the number of each KL was kept, but to denote the difference between the two alternate sequences a hyphen ("prime") was added to each 6nt-sequence-reversed KL. For example, where KL1 has the sequence 5'-CAAGGAGGCAAG, KL1' ("prime") has the sequence 5'-CAACGGAGGAG. All of the KL sequences were named in this manner, for now.

4.3 Hairpin design

To test the new KL sets, a GFPas-functionalized RNA fiber monomer was used as a basis (Figure 4.1, left panel). The 5'-end sequence of the "dumbbell" monomer from

the RNA fibers was kept and 3'-end of the monomer was removed, keeping the HIV-like motif intact. Furthermore, a GFP sense (GFPs) sequence was added at the 3'-end of the structure in place of GFPas toehold (Figure 4.1, right, and Figure 4.2, panel B), resulting in a hairpin structure with a toehold. Two uracil bases separate the hairpin from the GFP toehold sequence to allow flexibility the toehold when bound to its complement sequence. A GFPas displacement strand serves as a placeholder to prevent formation of any undesirable secondary structures on the toehold (sticky end) end of the hairpin. GFPs toehold on the structure would allow to test these hairpin monomers with existing nanoring¹⁰ monomer strands, and further allow them to be used as linkers to attach nanorings^{10, 109} or nanocubes^{11, 109} to a fiber structure via the GFP sticky ends.

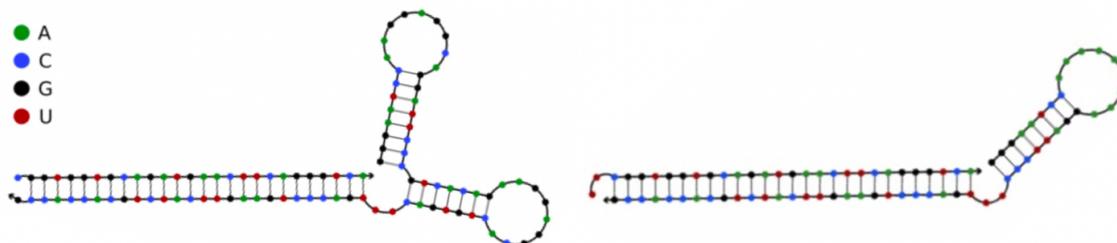


Figure 4.1: Secondary structure of a functionalized RNA fiber and KL hairpin. Secondary structures of HIV1_GFPs strand (left), and HIV0_GFPs strand (right) produced by NUPACK¹⁰⁶.



Figure 4.2: Template for hairpin strand sequence and its development. (see Figure 4.1)

- In panel **A**, original RNA fiber monomer strand 1, featuring KL1 marked in bold, red letters, and GFP antisense sequence (starting with an orange U) and ending in 3 bold, green nucleotides. The GFPas sequence can self-interact, highlighted with bold, green nucleotides and two bold interior C and G nucleotides.
- In panel **B**, modified hairpin sequence has only one kissing loop (from the 5'-end of the sequence) and a GFP sense (opposed to GFPas) sequence, starting with an orange U and ending with 3 bold, orange nucleotides. Region within the sequence starting at purple G (guanine) and ending at purple C (cytosine) in purple within the sequence of panel A has been removed for sequence of panel B. This resultant strand (panel B) would form the hairpin structure with the "top" portion of the strand sequence, and a toehold (sticky end) with the "bottom" portion of the strand sequence.
- Sequences in panels **C** and **D** are the same, except they are shown in different directions. By convention, nucleic acid sequences are listed in a 5' to 3' direction, except in panel C it is listed in opposite direction to more easily see that the 27nt Displacement Strand (in panel C, and D) is complementary to the 3' end of the sequence in panel B.

4.4 Kissing Loop crossTalk Results Matrix design

When checking KL1 against other 31 KL monomers, plus three controls, 34 wells would be needed on a single gel (or a gel can be separated in multiple smaller gels), where the maximum number of wells on a large gel is 20 wells, thus creating a problem. Bigger problem arises in the arrangement of samples for analysis on gels. When looking at the

subsequent gel setups, every next gel would have already checked for one of the previous combinations, that is, KL2 no longer needs to be checked against KL1, but rather KL3, KL4, etc. In this manner, the next gel would need three controls and KL2 monomer checked against 30 KL monomers, for a total of 33 wells. The following gel would need 32 wells, 31 wells, 30 wells, etc., until some of the last gels would need only a few or just one well and combining combinations in this manner was undesired. The results would then look like a linear algebra matrix in echelon form of (a diagonal triangle) (see Table 4.2, below). Another issue with the KL combination matrix in Table 4.2 is that when each KL is crossed with every KL in the KL “pool”, every uniquely crossed KL combination shows up twice in the matrix (see the grayed out bottom “half” of the matrix in Table 4.2, in darker gray). The two sections about the echelon diagonal contain the same KL combination sets. Also, every KL mixed with itself is grayed out in dark gray, since mixing KL 1 with KL1 (KL1 + KL1) is not really mixing it, but rather just adding more of the same strand to the mix. These combinations were grayed out because they were removed from the combination matrix.

To simplify the matrix by removing redundant KL combinations, all “homogeneous” (KL non-prime + KL non-prime, or KL prime + KL prime) KL combinations were selected for in the top section of the matrix (see Figure 4.4) from the “heterogeneous” (KL non-prime + KL prime, or KL prime + KL non-prime) KL combinations which were selected for in the bottom section of the matrix. All selected KL combinations were kept in their positions, while non-selected KL combinations were deleted. Selected KL combinations from the matrix were then slightly rearranged (see

Table 4.3, below) which yielded a roughly the same amount of combination sets to test in every row and column. This set up is much easier to arrange in gel analysis, and much easier to see the organization of data than without the matrix. The matrix arrangement may be complicated, but is very organized, and may be a work of art in itself. It makes seeing the crossTalk data immensely easier to analyze. Upon analysis of the KL numbering trends, any needed combination can be easily found on the matrix. The main body of the matrix contains 496 unique KL combinations, not counting the 16 redundant combinations, all the controls, and the initial KL monomer tests.

Table 4.2: Initial KL crossTalk combination matrix. KL combinations in darker gray are redundant (due to self-combination) and are later removed (see Figure 4.3, below). Combination sets above and below the diagonal contain duplicate KL combinations that will need to be removed also due to redundancy.

	1	1'	2	2'	3	3'	4	4'	5	5'	6	6'	7	7'	8	8'	9	9'	10	10'	11	11'	12	12'	13	13'	14	14'	15	15'	16	16'
1	1	1+1	1+2	1+3	1+3	1+4	1+4	1+5	1+5	1+6	1+6	1+7	1+7	1+8	1+8	1+9	1+9	1+10	1+10	1+11	1+11	1+12	1+12	1+13	1+13	1+14	1+14	1+15	1+15	1+16	1+16	
1'	1+1	1	1+2	1+2	1+3	1+3	1+4	1+4	1+5	1+5	1+6	1+6	1+7	1+7	1+8	1+8	1+9	1+9	1+10	1+10	1+11	1+11	1+12	1+12	1+13	1+13	1+14	1+14	1+15	1+15	1+16	1+16
2	1+1	1+1	2	2	2+3	2+3	2+4	2+4	2+5	2+5	2+6	2+6	2+7	2+7	2+8	2+8	2+9	2+9	2+10	2+10	2+11	2+11	2+12	2+12	2+13	2+13	2+14	2+14	2+15	2+15	2+16	2+16
2'	2+1	2+1	2	2	2+3	2+3	2+4	2+4	2+5	2+5	2+6	2+6	2+7	2+7	2+8	2+8	2+9	2+9	2+10	2+10	2+11	2+11	2+12	2+12	2+13	2+13	2+14	2+14	2+15	2+15	2+16	2+16
3	2+1	2+1	2+2	2+2	2+3	2+3	2+4	2+4	2+5	2+5	2+6	2+6	2+7	2+7	2+8	2+8	2+9	2+9	2+10	2+10	2+11	2+11	2+12	2+12	2+13	2+13	2+14	2+14	2+15	2+15	2+16	2+16
3'	2+1	2+1	2+2	2+2	2+3	2+3	2+4	2+4	2+5	2+5	2+6	2+6	2+7	2+7	2+8	2+8	2+9	2+9	2+10	2+10	2+11	2+11	2+12	2+12	2+13	2+13	2+14	2+14	2+15	2+15	2+16	2+16
4	4+1	4+1	4+2	4+2	4+3	4+3	4+4	4+4	4+5	4+5	4+6	4+6	4+7	4+7	4+8	4+8	4+9	4+9	4+10	4+10	4+11	4+11	4+12	4+12	4+13	4+13	4+14	4+14	4+15	4+15	4+16	4+16
4'	4+1	4+1	4+2	4+2	4+3	4+3	4+4	4+4	4+5	4+5	4+6	4+6	4+7	4+7	4+8	4+8	4+9	4+9	4+10	4+10	4+11	4+11	4+12	4+12	4+13	4+13	4+14	4+14	4+15	4+15	4+16	4+16
5	5+1	5+1	5+2	5+2	5+3	5+3	5+4	5+4	5+5	5+5	5+6	5+6	5+7	5+7	5+8	5+8	5+9	5+9	5+10	5+10	5+11	5+11	5+12	5+12	5+13	5+13	5+14	5+14	5+15	5+15	5+16	5+16
5'	5+1	5+1	5+2	5+2	5+3	5+3	5+4	5+4	5+5	5+5	5+6	5+6	5+7	5+7	5+8	5+8	5+9	5+9	5+10	5+10	5+11	5+11	5+12	5+12	5+13	5+13	5+14	5+14	5+15	5+15	5+16	5+16
6	6+1	6+1	6+2	6+2	6+3	6+3	6+4	6+4	6+5	6+5	6+6	6+6	6+7	6+7	6+8	6+8	6+9	6+9	6+10	6+10	6+11	6+11	6+12	6+12	6+13	6+13	6+14	6+14	6+15	6+15	6+16	6+16
6'	6+1	6+1	6+2	6+2	6+3	6+3	6+4	6+4	6+5	6+5	6+6	6+6	6+7	6+7	6+8	6+8	6+9	6+9	6+10	6+10	6+11	6+11	6+12	6+12	6+13	6+13	6+14	6+14	6+15	6+15	6+16	6+16
7	7+1	7+1	7+2	7+2	7+3	7+3	7+4	7+4	7+5	7+5	7+6	7+6	7+7	7+7	7+8	7+8	7+9	7+9	7+10	7+10	7+11	7+11	7+12	7+12	7+13	7+13	7+14	7+14	7+15	7+15	7+16	7+16
7'	7+1	7+1	7+2	7+2	7+3	7+3	7+4	7+4	7+5	7+5	7+6	7+6	7+7	7+7	7+8	7+8	7+9	7+9	7+10	7+10	7+11	7+11	7+12	7+12	7+13	7+13	7+14	7+14	7+15	7+15	7+16	7+16
8	8+1	8+1	8+2	8+2	8+3	8+3	8+4	8+4	8+5	8+5	8+6	8+6	8+7	8+7	8+8	8+8	8+9	8+9	8+10	8+10	8+11	8+11	8+12	8+12	8+13	8+13	8+14	8+14	8+15	8+15	8+16	8+16
8'	8+1	8+1	8+2	8+2	8+3	8+3	8+4	8+4	8+5	8+5	8+6	8+6	8+7	8+7	8+8	8+8	8+9	8+9	8+10	8+10	8+11	8+11	8+12	8+12	8+13	8+13	8+14	8+14	8+15	8+15	8+16	8+16
9	9+1	9+1	9+2	9+2	9+3	9+3	9+4	9+4	9+5	9+5	9+6	9+6	9+7	9+7	9+8	9+8	9+9	9+9	9+10	9+10	9+11	9+11	9+12	9+12	9+13	9+13	9+14	9+14	9+15	9+15	9+16	9+16
9'	9+1	9+1	9+2	9+2	9+3	9+3	9+4	9+4	9+5	9+5	9+6	9+6	9+7	9+7	9+8	9+8	9+9	9+9	9+10	9+10	9+11	9+11	9+12	9+12	9+13	9+13	9+14	9+14	9+15	9+15	9+16	9+16
10	10+1	10+1	10+2	10+2	10+3	10+3	10+4	10+4	10+5	10+5	10+6	10+6	10+7	10+7	10+8	10+8	10+9	10+9	10+10	10+10	10+11	10+11	10+12	10+12	10+13	10+13	10+14	10+14	10+15	10+15	10+16	10+16
10'	10+1	10+1	10+2	10+2	10+3	10+3	10+4	10+4	10+5	10+5	10+6	10+6	10+7	10+7	10+8	10+8	10+9	10+9	10+10	10+10	10+11	10+11	10+12	10+12	10+13	10+13	10+14	10+14	10+15	10+15	10+16	10+16
11	11+1	11+1	11+2	11+2	11+3	11+3	11+4	11+4	11+5	11+5	11+6	11+6	11+7	11+7	11+8	11+8	11+9	11+9	11+10	11+10	11+11	11+11	11+12	11+12	11+13	11+13	11+14	11+14	11+15	11+15	11+16	11+16
11'	11+1	11+1	11+2	11+2	11+3	11+3	11+4	11+4	11+5	11+5	11+6	11+6	11+7	11+7	11+8	11+8	11+9	11+9	11+10	11+10	11+11	11+11	11+12	11+12	11+13	11+13	11+14	11+14	11+15	11+15	11+16	11+16
12	12+1	12+1	12+2	12+2	12+3	12+3	12+4	12+4	12+5	12+5	12+6	12+6	12+7	12+7	12+8	12+8	12+9	12+9	12+10	12+10	12+11	12+11	12+12	12+12	12+13	12+13	12+14	12+14	12+15	12+15	12+16	12+16
12'	12+1	12+1	12+2	12+2	12+3	12+3	12+4	12+4	12+5	12+5	12+6	12+6	12+7	12+7	12+8	12+8	12+9	12+9	12+10	12+10	12+11	12+11	12+12	12+12	12+13	12+13	12+14	12+14	12+15	12+15	12+16	12+16
13	13+1	13+1	13+2	13+2	13+3	13+3	13+4	13+4	13+5	13+5	13+6	13+6	13+7	13+7	13+8	13+8	13+9	13+9	13+10	13+10	13+11	13+11	13+12	13+12	13+13	13+13	13+14	13+14	13+15	13+15	13+16	13+16
13'	13+1	13+1	13+2	13+2	13+3	13+3	13+4	13+4	13+5	13+5	13+6	13+6	13+7	13+7	13+8	13+8	13+9	13+9	13+10	13+10	13+11	13+11	13+12	13+12	13+13	13+13	13+14	13+14	13+15	13+15	13+16	13+16
14	14+1	14+1	14+2	14+2	14+3	14+3	14+4	14+4	14+5	14+5	14+6	14+6	14+7	14+7	14+8	14+8	14+9	14+9	14+10	14+10	14+11	14+11	14+12	14+12	14+13	14+13	14+14	14+14	14+15	14+15	14+16	14+16
14'	14+1	14+1	14+2	14+2	14+3	14+3	14+4	14+4	14+5	14+5	14+6	14+6	14+7	14+7	14+8	14+8	14+9	14+9	14+10	14+10	14+11	14+11	14+12	14+12	14+13	14+13	14+14	14+14	14+15	14+15	14+16	14+16
15	15+1	15+1	15+2	15+2	15+3	15+3	15+4	15+4	15+5	15+5	15+6	15+6	15+7	15+7	15+8	15+8	15+9	15+9	15+10	15+10	15+11	15+11	15+12	15+12	15+13	15+13	15+14	15+14	15+15	15+15	15+16	15+16
15'	15+1	15+1	15+2	15+2	15+3	15+3	15+4	15+4	15+5	15+5	15+6	15+6	15+7	15+7	15+8	15+8	15+9	15+9	15+10	15+10	15+11	15+11	15+12	15+12	15+13	15+13	15+14	15+14	15+15	15+15	15+16	15+16
16	16+1	16+1	16+2	16+2	16+3	16+3	16+4	16+4	16+5	16+5	16+6	16+6	16+7	16+7	16+8	16+8	16+9	16+9	16+10	16+10	16+11	16+11	16+12	16+12	16+13	16+13	16+14	16+14	16+15	16+15	16+16	16+16
16'	16+1	16+1	16+2	16+2	16+3	16+3	16+4	16+4	16+5	16+5	16+6	16+6	16+7	16+7	16+8	16+8	16+9	16+9	16+10	16+10	16+11	16+11	16+12	16+12	16+13	16+13	16+14	16+14	16+15	16+15	16+16	16+16

Table 4.3: Reduced KL crossTalk combination matrix. Self KL combinations (from dark gray cells, see Figure 4.2, above) are removed. “Heterogeneous” KL combinations are removed from the combination “pool” above the dark diagonal cells, and “homogeneous” KL combinations are removed from the combination “pool” below the dark diagonal cells.

	1	1'	2	2'	3	3'	4	4'	5	5'	6	6'	7	7'	8	8'	9	9'	10	10'	11	11'	12	12'	13	13'	14	14'	15	15'	16	16'	
1																																	
1'	1+1																																
2	2+1	1+2																															
2'	2'+1	2+2	1+3																														
3	3+1	3+2	3+3	1+4																													
3'	3'+1	3'+2	3'+3	3'+4	1+5																												
4	4+1	4+2	4+3	4+4	4+5	1+6																											
4'	4'+1	4'+2	4'+3	4'+4	4'+5	4'+6	1+7																										
5	5+1	5+2	5+3	5+4	5+5	5+6	5+7	1+8																									
5'	5'+1	5'+2	5'+3	5'+4	5'+5	5'+6	5'+7	5'+8	1+9																								
6	6+1	6+2	6+3	6+4	6+5	6+6	6+7	6+8	6+9	1+10																							
6'	6'+1	6'+2	6'+3	6'+4	6'+5	6'+6	6'+7	6'+8	6'+9	6'+10	1+11																						
7	7+1	7+2	7+3	7+4	7+5	7+6	7+7	7+8	7+9	7+10	7+11	1+12																					
7'	7'+1	7'+2	7'+3	7'+4	7'+5	7'+6	7'+7	7'+8	7'+9	7'+10	7'+11	7'+12	1+13																				
8	8+1	8+2	8+3	8+4	8+5	8+6	8+7	8+8	8+9	8+10	8+11	8+12	8+13	1+14																			
8'	8'+1	8'+2	8'+3	8'+4	8'+5	8'+6	8'+7	8'+8	8'+9	8'+10	8'+11	8'+12	8'+13	8'+14	1+15																		
9	9+1	9+2	9+3	9+4	9+5	9+6	9+7	9+8	9+9	9+10	9+11	9+12	9+13	9+14	9+15	1+16																	
9'	9'+1	9'+2	9'+3	9'+4	9'+5	9'+6	9'+7	9'+8	9'+9	9'+10	9'+11	9'+12	9'+13	9'+14	9'+15	9'+16																	
10	10+1	10+2	10+3	10+4	10+5	10+6	10+7	10+8	10+9	10+10	10+11	10+12	10+13	10+14	10+15	10+16																	
10'	10'+1	10'+2	10'+3	10'+4	10'+5	10'+6	10'+7	10'+8	10'+9	10'+10	10'+11	10'+12	10'+13	10'+14	10'+15	10'+16																	
11	11+1	11+2	11+3	11+4	11+5	11+6	11+7	11+8	11+9	11+10	11+11	11+12	11+13	11+14	11+15	11+16																	
11'	11'+1	11'+2	11'+3	11'+4	11'+5	11'+6	11'+7	11'+8	11'+9	11'+10	11'+11	11'+12	11'+13	11'+14	11'+15	11'+16																	
12	12+1	12+2	12+3	12+4	12+5	12+6	12+7	12+8	12+9	12+10	12+11	12+12	12+13	12+14	12+15	12+16																	
12'	12'+1	12'+2	12'+3	12'+4	12'+5	12'+6	12'+7	12'+8	12'+9	12'+10	12'+11	12'+12	12'+13	12'+14	12'+15	12'+16																	
13	13+1	13+2	13+3	13+4	13+5	13+6	13+7	13+8	13+9	13+10	13+11	13+12	13+13	13+14	13+15	13+16																	
13'	13'+1	13'+2	13'+3	13'+4	13'+5	13'+6	13'+7	13'+8	13'+9	13'+10	13'+11	13'+12	13'+13	13'+14	13'+15	13'+16																	
14	14+1	14+2	14+3	14+4	14+5	14+6	14+7	14+8	14+9	14+10	14+11	14+12	14+13	14+14	14+15	14+16																	
14'	14'+1	14'+2	14'+3	14'+4	14'+5	14'+6	14'+7	14'+8	14'+9	14'+10	14'+11	14'+12	14'+13	14'+14	14'+15	14'+16																	
15	15+1	15+2	15+3	15+4	15+5	15+6	15+7	15+8	15+9	15+10	15+11	15+12	15+13	15+14	15+15	15+16																	
15'	15'+1	15'+2	15'+3	15'+4	15'+5	15'+6	15'+7	15'+8	15'+9	15'+10	15'+11	15'+12	15'+13	15'+14	15'+15	15'+16																	
16	16+1	16+2	16+3	16+4	16+5	16+6	16+7	16+8	16+9	16+10	16+11	16+12	16+13	16+14	16+15	16+16																	
16'	16'+1	16'+2	16'+3	16'+4	16'+5	16'+6	16'+7	16'+8	16'+9	16'+10	16'+11	16'+12	16'+13	16'+14	16'+15	16'+16																	

4.5 KL strand crossTalk experiments

This experiment was set up to check strand inter-compatibility as a group of KL sets to be used simultaneously, and to assess whether crosstalk between KL strands exists and how much. As mentioned above, strand KL0 was designed as a negative control and was meant to not interact with any other KL strands. Each KL strand (KL1, KL1', KL2, KL2', etc.) was tested against every KL strand separately in an Eppendorf tube to visualize whether they only interact as intended (KL1 + KL2, KL3' + KL4', etc.) or there is crosstalk present as in KL4' & KL9' interacting, when they were not supposed to (see Table on p64). Each KL strand having a GFPs toehold (for future use) was mixed (1:1) with a 27nt DNA displacement strand, which binds from the 3' end of each KL strand

(see Figure 4.1, right). The purpose of the 27nt displacement strand is to bind to the GFPs toehold to prevent unwanted secondary structure formation and help with gel staining. GFPs toehold on the KL strands is not required for testing crosstalk, but it greatly helps visualize KL strands with ethidium bromide total staining, and it gives an option to easily utilize them with other projects.

4.6 Methods for crossTalk experiments

Concentration of purified RNA strands in primary stocks was determined using the NanoDrop 2000 spectrophotometer. Although micropipettes are accurate instruments, they require calibration on regular intervals, and their accuracy may have inaccuracies. Using primary stocks of RNA material to assemble nanoconstructs may add inaccuracies, since every amount of liquid pipetted will require adjustment of the micropipette. This introduces two major issues: 1) frequent micropipette volume adjustments adds tedious work, take time, and slowly “wears out” the calibration of the micropipette, and 2) frequent micropipette volume adjustments add to the pipetting error, resulting in less accurate fluid additions into the assembling mixture. For this reason, secondary RNA stocks were prepared, where every fluid addition to the mixture would be at the same volume (with few exceptions). This prevents using different pipettes (which may introduce their own systemic errors), where a specific micropipette would be the only instrument in use throughout the experiment, minimizing the overall error in pipetting aliquot volumes.

RNA molecules are much more sensitive to temperature and may spontaneously hydrolyze bonds,¹¹⁰ for which reason RNA samples are best kept on ice (at 4 °C) during

benchtop work with them. For crossTalk experiments the working volumes of each final assembly were 5 μ L, which may be very sensitive to any environmental conditions. Liquid aliquots may evaporate much faster, Eppendorf tube temperatures may play a role, keeping the lids of Eppendorf tubes open or closed, and reagent aliquot temperatures may be important in preservation of sample integrity. For these reasons multiple “safety approaches” were implemented. Plenty of ice was readily available to keep secondary stocks cold. Empty 1.5 mL Eppendorf tubes were labelled appropriately for each assembly and placed on ice to pre-cool. Ultra-Pure water (endotoxin-free) used for assembly was pre-cooled on ice prior to assembly, to ensure that the RNA samples were added into water of the same temperature as RNA sample aliquots. Assembly buffer (5x AB) was also pre-cooled on ice prior to its addition to the mixtures to ensure isothermic conditions upon snap-cooling (at 4 °C) of samples right after melting step (at 95 °C), which may alter the dynamics of assembling molecules that need reduced entropy during this step.¹¹¹ After assembly of samples and before their loading onto a native-PAG, samples were mixed with loading buffer, which was also pre-cooled on ice to ensure isothermic conditions and to reduce temperature fluctuations.

Native polyacrylamide gels (native-PAGs) were prepared fresh for every assembly, typically in sets of 2 or 4, depending on availability of equipment. Gels were prepared in sets to ensure that the results would be as similar between gels as possible. Gels polymerizing for different duration of time may behave slightly different; this may happen if gels are not completely polymerized, due to short polymerization time allotted or when integrity of certain reagents is questionable or compromised.

Since RNA dimers are very small, 19:1 acrylamide to bisacrylamide (A:bisA) was selected for use, respectively. All gels for crossTalk experiments were 8% native-PAGs using 19:1 A:bisA and having a 2 mM Mg^{2+} ion concentration.

All RNA monomers prepared for these experiments were prepared at the final concentration of 1 μ M, and all RNA dimers were prepared at the final concentration of 0.5 μ M. This ensures that using same volumes to load onto the gels results in equal total RNA (or nucleic acid (NA) for that matter) content per well. If total NA content per well is not equal, then bands in different well lanes do not stain equally, resulting in some bands (with more NA content in that well lane) showing up dark and other bands (with less NA content in that well lane) showing up light upon visualization. This may result in issues when imaging the gel in terms of adjusting contrast and brightness of the whole gel. Therefore, ensuring that total RNA (or NA) content per each loaded well reduces the potential issue upon total gel staining with ethidium bromide and subsequent imaging of the gel on ChemiDoc™ MP Imaging System.

4.7 Kissing loop T_m determination

It is important to know the stability of various components to construct a stable and reliable nanoparticle or nanoconstruct that would have predictable characteristics, because that can be used as a leverage in controlling the dynamics, biodistribution or cargo release from the nanoconstruct. For this reason, the melting temperature (T_m) of KL complexes needed to be determined. Thermal-gradient gel electrophoresis (TGGE) apparatus was available to perform that task (see Figure 4.3, below). It has a unique horizontal platform with a thermal block for gel placement with integrated heating

elements that can heat the platform in a gradient fashion, where temperature can be set within 30 °C ranging from left to right of the platform. The lower temperature is limited by environment (room temperature, or possible cold room) and upper temperature is 80 °C, which may potentially be a problem as the integrity of the gel may be compromised at that temperature. The thermal block also has two removable and orientation adjustable buffer containers that can be positioned to be on north and south, or on east and west positions of the thermal block platform. There are different gel positions and orientations that can be set up, which are provided in the manufacturer's apparatus manual. The apparatus also has a unique power source that needs to have a custom program protocol set up to run different nanoconstructs on the gel. A native-PAG positioned on the thermal block is shown below in Figure 4.4, left. The same gel visualized under UV is also shown below in Figure 4.4, right.

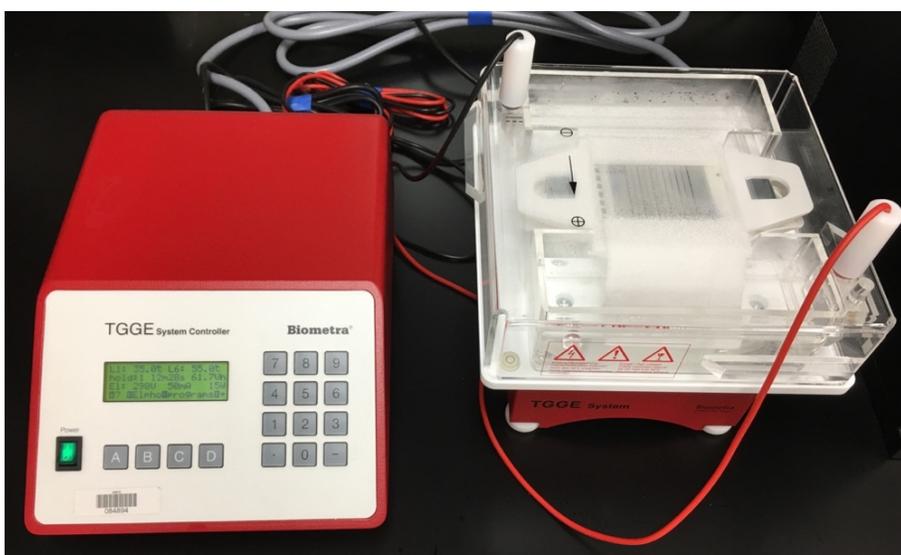


Figure 4.3: Temperature-gradient gel electrophoresis (TGGE) apparatus during an active experiment. On the left is the power source, and gel platform with thermal block is on the right.

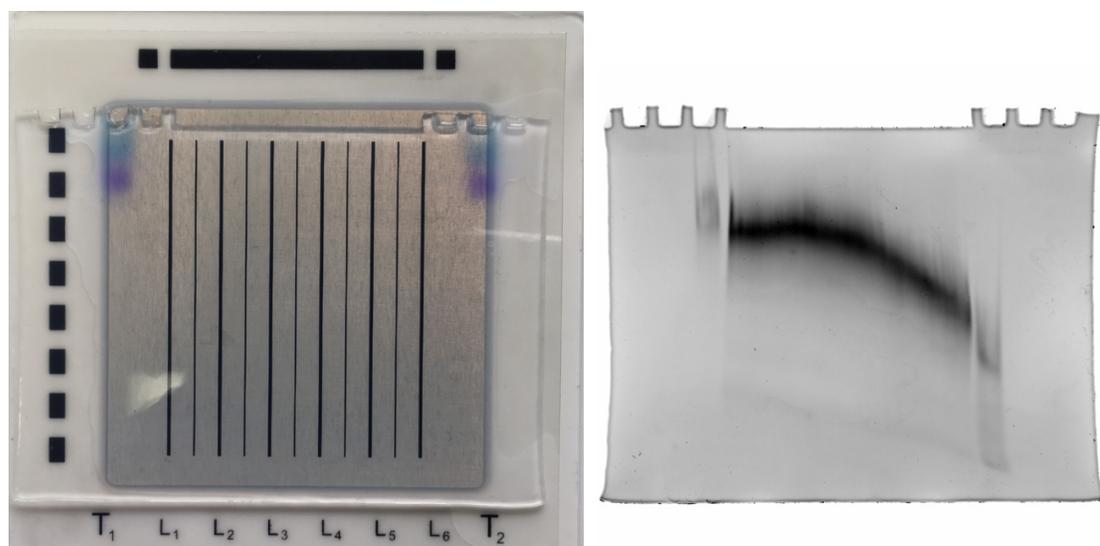


Figure 4.4: TGGE thermal block and a sample native-PAG. Sample native-PAG positioned on the TGGE thermal block (left), and the same gel stained with ethidium bromide and visualized on UV transilluminator ChemiDoc (right).

CHAPTER 5: RESULTS AND DISCUSSION

5.1 Results from *in vitro* transcription optimization

To verify and optimize the transcription recipe in use by the lab an experiment was performed altering the standard recipe for *in vitro* transcription, or RNA synthesis using T7 RNA polymerase. Typically, T7 RNA polymerase is added a standard amount of 5 μL per 50 μL of transcription reaction. Questioning whether rNTPs are also overused and what the optimal amount of T7 RNA polymerase per reaction is, an experiment was set up. rNTP amounts were varied from 1 μL to the standard 10 μL additions, and a newly in-house produced T7 RNA polymerase amounts were varied from 0.5 μL to 5 μL per 50 μL transcription reaction, while keeping all other conditions constant.

Curiosity about concentrations between the in-house expressed T7 polymerase and the commercially available T7 RNA polymerase (by Promega) necessitated further investigation. Commercially available T7 RNA polymerase was found to be at 31.79 μM , compared to the in-house expressed T7 polymerase which was at 37.86 μM . The in-house expressed T7 polymerase was later mixed with glycerol (50 % v/v), which reduced the concentration by half. The blank to measure the concentration of both enzymes was prepared by mixing 1x T7 Purification buffer A with glycerol (50% v/v) to give the baseline absorbance to the NanoDrop 2000 spectrophotometer.

Results from this experiment confirmed that rNTP amount of 10 μL (25 mM each nucleotide) is the best amount to use, given the RNA product yields using those amounts (see Table 5.1 and Figure 5.1, below). In regard to T7 RNA polymerase optimal concentration per transcription reaction, results of this experiment suggested that using

more enzyme is not better. It is clear that using less T7 polymerase per reaction produced better RNA product yields (see Table 5.1 and Figure 5.1, below). Based on these results the best concentration of T7 RNA polymerase per 50 μL transcription reaction may be out of scope of the designed experiment, which may need another optimization to determine the optimal concentration. With the conditions tested best final T7 polymerase concentration per transcription reaction was 0.2 μM of the enzyme, while all other conditions were as currently accepted and used by the lab.

Table 5.1: *In vitro* transcription optimization recipes.

sample	[T7 RNA pol.] (μM)	T7 RNA pol. amount (μL)	[rNTPs] (mM each)	rNTPs amount (μL)	Trx Vol (μL)	Trx T7p (μM)	[μM]
strand A	18.93	0.5	25	1.0	50	0.2	2.79
strand A	18.93	0.5	25	2.0	50	0.2	15.51
strand A	18.93	0.5	25	5.0	50	0.2	31.58
strand A	18.93	0.5	25	10.0	50	0.2	149.14
strand A	18.93	1.0	25	1.0	50	0.4	3.44
strand A	18.93	1.0	25	2.0	50	0.4	12.04
strand A	18.93	1.0	25	5.0	50	0.4	30.28
strand A	18.93	1.0	25	10.0	50	0.4	82.52
strand A	18.93	2.0	25	1.0	50	0.8	2.68
strand A	18.93	2.0	25	2.0	50	0.8	9.97
strand A	18.93	2.0	25	5.0	50	0.8	25.98
strand A	18.93	2.0	25	10.0	50	0.8	73.75
strand A	18.93	5.0	25	1.0	50	1.9	2.00
strand A	18.93	5.0	25	2.0	50	1.9	5.41
strand A	18.93	5.0	25	5.0	50	1.9	24.12
strand A	18.93	5.0	25	10.0	50	1.9	50.17

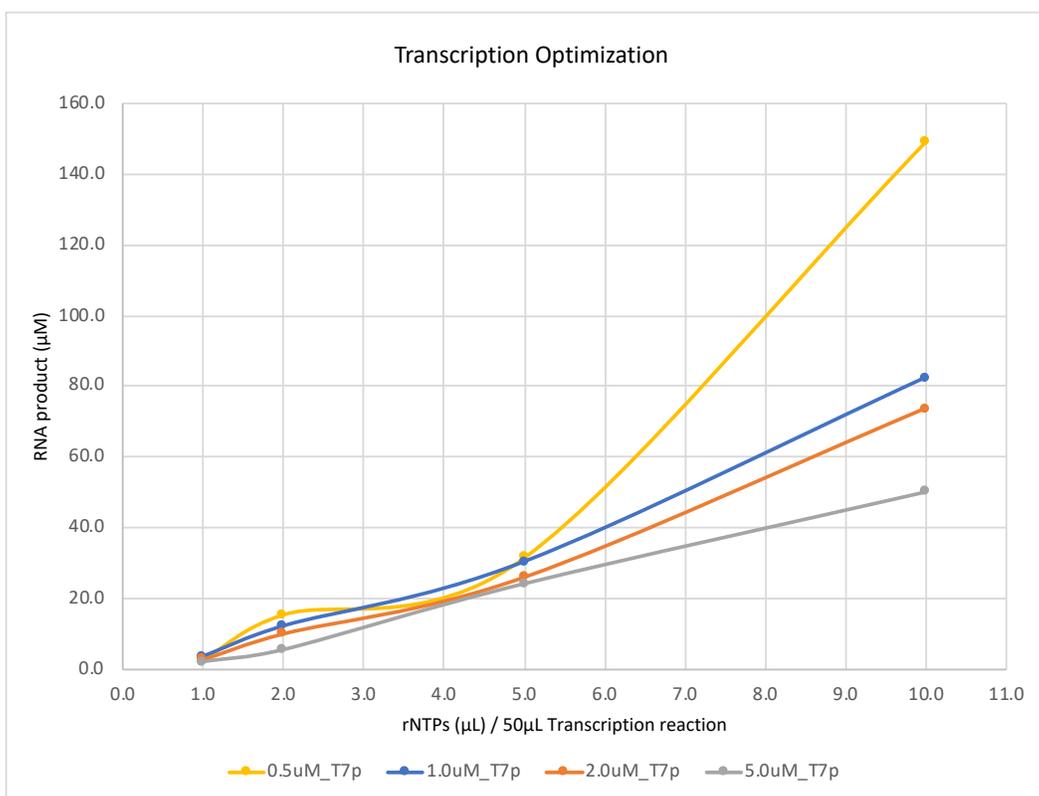


Figure 5.1: *In vitro* optimization RNA product yields at different conditions. Results are based on one experiment replicate.

5.2 Results from T7 RNA Polymerase Optimization for *in vitro* transcription

The previous T7 RNA polymerase optimization experiment required further investigating into the optimal T7 polymerase concentration per transcription reaction. After another in-house T7 RNA polymerase expression, another optimization experiment was devised. Although the initial T7 polymerase concentration was calculated to be 64.46 µM, a small secondary stock of the enzyme was prepared at 10.0 µM to be able to test lower final enzyme concentration per transcription reaction. Final enzyme concentrations per reaction were tested from 0.1 µM to 0.8 µM of T7 RNA polymerase. This time, a peak in RNA product yield can be seen at ~0.3 µM final T7 RNA concentration (see Table 5.2 and Figure 5.2, below), where all other conditions were

held constant. These findings may need to be repeated multiple times and with different samples to confirm.

Table 5.2: T7 RNA polymerase optimization recipe for *in vitro* transcription.

sample	[T7 RNA pol.] (μM)	T7 RNA pol. amount (μL)	[rNTPs] (mM each)	rNTPs amount (μL)	Trx Vol (μL)	Trx T7p (μM)	[μM]
strand A	10.0	0.5	25	10.0	50	0.1	85.95
strand A	10.0	1.0	25	10.0	50	0.2	99.07
strand A	10.0	1.5	25	10.0	50	0.3	108.79
strand A	10.0	2.0	25	10.0	50	0.4	105.64
strand A	10.0	2.5	25	10.0	50	0.5	98.01
strand A	10.0	3.0	25	10.0	50	0.6	93.78
strand A	10.0	3.5	25	10.0	50	0.7	94.83
strand A	10.0	4.0	25	10.0	50	0.8	82.38

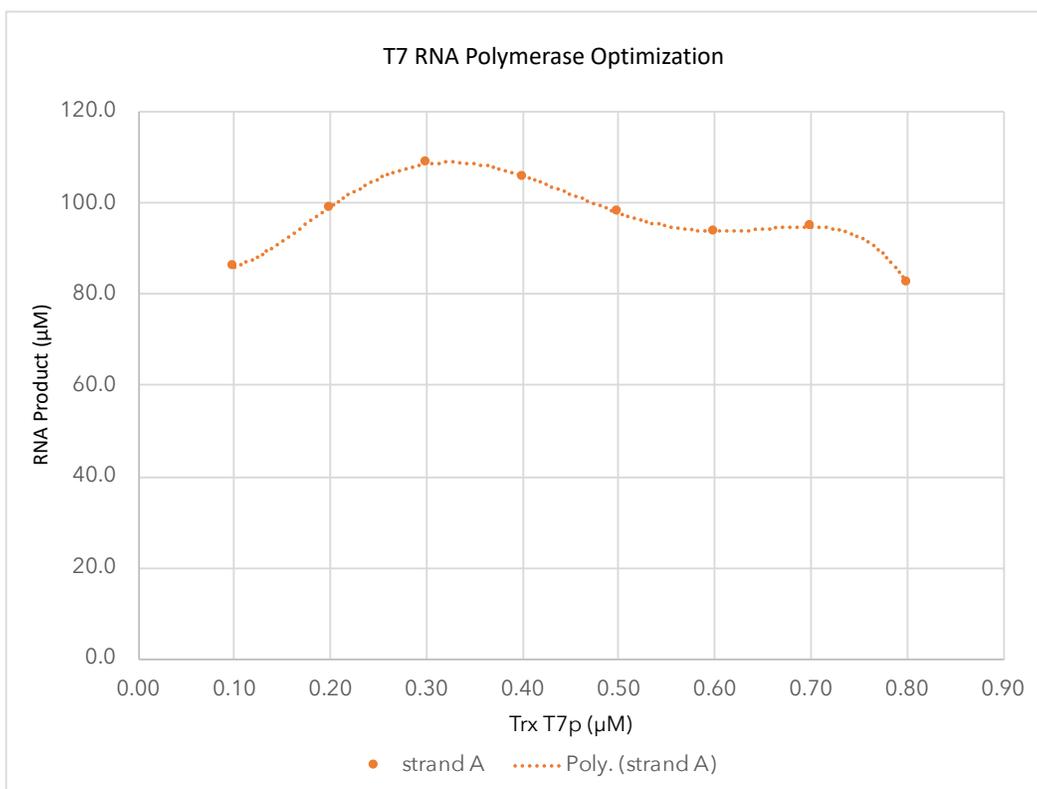


Figure 5.2: RNA product yields from T7 RNA polymerase optimization. Results are based on one experiment replicate.

5.3 Color code decryption of crossTalk Results Matrix

Each of the kissing loop crossTalk experiments is extensive and generates lots of data on many native gels. A system needed to be developed to represent the electrophoretic mobility of nanoconstruct species on the gel and for all data to be readily available for viewing at one time. Each repeat of the crossTalk experiment was either 34 or 68 gels, depending on the size of gels used. It is very difficult to view all gels at one time to analyze data, for which reason a matrix was developed. When analyzing each of the gels, it was necessary to develop a system report each KL interaction as intended or not intended. A matrix or a chart has many cells, where each cell of the matrix represents a reaction of two different KLs (see Tables 5.3, 5.4, A.3.1, and A.3.2; the latter 2 tables are in the Appendix). The green color designates an intended/desired (predicted) KL interaction forming a dimer, and the red color designates an unintended/undesired (not predicted) KL interaction, where the assembly product migrates through the native gel almost identical to the migration of an intended assembly product (a hairpin dimer) (see Figure 5.3, below). A range of shades from red to pale yellow are in decreasing degree of KL monomer interactions, cross-reactivity, or crosstalk. Non-colored cells represent a failed addition of 27nt displacement strand, where assembly products are difficult to visualize. It is difficult to see interactions lacking the displacement strand because it can act as a placeholder in a sticky end region and form double stranded (DNA-RNA) region, which can be intercalated by ethidium bromide dye that is used for gel visualization. And the pale blue/gray color designates no KL interaction.

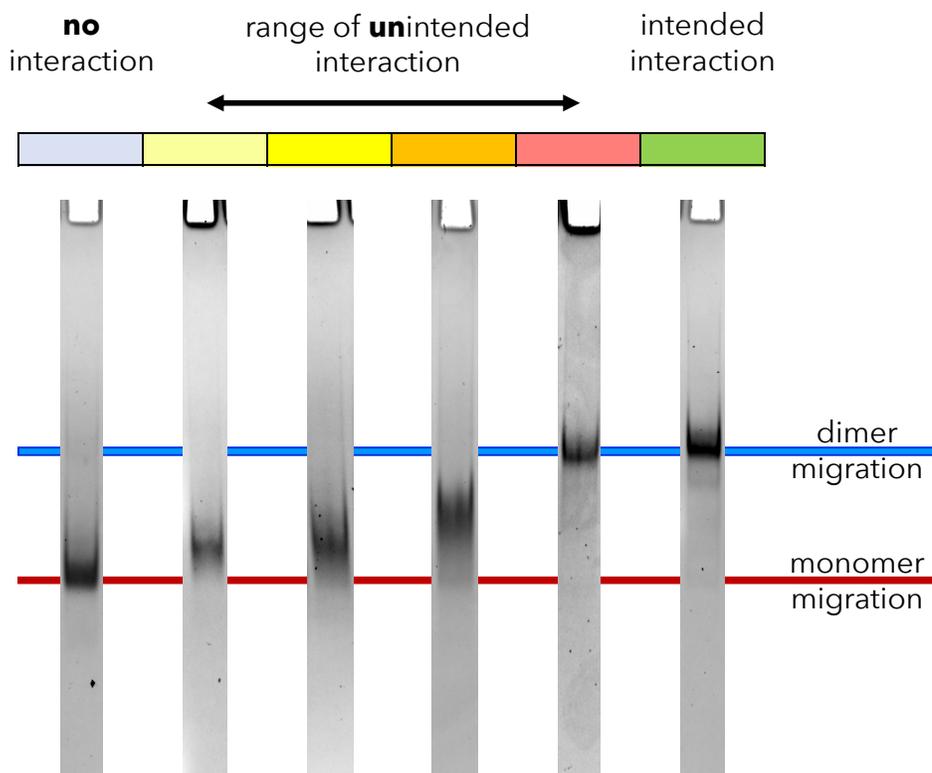


Figure 5.3: Decryption of color codes representing the degree of interaction or crosstalk between a combination of two monomer hairpins. Red line represents the electrophoretic mobility of a hairpin monomer (or two non-interacting hairpin KLS migrating as monomers) on the native-PAG (represented by a light blue/gray color, first sample from the left). Blue line represents the electrophoretic mobility of a dimer (formed by two hairpin KLS interacting in a KL complex (KLC)) on the native-PAG (represented by green color, first sample from the right). Range of colors from pale yellow to red (left to right) represent unintended/undesired KL interactions/crosstalk between two hairpin KLS.

5.4 Kissing Loop crossTalk Results

In the first crossTalk experiment, two gels (Table 5.3 below, gels 34 and 35) were run initially to assess each KL monomer (KL strand + 27nt displacement strand 1:1) formation. After these initial 2 gels, another 32 gels were run to check crosstalk of each KL monomer with the KL “pool”. For example, in one gel, KL0 monomer, KL1 monomer, and KL1 + KL0 were prepared as controls, followed by KL1 + KL2, KL1 + KL3, KL1 + KL4,

until KL1 + KL16. In this manner KL1 was checked against other KL monomers. On the next gel, KL1' was checked against other KL monomers, in the same fashion. KL1 and KL1' loops were also checked against other KLs along the 2nd column of the main body of the matrix. All KL monomers were checked against all other KL monomers (see Table 5.3 below, and Table A.3.1 in Appendix A.3 for a slightly different view of the matrix). And while the layout may be a little confusing, efforts were made to have the results systematically laid out. All "prime" KLs were marked in bold blue text, where non-prime KLs were not.

For the first crossTalk experiment, every row of cells represents the results for one gel, where the first clear cell with a number refers to a gel number (internal nomenclature). First three cells (left → right) are controls, followed by combinations of KL monomer against other KL monomers. Each gel (row) checks a single monomer (for example, gel 36 is checking KL monomer against other KL monomers). At a closer look, KL "prime" strands are absent in gel 36, as they are dispersed among other gels, but KL1 monomer is still checked against "prime" KL monomers, but in the columns; observe that KL1' + KL1 is in gel 37, KL2' + KL1 is in gel 39, KL3' + KL1 is in gel 41, etc. (every other gel, but still in the same column). The odd diagonal in the chart is there for a reason, that is, it is a breaking point at which "normal" and "prime" strand combinations are listed either horizontal or vertical across the gels. Sample gels from crossTalk #1 experiment can be seen in Appendix A.4.

For the second crossTalk experiment (see Table 5.4 below, and Table A.3.2 in Appendix A.3 for a slightly different view of the matrix), every row of cells represents

the results for two gels, doubling the number of gels from 34 to 68 for the whole crossTalk #2 experiment. The body of the matrix is split with a bold line roughly through the middle, that is, between KL combinations containing KL8 and KL9, and KL combinations containing KL8' and KL9'. Corresponding gel numbers (internal nomenclature) are listed on left and right sides of the matrix. Positive (+ C) and negative (– C) controls were loaded on every gel. Sample gels from crossTalk #2 experiment can be seen in Appendix A.5.

Table 5.3: Results from crossTalk #1 experiment. Each row represents one gel from the experiment. Gel number can be seen in the very left column. Nomenclature is greatly simplified, where each hairpin structure fully named KL_1_GFPs is simplified to 1, respectively. Additionally, each hairpin structure is also bound with a GFPasDisplaceStrand. The top two gels (34 and 35) and two left columns with controls only have the monomers, hairpin structures with displacement strands. Additionally, in the 3rd column, KL of that row is tested against KLO, as a negative control. The main matrix shows dimers, where 1 + 0, 2, 3, ... designates KLC (kissing loop complex) 1 + 0, 1 + 2, 1 + 3, ... See Figure 5.3 above for understanding color coding. Also see Table A.3.1 in Appendix A.3 for this crossTalk #1 data in a matrix with a different view.

Gel	(-C)	(-C)	crossTalk #1																	
34	0	1	1'	2	2'	3	3'	4	4'	5	5'	6	6'	7	7'	8	8'			
35	0	9	9'	10	10'	11	11'	12	12'	13	13'	14	14'	15	15'	16	16'			
36	0	1	1+	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
37	0	1'	1'+	0	1	2'	3'	4'	5'	6'	7'	8'	9'	10'	11'	12'	13'	14'	15'	16'
38	0	2	2+	0	1'	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
39	0	2'	2'+	0	1	2	3'	4'	5'	6'	7'	8'	9'	10'	11'	12'	13'	14'	15'	16'
40	0	3	3+	0	1'	2'	3	4	5	6	7	8	9	10	11	12	13	14	15	16
41	0	3'	3'+	0	1	2	3	4'	5'	6'	7'	8'	9'	10'	11'	12'	13'	14'	15'	16'
42	0	4	4+	0	1'	2'	3'	4	5	6	7	8	9	10	11	12	13	14	15	16
43	0	4'	4'+	0	1	2	3	4	5'	6'	7'	8'	9'	10'	11'	12'	13'	14'	15'	16'
44	0	5	5+	0	1'	2'	3'	4'	5	6	7	8	9	10	11	12	13	14	15	16
45	0	5'	5'+	0	1	2	3	4	5	6'	7'	8'	9'	10'	11'	12'	13'	14'	15'	16'
46	0	6	6+	0	1'	2'	3'	4'	5'	6	7	8	9	10	11	12	13	14	15	16
47	0	6'	6'+	0	1	2	3	4	5	6	7'	8'	9'	10'	11'	12'	13'	14'	15'	16'
48	0	7	7+	0	1'	2'	3'	4'	5'	6'	7	8	9	10	11	12	13	14	15	16
49	0	7'	7'+	0	1	2	3	4	5	6	7	8'	9'	10'	11'	12'	13'	14'	15'	16'
50	0	8	8+	0	1'	2'	3'	4'	5'	6'	7'	8	9	10	11	12	13	14	15	16
51	0	8'	8'+	0	1	2	3	4	5	6	7	8	9'	10'	11'	12'	13'	14'	15'	16'
52	0	9	9+	0	1'	2'	3'	4'	5'	6'	7'	8'	9	10	11	12	13	14	15	16
53	0	9'	9'+	0	1	2	3	4	5	6	7	8	9	10'	11'	12'	13'	14'	15'	16'
54	0	10	10+	0	1'	2'	3'	4'	5'	6'	7'	8'	9'	10	11	12	13	14	15	16
55	0	10'	10'+	0	1	2	3	4	5	6	7	8	9	10	11'	12'	13'	14'	15'	16'
56	0	11	11+	0	1'	2'	3'	4'	5'	6'	7'	8'	9'	10'	11	12	13	14	15	16
57	0	11'	11'+	0	1	2	3	4	5	6	7	8	9	10	11	12'	13'	14'	15'	16'
58	0	12	12+	0	1'	2'	3'	4'	5'	6'	7'	8'	9'	10'	11'	12	13	14	15	16
59	0	12'	12'+	0	1	2	3	4	5	6	7	8	9	10	11	12	13'	14'	15'	16'
60	0	13	13+	0	1'	2'	3'	4'	5'	6'	7'	8'	9'	10'	11'	12'	13	14	15	16
61	0	13'	13'+	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14'	15'	16'
62	0	14	14+	0	1'	2'	3'	4'	5'	6'	7'	8'	9'	10'	11'	12'	13'	14	15	16
63	0	14'	14'+	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15'	16'
64	0	15	15+	0	1'	2'	3'	4'	5'	6'	7'	8'	9'	10'	11'	12'	13'	14'	15	16
65	0	15'	15'+	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16'
66	0	16	16+	0	1'	2'	3'	4'	5'	6'	7'	8'	9'	10'	11'	12'	13'	14'	15'	16
67	0	16'	16'+	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16

Table 5.4: Results from crossTalk #2 experiment. Each row represents two gels from the experiment, where separation between the two gels is the bold line running down the whole data matrix. Gel numbers can be seen in the very left and very right columns. Nomenclature is greatly simplified, where each hairpin structure fully named KL_1_GFPs is simplified to **1**, respectively. Additionally, each hairpin structure is also bound with a GFPasDisplaceStrand. The top four gels (96, 97, 98, and 99) only have the hairpin with displacement strand structures. The main matrix shows dimers, where 1 + 0, 2, 3, ... designates KLC (kissing loop complex) 1 + 0, 1 + 2, 1 + 3, ... Columns of positive (dimer formation) and negative (only monomer) controls are marked at the top left of the matrix. Negative control is only a monomer, a hairpin structure with a displacement strand. See Figure 5.3 for understanding color coding. Also see Table A.3.2 in Appendix A.3 for this crossTalk #2 data in a matrix with a different view.

Gel (+C)	(-C)	crossTalk #2																Gel		
96	1+2	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	97	
98	1+2	0	1'	2'	3'	4'	5'	6'	7'	8'	9'	10'	11'	12'	13'	14'	15'	16'	99	
100	1+2	1+	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	101
102	1+2	1'+	0	1	2'	3'	4'	5'	6'	7'	8'	9'	10'	11'	12'	13'	14'	15'	16'	103
104	1+2	2+	0	1'	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	105
106	1+2	2'+	0	1	2	3'	4'	5'	6'	7'	8'	9'	10'	11'	12'	13'	14'	15'	16'	107
108	1+2	3+	0	1'	2'	3	4	5	6	7	8	9	10	11	12	13	14	15	16	109
110	1+2	3'+	0	1	2	3	4'	5'	6'	7'	8'	9'	10'	11'	12'	13'	14'	15'	16'	111
112	1+2	4+	0	1'	2'	3'	4	5	6	7	8	9	10	11	12	13	14	15	16	113
114	1+2	4'+	0	1	2	3	4	5'	6'	7'	8'	9'	10'	11'	12'	13'	14'	15'	16'	115
116	1+2	5+	0	1'	2'	3'	4'	5	6	7	8	9	10	11	12	13	14	15	16	117
118	1+2	5'+	0	1	2	3	4	5	6'	7'	8'	9'	10'	11'	12'	13'	14'	15'	16'	119
120	1+2	6+	0	1'	2'	3'	4'	5'	6	7	8	9	10	11	12	13	14	15	16	121
122	1+2	6'+	0	1	2	3	4	5	6	7'	8'	9'	10'	11'	12'	13'	14'	15'	16'	123
124	1+2	7+	0	1'	2'	3'	4'	5'	6'	7	8	9	10	11	12	13	14	15	16	125
126	1+2	7'+	0	1	2	3	4	5	6	7	8'	9'	10'	11'	12'	13'	14'	15'	16'	127
128	1+2	8+	0	1'	2'	3'	4'	5'	6'	7'	8	9	10	11	12	13	14	15	16	129
130	1+2	8'+	0	1	2	3	4	5	6	7	8	9'	10'	11'	12'	13'	14'	15'	16'	131
132	1+2	9+	0	1'	2'	3'	4'	5'	6'	7'	8'	9	10	11	12	13	14	15	16	133
134	1+2	9'+	0	1	2	3	4	5	6	7	8	9	10'	11'	12'	13'	14'	15'	16'	135
136	1+2	10+	0	1'	2'	3'	4'	5'	6'	7'	8'	9'	10	11	12	13	14	15	16	137
138	1+2	10'+	0	1	2	3	4	5	6	7	8	9	10	11'	12'	13'	14'	15'	16'	139
140	1+2	11+	0	1'	2'	3'	4'	5'	6'	7'	8'	9'	10'	11	12	13	14	15	16	141
142	1+2	11'+	0	1	2	3	4	5	6	7	8	9	10	11	12'	13'	14'	15'	16'	143
144	1+2	12+	0	1'	2'	3'	4'	5'	6'	7'	8'	9'	10'	11'	12	13	14	15	16	145
146	1+2	12'+	0	1	2	3	4	5	6	7	8	9	10	11	12	13'	14'	15'	16'	147
148	1+2	13+	0	1'	2'	3'	4'	5'	6'	7'	8'	9'	10'	11'	12'	13	14	15	16	149
150	1+2	13'+	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14'	15'	16'	151
152	1+2	14+	0	1'	2'	3'	4'	5'	6'	7'	8'	9'	10'	11'	12'	13'	14	15	16	153
154	1+2	14'+	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15'	16'	155
156	1+2	15+	0	1'	2'	3'	4'	5'	6'	7'	8'	9'	10'	11'	12'	13'	14'	15	16	157
158	1+2	15'+	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16'	159
160	1+2	16+	0	1'	2'	3'	4'	5'	6'	7'	8'	9'	10'	11'	12'	13'	14'	15'	16	161
162	1+2	16'+	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	163

5.5 CrossTalk Results filtering

After crossTalk #1 experiment was complete, there was natural interest to filter out all undesired combinations due to their crosstalk interactions. KL1 and KL2 strands were preferentially kept, as they are the basis of the experiment, because the purpose of experiments is to expand one KL complex (KLC) set into a library of KL sets. First, all interactions that were shaded (encoded) in red were removed, followed by other interactions which were shaded orange, and finally yellow encoded interactions. Results from crossTalk #1 experiment underwent four filtering processes (see filtered results of crossTalk #1 experiment, Table 5.5, below). In the first filter, 8 KLCs sets were kept out of 16, where KL5 and KL6 seemed to pose a problem across the board. Looking back at first two gels (34 and 35) of KL monomers, these two monomers (KL5 and KL6) did not appear to have crisp bands to start with, suggesting strand misfolding or self-interaction. If KLC (5 + 6) was to be removed from the KLC group set, then 7 unique KL sets are available. This suggested filtering all results for the second time (see Table 5.5, below), at which point the aforementioned KLC (5 + 6) set was removed and other KLCs were considered. Two other result filters (3 and 4) were performed, to try and find the optimal KLC set group with maximum number of unique kissing loop complexes.

Table 5.5: Kissing loops complex sets kept as a result of 4 different results filters. Kissing loop complexes that cannot coexist in the KL complex set (in vertical columns). All KL dimer combinations form KL complexes well, designated by green color. Different “filtered” KLC sets can potentially coexist in one-pot assembly. These Filter results are based on the data gathered in crossTalk experiment #1.

All KL sets	Filter #1	Filter #2	Filter #3	Filter #4
1 + 2	1 + 2	1 + 2	1 + 2	1 + 2
1' + 2'	1' + 2'	1' + 2'	1' + 2'	1' + 2'
3 + 4	3 + 4	3 + 4	3 + 4	3 + 4
3' + 4'	3' + 4'	3' + 4'	3' + 4'	3' + 4'
5 + 6	5 + 6	5 + 6	5 + 6	5 + 6
5' + 6'	5' + 6'	5' + 6'	5' + 6'	5' + 6'
7 + 8	7 + 8	7 + 8	7 + 8	7 + 8
7' + 8'	7' + 8'	7' + 8'	7' + 8'	7' + 8'
9 + 10	9 + 10	9 + 10	9 + 10	9 + 10
9' + 10'	9' + 10'	9' + 10'	9' + 10'	9' + 10'
11 + 12	11 + 12	11 + 12	11 + 12	11 + 12
11' + 12'	11' + 12'	11' + 12'	11' + 12'	11' + 12'
13 + 14	13 + 14	13 + 14	13 + 14	13 + 14
13' + 14'	13' + 14'	13' + 14'	13' + 14'	13' + 14'
15 + 16	15 + 16	15 + 16	15 + 16	15 + 16
15' + 16'	15' + 16'	15' + 16'	15' + 16'	15' + 16'
	7	8	6	7
All KLCs	Number of KL sets kept in each group after filtering			

5.6 T_m Results of Kissing Loop Complexes from TGGE analysis

Melting temperature (T_m) of KLCs was determined using Temperature Gradient Gel Electrophoresis (TGGE). All KL hairpin strands have a predicted T_m provided by IDT company (see Table A.1 in Appendix A.1). Native-PAG was prepared for each KLC tested, where each KLC dimer was run into the gel vertically on Bio-Rad Mini-PROTEAN® gel

system first, followed by running the gel horizontally on the Biometra TGGE system. The temperature range for T_m #1 experiment was within L₁ and L₆ thermal block bounds and set between 45 °C and 65 °C, respectively (see Figure 4.4, and Figure 5.4).

An approach to calculate the melting temperature is suggested by the manufacturer. The prepared native-PAG was adjusted to have a wide central well that is the width of the adjusted temperature gradient within the bounds of L₁ and L₆ (see Figure 4.4, and Figure 5.4). The wells on the left and right of the central wide well had dimer and monomer controls, respectively (see Figure 5.4). The left-side control (cool side) is a dimer, and the right-side control (hot side) is a monomer. The principle is that the sample in the central wide well will (presumably) melt the nanostructure (KLC), where the left side of the curve (sample in the wide central well) would migrate through the gel at about the rate of the dimer (left control) and the right side of the curve would migrate through the gel at about the rate of the monomer (right control).

To analyze each T_m native-PAG, bounds (left and right, L₁ and L₆, respectively) need to be set (see Figure 5.4). The distance between the bounds needs to be divided into 5 even parts. The temperature difference within the bounds is set and is programmed into the program protocol. The temperature difference of the thermal gradient is then divided by 5 to get the temperature increase rate with every unit. The bounds of main sample (KLC) migration on the gel (top and bottom of the curve) are set and the midpoint is defined horizontally. A vertical line is passed through the intersection of the sample curve and the horizontal curve midpoint. The relative distance from the left bound (L₁) to the crosspoint (Figure 5.4, marked by red vertical

line) is determined. The temperature increase rate per unit is then multiplied by number of relative units from the left bound (L_1), yielding the temperature above the left bound temperature, which is the calculated T_m of the KLC.

Melting temperature of each KLC was determined using the TGGE manufacturer's recommended approach as shown in Figure 5.4 below. Since each KLC is formed by two hairpins, an average of the two KLC components for KLC set was calculated (see Table 5.6, below) out of curiosity. Melting temperature for RNA fibers reported ⁷⁷ was ~ 42 °C for the KL interaction and ~ 59 °C for intramolecular "dumbbell" monomer denaturation. These new KLs have mutated 6nt interacting regions, where KLC 1 + 2 still has the same KLC and the interacting sequence as used in RNA fibers. This suggests that the experimental melting temperatures in Table 5.6 are likely that of the hairpin duplex denaturing and/or the displacement strand melting off of the toehold. Once again, out of curiosity, the difference between the calculated T_m average and experimental T_m for each KLC was calculated (see Table 5.6, below), where the average was ~ 9 °C, confirming the suspicion that the T_m is not true and the experiment needs to be repeated at a lower temperature range.

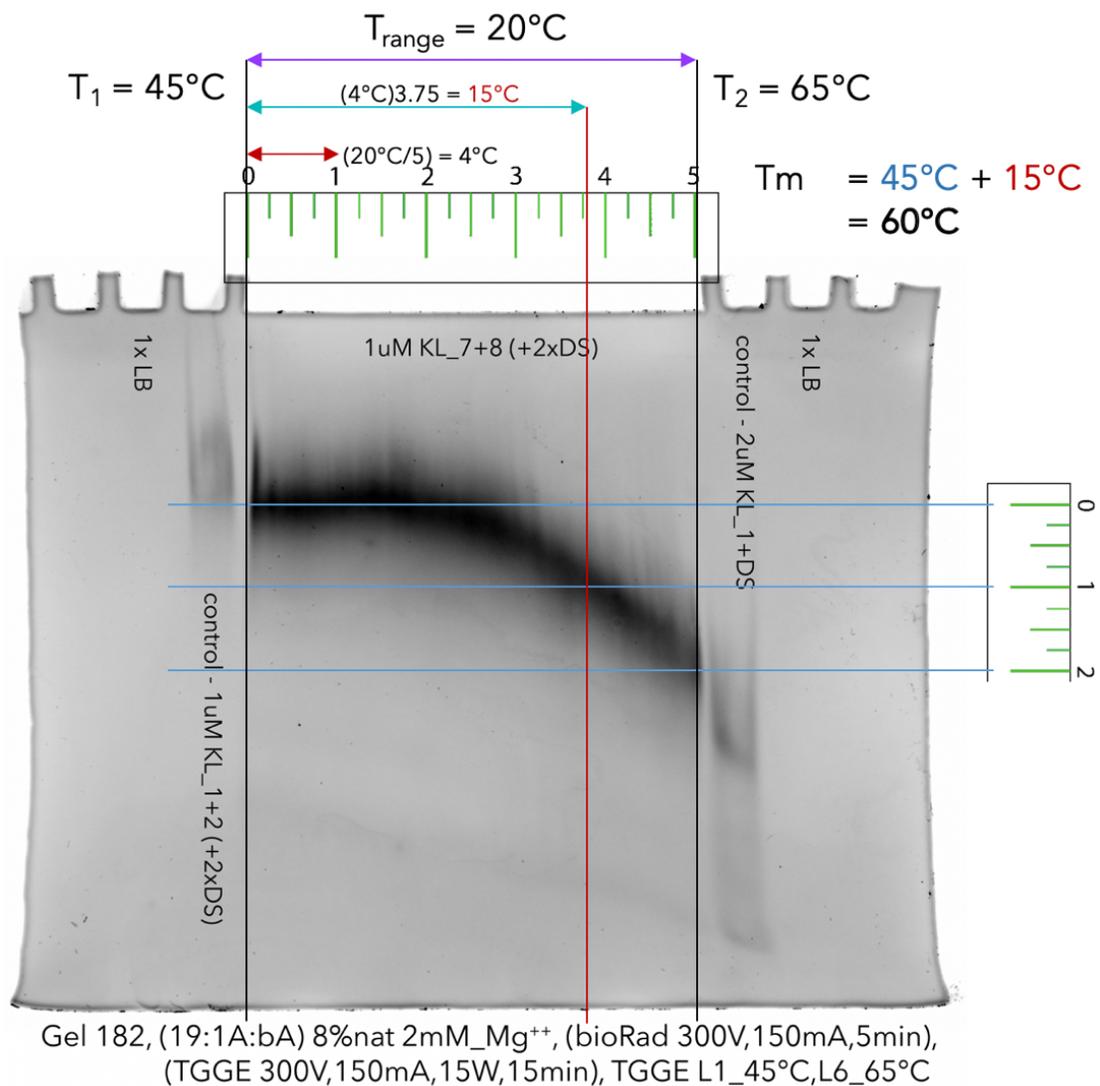


Figure 5.4: Melting temperature determination approach. See Section 5.6 above for detailed explanation of the approach.

Table 5.6: Melting temperature of all tested KLCs. All kissing loop complexes tested for melting temperature are listed on the left side of the table with their respective T_m values found experimentally. Predicted T_m values of each kissing loop provided by the manufacturer are listed on the right side of the table. Calculated average T_m values of each two components of each KLC are also shown on the right side of the table. Difference between calculated average T_m and experimental T_m of each KLC is shown in the left side of the table.

KLC	experimental T_m (°C)	diff. between T_m & $T_{m_{avg}}$ (°C)
1 + 2	60	9.4
1' + 2'	60	8.9
3 + 4	58	10.45
3' + 4'	60	8.9
5 + 6	61	7.85
5' + 6'	57	10.85
7 + 8	60	9.6
7' + 8'	58.2	10.4
9 + 10	59	9.7
9' + 10'	61.8	6.75
11 + 12	57	11.6
11' + 12'	61	8.5
13 + 14	61.8	7.5
13' + 14'	59	9.75
15 + 16	61	7.5
15' + 16'	57.4	11.5
4 + 9	62	6.25
4' + 9'	63	5.65
Avg. T_m diff. (°C)		8.95

KL	predicted T_m (°C)	$T_{m_{avg}}$ (°C)	predicted T_m (°C)	KL
1	70	69.4	68.8	2
1'	69.5	68.9	68.3	2'
3	69.2	68.45	67.7	4
3'	69	68.9	68.8	4'
5	68.9	68.85	68.8	6
5'	67.3	67.85	68.4	6'
7	69.7	69.6	69.5	8
7'	69.3	68.6	67.9	8'
9	68.8	68.7	68.6	10
9'	68.5	68.55	68.6	10'
11	68.5	68.6	68.7	12
11'	69.8	69.5	69.2	12'
13	68.7	69.3	69.9	14
13'	68	68.75	69.5	14'
15	67.7	68.5	69.3	16
15'	68.5	68.9	69.3	16'
4	67.7	68.25	68.8	9
4'	68.8	68.65	68.5	9'

CHAPTER 6: APPLICATIONS OF KISSING LOOPS

6.1 New KLs crosstalk with nanoring KLs

RNA nanorings¹⁰ are a unique RNA structure in that it is composed of 6 monomers each having two unique 120° KLs and able to self-assemble into a planar, hexagonal structure. Each KL utilized in nanorings has 7 actively participating nucleotides (compared to 6nts of the RNA fiber KL, with 3 stacking adenines) that interact (between two KLs) via hydrogen-bonding and are stabilized by magnesium ions (Mg^{2+}) in the solution. Curious to check whether KLs utilized in nanorings would potentially be used in conjunction with KL library, each new hairpin KL was run against each nanoring KL in a NUPACK¹⁰⁶ (nupack.org) online simulation. More specifically, the actively participating KL nucleotides were “excised” from RNA monomers (containing them) for the simulation, as NUPACK tool cannot perform and/or show inter-strand H-bond interaction, but only complementary binding strand interactions. Each nanoring KL active sequence fragment was set to 1 μ M “virtual” concentration and run against another KL active sequence fragment from the KL library also at 1 μ M “virtual” concentration. These “digital” crosstalk results are shown in the Table 6.1, below. These results show that nanoring KLs have some cross-reactivity with most of the new KLs, but there are a few KLs that did not show any reactivity with nanoring’s KLs (marked with nR in the table), and many new KLs that had some cross-reactivity. Some new KLs had a significant cross-reactivity to nanoring’s KLs, based on the simulations. For an application of both types of KLs in a one-pot assembly, experimental work will need to be performed to verify new KLs of interest for this type of platform.

Table 6.1: CrossTalk between active KL fragment sequences of fiber KLs and nanoring KLs.

mix each of two active fragments of sequences @ 1uM		concentration of interacting duplexes	# interacting nucleotides	mix each of two active fragments of sequences @ 1uM		concentration of interacting duplexes	# interacting nucleotides
KL_1 vs	Ab (A3')	0.0087 uM	4	KL_7' vs	fA (A5')	0.0063 uM	4
KL_1 vs	bC (C5')	0.0017 uM	4 + 2	KL_8 vs		nR	
KL_1 vs	Cd (C3')	0.0038 uM	5	KL_8' vs	KL_8'	0.0014 uM	2 + 2
KL_1' vs		nR		KL_8' vs	Fa (F3')	0.0035 uM	5
KL_2 vs	aB (B5')	0.015 uM	5 + 1	KL_9 vs	Ab (A3')	0.0052 uM	5
KL_2' vs		nR		KL_9' vs	bC (C5')	0.0018 uM	6
KL_3 vs	fA (A5')	0.039 uM	4	KL_9' vs	Cd (C3')	0.0045 uM	5
KL_3' vs	Ab (A3')	0.072 uM	5	KL_9' vs	fA (A5')	0.0061 uM	4
KL_3' vs	Cd (C3')	0.003 uM	5	KL_10 vs	De (D3')	0.0016 uM	4
KL_4 vs	aB (B5')	0.001 uM	4 + 2	KL_10 vs	Fa (F3')	0.0029 uM	4
KL_4 vs	Fa (F3')	0.19 uM	6	KL_10' vs	bC (C5')	0.0057 uM	5
KL_4' vs	aB (B5')	0.07 uM	5	KL_11 vs		nR	
KL_4' vs	bC (C5')	0.0033 uM	5	KL_11' vs	Ab (A3')	0.0016 uM	4
KL_5 vs	KL_5	0.39 uM	6	KL_11' vs	fA (A5')	0.0026 uM	4
KL_5 vs	bC (C5')	0.001 uM	5	KL_12 vs	De (D3')	0.0038 uM	3 + 3
KL_5' vs	KL_5'	0.18 uM	6	KL_12' vs		nR	
KL_5' vs	bC (C5')	0.0011 uM	3 + 2	KL_13 vs		nR	
KL_5' vs	De (D3')	0.00088 uM	5	KL_13' vs		nR	
KL_6 vs	KL_6	0.19 uM	4	KL_14 vs	Cd (C3')	0.0017 uM	5
KL_6 vs	bC (C5')	0.008 uM	5	KL_14 vs	fA (A5')	0.0065 uM	4
KL_6' vs	KL_6'	0.019 uM	4	KL_14' vs		nR	
KL_6' vs	bC (C5')	0.015 uM	5	KL_15 vs		nR	
KL_6' vs	De (D3')	0.0051 uM	5	KL_15' vs	Fa (F3')	0.0076 uM	5
KL_7 vs	aB (B5')	0.0037 uM	4	KL_16 vs	bC (C5')	0.0037 uM	6
KL_7 vs	eF (F5')	0.0019 uM	5	KL_16 vs	fA (A5')	0.0063 uM	4
KL_7' vs	KL_7'	0.001 uM	2 + 2	KL_16' vs	Ab (A3')	0.0071 uM	5

nR : no Reaction with other strands

[uM] : considerable interference between two KL active fragment sequences

[uM] : moderate to high interference between two KL active fragment sequences

note: : nanorings consist of 6 strands, labelled from A to F (A, B, C, D, E, F).

: each nanoring monomer interacts with the next via a KL that forms ~120° geometry interaction.

ex: Ab (A3') : denotes in 2 different formats, KL on strand A (located on the 3' end of the strand) that interacts with KL on strand B (located at the 5' end of strand B).

ex: 2 + 2 : two KL active fragment sequences interact with 2 nucleotides,

followed by a bulge of at least 1 nucleotide, then 2 more interacting nucleotides.

6.2 RNA and RNA/DNA hybrid Fiber strands

Hybrid DNA/RNA monomer strands were tested for the RNA fibers. Different length Fiber monomers were developed (8 or 9 base pair, 15bp or 30bp). Each hybrid Fiber monomer in a “dumbbell” shape consists of a DNA helical region and RNA KL regions, where all monomers have the same KLC (1 + 2). RNA helix was used for hybrid Fiber strands due to the need of RNA base interactions for the KL complex to form. Due to the fact that DNA can assume at least 3 different forms, A-form (RNA), B-form (DNA), and Z-form, it was interesting to calculate the helical-twist rotation difference between helices of various available strands. Since A-form DNA (and RNA) has 11bp rise, while B-form DNA has ~ 10.4 bp rise, one can calculate what the difference in helix rotation would be, taking into account the number of base-pairs in helices of each of the RNA fiber strands. For example, the Fiber KL1 strand (see Table 6.2, below) with an 8bp stem region available in all RNA or RNA/DNA hybrid has a calculated difference in helical-twist rotation of $\sim 15^\circ$ between the two otherwise the same strands (same sequence, different nucleotides, rNTPs vs dNTPs). Respective differences between “relative” fiber strands are highlighted in Table 6.2 below.

Table 6.2: Differences in helix rotation between various HIV strands. RNA fiber can be assembled from a combination of any Fiber KL 1 + 2 strands. Hybrid strands are not all RNA composition, but rather, RNA KL regions and DNA helical / stem region of the “dumbbell” structure. (see Figure 1.4)

all-RNA strand:	helical rotation difference (°)	hybrid strand:	all-RNA strand:	helical rotation difference (°)	hybrid strand:
FiberKL_HIV1_8bp	~15°	FiberKL_HIV1 ^h _8bp	FiberKL_HIV2_9bp	~17.6°	FiberKL_HIV2 ^h _8bp
FiberKL_HIV1_15bp	~28.3° (net)	FiberKL_HIV1 ^h _15bp	FiberKL_HIV2_15bp	~28.3° (net)	FiberKL_HIV2 ^h _15bp
FiberKL_HIV1_30bp	~56.6° (net)	FiberKL_HIV1 ^h _30bp	FiberKL_HIV2_30bp	~56.6° (net)	FiberKL_HIV2 ^h _30bp

6.3 Modifications based on hybrid strands

Having different composition Fiber strands led to ideas of what can be done with adjusting either the composition of helical nucleotides or helical length. As already mentioned, Fiber strands can be modified with a functionalized toehold (such as GFP sense or antisense), but the location/position of the toehold on the folded strand can be adjusted for further quaternary (4°) structures. For this purpose, one can calculate helical rotation for different length helices and record the overall or net rotation, testing with 11 bp and 10.4 bp helical pitch. Functionalized group is assumed to be in the middle of the helix (but whose location/position can be designed in a different location along the Fiber strand), so further calculations may be needed. With this approach, Fibers can be designed to have specific “rotational” geometry, where functionalized toehold groups would come out of fiber as predicted, and not by chance. As shown in Figure 6.4, functionalized arms/groups are coming out of fiber at different “clock arm/hand” rotations. It does appear as though from left to right of the figure the functional arms rotate clockwise as one’s view progresses down the fiber length. But

this could be controlled for specific rotational geometry by utilizing DNA or RNA composition of the helix and its length.

To calculate functional arm locations and their rotational geometry, one needs to also consider KL interactions between different Fiber monomers and the rotation that they impose on the overall helical fiber. For this, KL interaction (PDB: 2FCX) was looked at (Figure 6.5), highlighting terminal interacting KL base-pairs as space-filling structures, as they would also be terminal helical base pairs. Other interacting base-pairs or non-interacting bases were shown (or deactivated) as licorice (line) structures to clear up the view. The image shows KL interaction as one would look into the helix, as into a barrel. Rotation that the KL imposes can be deduced from that image and is about 45° counterclockwise as moving into the helix, also keeping in mind that A- and B-form DNA are right-handed, that is, rotation is clockwise as one looks into the helix. Thus, left-handed (counterclockwise, or negative) KL “rotation” needs to be considered when designing further Fiber strands, if specific rotational geometry is desired.

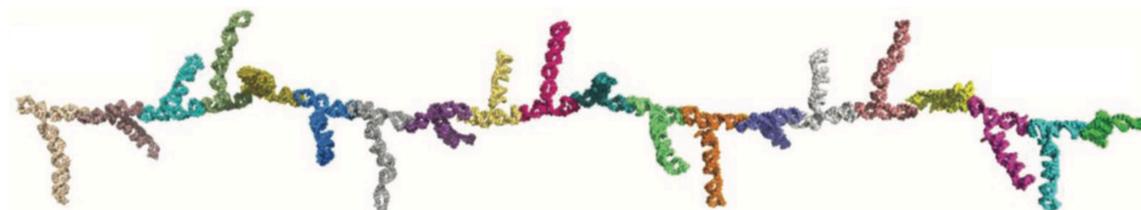


Figure 6.1: RNA fiber depicted with computer graphics, showing functionalized groups coming out of fiber at various “clock arm” rotations. Figure adapted from paper by Rackley et al.⁷⁷

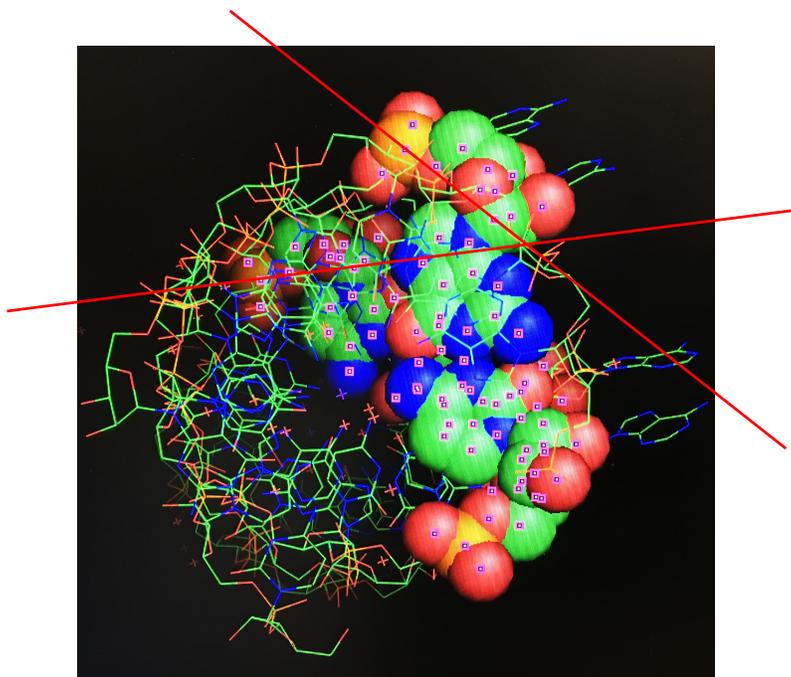


Figure 6.2: Counterclockwise rotation of the kissing loop. KL (PDB: 2FCX⁸⁷) as it appears when looking into it as one would look into a barrel. Terminal base interacting pairs are shown as space-filling model, other interacting base pairs or non-interacting bases are shown as licorice (line) structures. A red line is passed through the sugar-phosphate of the base-pair in the background as well as through the sugar-phosphate of the base-pair in the foreground. These are terminal base-pairs (**C** and **G**) of the helix before the KL motif (5'-**CAAXXXXXXAG**-3'). The rotation difference is $\sim 45^\circ$ counterclockwise, as opposed to DNA having a clockwise rotation, in this instance as one would move into the “board” / paper-plane.

6.4 SuperFibers / Fiber Bundles

As mentioned above, one of the challenges currently standing with Fiber A-B system is having only one KLC set, with the help of which more controlled fibers could be designed. After filtering initial results of crossTalk experiments four times, an average of 7 unique KL sets arise. One can design a programmable Fiber that will assemble exactly the same way and be the same length every time, expanding the A-B system into an ABCDE-type system, that is, with seven unique KL interacting sets, a fiber could be up

to eight monomers in length. Controlling and limiting the length to eight monomers may be good for some applications, except for some where a certain minimum molecular weight needs to be achieved. For this reason, shorter, controlled Fibers could be bundled together like a bundle of “twizzlers”, bringing individual fibers together either via toeholds or by incorporating Ks into functionalized “arms” on individual monomers of the RNA fibers. Of course, utilizing Ks for functional arms of individual monomers would take away from the total Ks sets utilized for length of fiber, as some would be used for bundling of fibers together. These could be fiber bundles or superFibers. Nevertheless, playing with hybrid strands as described here can yield interesting applications of how strands are designed, and which material is utilized.

CHAPTER 7: CONCLUSIONS

In regard to the use of T7 RNA polymerase for *in vitro* transcriptions, it was interesting to note that using more enzyme to transcribe RNA does not increase the yields of transcription product. Rather, an opposite was true, where lower final T7 RNA polymerase concentrations produced higher yields. Interestingly, commercial T7 RNA polymerase is available in even higher concentrations, provided the amounts suggested by manufacturers results in high final enzyme concentration per transcription reaction. Optimization of T7 RNA polymerase helped gain valuable information not to overuse the enzyme and not to oversaturate the transcription reactions.

The kissing loop library development had proved to be a much more involved process than initially perceived. Overlooking the extensiveness of the project, good prospects of utilizing the KL library are foreseen. After filtering the crosstalk experimental data 4 different times, it appears that between 6 – 8 KLC sets are unique and inter-compatible for use in larger one-pot assemblies. This will be helpful in development of precise RNA fibers for intracellular poly-heterogeneous siRNA cargo load delivery.

REFERENCES

1. Mendel, G. J. A.-I., Versuche über Pflanzenhybriden Verhandlungen des naturforschenden Vereines in Brünn, Bd. IV für das Jahr, 1865. **1866**, 3, 47.
2. Dahm, R. J. H. g., Discovering DNA: Friedrich Miescher and the early years of nucleic acid research. **2008**, 122 (6), 565-581.
3. Morgan, T. H. J. T. A. N., Chromosomes and heredity. **1910**, 44 (524), 449-496.
4. Wilhelm, J. J. t. I. O. i. S. N., Elemente der exakten Erblchkeitslehre. **1909**.
5. Avery, O. T.; MacLeod, C. M.; McCarty, M. J. T. J. o. e. m., Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. **1944**, 79 (2), 137-158.
6. Chargaff, E. J. E., Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. **1950**, 6 (6), 201-209.
7. Watson, J. D.; Crick, F. H. J. N., Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. **1953**, 171 (4356), 737-738.
8. Leontis, N. B.; Stombaugh, J.; Westhof, E. J. N. a. r., The non-Watson–Crick base pairs and their associated isostericity matrices. **2002**, 30 (16), 3497-3531.
9. Seeman, N. C. J. J. o. t. b., Nucleic acid junctions and lattices. **1982**, 99 (2), 237-247.
10. Grabow, W. W.; Zakrevsky, P.; Afonin, K. A.; Chworos, A.; Shapiro, B. A.; Jaeger, L. J. N. I., Self-assembling RNA nanorings based on RNAI/II inverse kissing complexes. **2011**, 11 (2), 878-887.
11. Afonin, K. A.; Kasprzak, W.; Bindewald, E.; Puppala, P. S.; Diehl, A. R.; Hall, K. T.; Kim, T. J.; Zimmermann, M. T.; Jernigan, R. L.; Jaeger, L. J. M., Computational and experimental characterization of RNA cubic nanoscaffolds. **2014**, 67 (2), 256-265.
12. Chandler, M.; Lyalina, T.; Halman, J.; Rackley, L.; Lee, L.; Dang, D.; Ke, W.; Sajja, S.; Woods, S.; Acharya, S. J. M., Broccoli fluorets: split aptamers as a user-friendly fluorescent toolkit for dynamic RNA Nanotechnology. **2018**, 23 (12), 3178.

13. Gwinn, E. G.; O'Neill, P.; Guerrero, A. J.; Bouwmeester, D.; Fygenson, D. K. J. A. M., Sequence-Dependent fluorescence of DNA-Hosted silver nanoclusters. **2008**, *20* (2), 279-283.
14. Yourston, L. E.; Lushnikov, A. Y.; Shevchenko, O. A.; Afonin, K. A.; Krasnoslobodtsev, A. V. J. N., First Step Towards Larger DNA-Based Assemblies of Fluorescent Silver Nanoclusters: Template Design and Detailed Characterization of Optical Properties. **2019**, *9* (4), 613.
15. Chandler, M.; Shevchenko, O.; Vivero-Escoto, J. L.; Striplin, C. D.; Afonin, K. A. J. J. o. C. E., DNA-Templated Synthesis of Fluorescent Silver Nanoclusters. **2020**.
16. Saenger, W., Defining terms for the nucleic acids. In *Principles of nucleic acid structure*, Springer: 1984; pp 9-28.
17. Saenger, W. J. N. Y., Berlin, Heidelberg, Principles of Nucleic Acid Structure. Springer. **1984**.
18. Saenger, W., Why Study Nucleotide and Nucleic Acid Structure? In *Principles of Nucleic Acid Structure*, Springer: 1984; pp 1-8.
19. Klein, A.; Bonhoeffer, F. J. A. R. o. B., DNA replication. **1972**, *41* (1), 301-332.
20. Saenger, W., Structures and conformational properties of bases, furanose sugars, and phosphate groups. In *Principles of nucleic acid structure*, Springer: 1984; pp 51-104.
21. Cushley, R. J.; Codrington, J. F.; Fox, J. J. J. C. J. o. C., Nucleosides. XLIX. Nuclear magnetic resonance studies of 2'-and 3'-halogeno nucleosides. The conformations of 2'-deoxy-2'-fluorouridine and 3'-deoxy-3'-fluoro- β -d-arabinofuranosyluracil. **1968**, *46* (7), 1131-1140.
22. Ekiel, I.; Darzynkiewicz, E.; Dudycz, L.; Shugar, D. J. B., Solution conformation and relative acidities of the sugar hydroxyls of the O'-methylated derivatives of the antimetabolite 9- β -D-xylofuranosyladenine. **1978**, *17* (8), 1530-1536.
23. IMAZAWA, M.; UEDA, T.; UKITA, T. J. C.; Bulletin, P., Nucleosides and nucleotides. XII. Synthesis and properties of 2'-deoxy-2'-mercaptouridine and its derivatives. **1975**, *23* (3), 604-610.
24. DE LEEUW, H. P.; DE JAGER, J. R.; KOENERS, H. J.; VAN BOOM, J. H.; ALTONA, C. J. E. j. o. b., Puromycin and Some of Its Analogues: Conformational Properties in Solution: A 360-MHz Nuclear-Magnetic-Resonance Investigation. **1977**, *76* (1), 209-217.

25. PLACH, H.; WESTHOF, E.; LÜDEMANN, H. D.; MENGEL, R. J. E. j. o. b., Solution Conformational Analysis of 2'-Amino-2'-Deoxyadenosine, 3'-Amino-3'-Deoxyadenosine and Puromycin by Pulsed Nuclear-Magnetic-Resonance Methods. **1977**, *80* (1), 295-304.
26. Sundaralingam, M. J. S.; Acids, C. o. N.; Interactions, P.-N. A., Principles governing nucleic acid and polynucleotide conformations. **1975**, 487-524.
27. Saenger, W., Polymorphism of DNA versus structural conservatism of RNA: Classification of A-, B-, and Z-type double helices. In *Principles of Nucleic Acid Structure*, Springer: 1984; pp 220-241.
28. Saenger, W., Forces stabilizing associations between bases: hydrogen bonding and base stacking. In *Principles of nucleic acid structure*, Springer: 1984; pp 116-158.
29. Saenger, W., Physical properties of nucleotides: charge densities, pK values, spectra, and tautomerism. In *Principles of nucleic acid structure*, Springer: 1984; pp 105-115.
30. Clauwaert, J.; Stockx, J. J. Z. f. N. B., Interactions of polynucleotides and their components. **1968**, *23* (1), 25-30.
31. Fasman, G. D., Handbook of biochemistry and molecular biology. Nucleic acids-v. 1-3. **1975**.
32. Sober, H. J. B.-. Handbook of Biochemistry (Chemical Rubber Co., Cleveland, Ohio), pp. **1970**.
33. Angell, C. J. J. o. t. C. S., 100. An infrared spectroscopic investigation of nucleic acid constituents. **1961**, 504-515.
34. Tsuboi, M.; Kyogoku, Y.; Shimanouchi, T. J. B. e. B. A.-S. S. o. N. A.; Subjects, R., Infrared absorption spectra of protonated and deprotonated nucleosides. **1962**, *55* (1-2), 1-12.
35. Lord, c.; Thomas Jr, G. J. S. A. P. A. M. S., Raman spectral studies of nucleic acids and related molecules—I Ribonucleic acid derivatives. **1967**, *23* (9), 2551-2591.
36. Fox, J. J.; Shugar, D. J. B. e. b. a., Spectrophotometric studies on nucleic acid derivatives and related compounds as a function of pH. II. Natural and synthetic pyrimidine nucleosides. **1952**, *9* (C), 369-384.
37. Jardetzky, C. D.; Jardetzky, O. J. J. o. t. A. C. S., Investigation of the Structure of Purines, Pyrimidines, Ribose Nucleosides and Nucleotides by Proton Magnetic Resonance. II1. **1960**, *82* (1), 222-229.

38. Katritzky, A.; Waring, A. J. J. o. t. C. S., 299. Tautomeric azines. Part I. The tautomerism of 1-methyluracil and 5-bromo-1-methyluracil. **1962**, 1540-1544.
39. Poulter, C. D.; Anderson, R. B. J. T. L., Direct observation of uracil dication and related derivatives. **1972**, 13 (36), 3823-3826.
40. Wagner, R.; Von Philipsborn, W. J. H. c. a., Protonierung von Amino- und Hydroxypyrimidinen NMR-Spektren und Strukturen der Mono- und Dikationen. **1970**, 53 (2), 299-320.
41. Phillips, R.; Eisenberg, S. P.; George, P.; Rutman, R. J. J. o. B. C., Thermodynamic data for the secondary phosphate ionizations of adenosine, guanosine, inosine, cytidine, and uridine nucleotides and triphosphate. **1965**, 240 (11), 4393-4397.
42. Jonas, J.; Gut, J. J. C. o. C. C. C., Nucleic acid components and their analogues. XVI. Dissociation constants of uracil, 6-azauracil, 5-azauracil and related compounds. **1962**, 27 (3), 716-723.
43. Izatt, R.; Hansen, L.; Rytting, J.; Christensen, J. J. J. o. t. A. C. S., Proton ionization from adenosine. **1965**, 87 (12), 2760-2761.
44. Birnbaum, G. I.; Giziewicz, J.; Huber, C. P.; Shugar, D. J. J. o. t. A. C. S., Intramolecular hydrogen bonding and acidities of nucleoside sugar hydroxyls. Crystal structure and conformation of O², 2'-anhydro-1- α -D-xylofuranosyluracil. **1976**, 98 (15), 4640-4644.
45. Voet, D.; Rich, A., The crystal structures of purines, pyrimidines and their intermolecular complexes. In *Progress in nucleic acid research and molecular biology*, Elsevier: 1970; Vol. 10, pp 183-265.
46. Pauling, L., *The Nature of the Chemical Bond*. Cornell university press Ithaca, NY: 1960; Vol. 260.
47. Vinogradov, S. N.; Linnell, R. H., *Hydrogen bonding*. Van Nostrand Reinhold: 1971.
48. Jeffrey, G. A.; Takagi, S. J. A. o. C. R., Hydrogen-bond structure in carbohydrate crystals. **1978**, 11 (7), 264-270.
49. Kroon, J.; Kanters, J.; van Duijneveldt-van De Rijdt, J.; Van Duijneveldt, F.; Vliegthart, J. J. J. o. M. S., OH \cdots O Hydrogen bonds in molecular crystals a statistical and quantum-chemical analysis. **1975**, 24 (1), 109-129.

50. Bugg, C.; Thomas, J.; Sundaralingam, M.; Rao, S. J. B. O. R. o. B., Stereochemistry of nucleic acids and their constituents. X. solid-slate base-slacking patterns in nucleic acid constituents and polynucleotides. **1971**, *10* (1), 175-219.
51. Broom, A. D.; Schweizer, M. P.; Ts'o, P. O. J. J. o. t. A. C. S., Interaction and association of bases and nucleosides in aqueous solutions. V. Studies of the association of purine nucleosides by vapor pressure osmometry and by proton magnetic resonance. **1967**, *89* (14), 3612-3622.
52. Ts'o, P. O.; Kondo, N. S.; Robins, R. K.; Broom, A. D. J. J. o. t. A. C. S., Interaction and association of bases and nucleosides in aqueous solutions. VI. Properties of 7-methylinosine as related to the nature of the stacking interaction. **1969**, *91* (20), 5625-5631.
53. Ts'o, P. O. J. B. p. i. n. a. c., Bases, nucleosides, and nucleotides. **1974**, *1*, 453-584.
54. Gratzer, W. J. E. j. o. b., Association of Nucleic-Acid Bases in Aqueous Solution: a Solvent Partition Study. **1969**, *10* (1), 184-187.
55. Solie, T. t.; Schellman, J. J. J. o. m. b., The interaction of nucleosides in aqueous solution. **1968**, *33* (1), 61-77.
56. MITCHELL, P. R.; SIGEL, H. J. E. j. o. b., A proton nuclear-magnetic-resonance study of self-stacking in purine and pyrimidine nucleosides and nucleotides. **1978**, *88* (1), 149-154.
57. Topal, M. D.; Warshaw, M. M. J. B. O. R. o. B., Dinucleoside monophosphates. II. Nearest neighbor interactions. **1976**, *15* (9), 1775-1793.
58. Davies, D. B., Co-operative conformational properties of nucleosides, nucleotides and nucleotidyl units in solution. In *Nuclear Magnetic Resonance Spectroscopy in Molecular Biology*, Springer: 1978; pp 71-85.
59. Pörschke, D., Elementary steps of base recognition and helix-coil transitions in nucleic acids. In *Chemical relaxation in molecular biology*, Springer: 1977; pp 191-218.
60. Cantor, C. R.; Tinoco Jr, I. J. J. o. m. b., Absorption and optical rotatory dispersion of seven trinucleoside diphosphates. **1965**, *13* (1), 65-77.
61. Ackerman, M. M.; Ricciardi, C.; Weiss, D.; Chant, A.; Kraemer-Chant, C. M. J. J. o. C. E., Analyzing Exonuclease-Induced Hyperchromicity by UV Spectroscopy: An Undergraduate Biochemistry Laboratory Experiment. **2016**, *93* (12), 2089-2095.

62. Putnam, B.; Van Zandt, L.; Prohofsky, E.; Mei, W. J. B. j., Resonant and localized breathing modes in terminal regions of the DNA double helix. **1981**, *35* (2), 271-287.
63. Altan-Bonnet, G.; Libchaber, A.; Krichevsky, O. J. P. r. l., Bubble dynamics in double-stranded DNA. **2003**, *90* (13), 138101.
64. Watson, J. D. J. S., The human genome project: past, present, and future. **1990**, *248* (4951), 44-49.
65. Institute, N. H. G. R. The Human Genome Project.
<https://www.genome.gov/human-genome-project>
.
66. Feitelson, D. G.; Treinin, M. J. C., The blueprint for life? **2002**, *35* (7), 34-40.
67. Phillips, T. J. N. E., Regulation of transcription and gene expression in eukaryotes. **2008**, *1* (1), 199.
68. Costa, F. F. J. B., Non-coding RNAs: meet thy masters. **2010**, *32* (7), 599-608.
69. Ling, H.; Vincent, K.; Pichler, M.; Fodde, R.; Berindan-Neagoe, I.; Slack, F. J.; Calin, G. A. J. O., Junk DNA and the long non-coding RNA twist in cancer genetics. **2015**, *34* (39), 5003-5011.
70. Djebali, S.; Davis, C. A.; Merkel, A.; Dobin, A.; Lassmann, T.; Mortazavi, A.; Tanzer, A.; Lagarde, J.; Lin, W.; Schlesinger, F. J. N., Landscape of transcription in human cells. **2012**, *489* (7414), 101-108.
71. Pheasant, M.; Mattick, J. S. J. G. r., Raising the estimate of functional human sequences. **2007**, *17* (9), 1245-1253.
72. Carninci, P.; Kasukawa, T.; Katayama, S.; Gough, J.; Frith, M.; Maeda, N.; Oyama, R.; Ravasi, T.; Lenhard, B.; Wells, C. J. s., The transcriptional landscape of the mammalian genome. **2005**, *309* (5740), 1559-1563.
73. Comings, D. E., The structure and function of chromatin. In *Advances in human genetics*, Springer: 1972; pp 237-431.
74. Fire, A.; Xu, S.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E.; Mello, C. C. J. n., Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. **1998**, *391* (6669), 806-811.
75. Fire, A.; Mello, C., RNA Interference—Gene Silencing by Double-stranded RNA. **2006**.

76. Cullen, L. M.; Arndt, G. M. J. I.; biology, c., Genome-wide screening for gene function using RNAi in mammalian cells. **2005**, *83* (3), 217-223.
77. Rackley, L.; Stewart, J. M.; Salotti, J.; Krokhotin, A.; Shah, A.; Halman, J. R.; Juneja, R.; Smollett, J.; Lee, L.; Roark, K. J. A. f. m., RNA fibers as optimized nanoscaffolds for siRNA coordination and reduced immunological recognition. **2018**, *28* (48), 1805959.
78. Halman, J. R.; Satterwhite, E.; Roark, B.; Chandler, M.; Viard, M.; Ivanina, A.; Bindewald, E.; Kasprzak, W. K.; Panigaj, M.; Bui, M. N. J. N. a. r., Functionally-interdependent shape-switching nanoparticles with controllable properties. **2017**, *45* (4), 2210-2220.
79. Ke, W.; Hong, E.; Saito, R. F.; Rangel, M. C.; Wang, J.; Viard, M.; Richardson, M.; Khisamutdinov, E. F.; Panigaj, M.; Dokholyan, N. V. J. N. a. r., RNA–DNA fibers and polygons with controlled immunorecognition activate RNAi, FRET and transcriptional regulation of NF- κ B in human cells. **2019**, *47* (3), 1350-1361.
80. Dao, B. N.; Viard, M.; Martins, A. N.; Kasprzak, W. K.; Shapiro, B. A.; Afonin, K. A. J. D.; nanotechnology, R., Triggering RNAi with multifunctional RNA nanoparticles and their delivery. **2015**, *1* (open-issue), 1-12.
81. Afonin, K. A.; Viard, M.; Kagiampakis, I.; Case, C. L.; Dobrovolskaia, M. A.; Hofmann, J.; Vrzak, A.; Kireeva, M.; Kasprzak, W. K.; KewalRamani, V. N. J. A. n., Triggering of RNA Interference with RNA–RNA, RNA–DNA, and DNA–RNA Nanoparticles. **2015**, *9* (1), 251-259.
82. Afonin, K. A.; Bindewald, E.; Yaghoubian, A. J.; Voss, N.; Jacovetty, E.; Shapiro, B. A.; Jaeger, L. J. N. n., In vitro assembly of cubic RNA-based scaffolds designed in silico. **2010**, *5* (9), 676-682.
83. Li, M.; Zheng, M.; Wu, S.; Tian, C.; Liu, D.; Weizmann, Y.; Jiang, W.; Wang, G.; Mao, C. J. N. c., In vivo production of RNA nanostructures via programmed folding of single-stranded RNAs. **2018**, *9* (1), 1-9.
84. Liu, D.; Geary, C. W.; Chen, G.; Shao, Y.; Li, M.; Mao, C.; Andersen, E. S.; Piccirilli, J. A.; Rothmund, P. W.; Weizmann, Y. J. N. C., Branched kissing loops for the construction of diverse RNA homooligomeric nanostructures. **2020**, *12* (3), 249-259.
85. Grabow, W. W.; Jaeger, L. J. A. o. c. r., RNA self-assembly and RNA nanotechnology. **2014**, *47* (6), 1871-1880.

86. Paillart, J.-C.; Westhof, E.; Ehresmann, C.; Ehresmann, B.; Marquet, R. J. J. o. m. b., Non-canonical interactions in a kissing loop complex: the dimerization initiation site of HIV-1 genomic RNA. **1997**, *270* (1), 36-49.
87. Ennifar, E.; Paillart, J.-C.; Bodlenner, A.; Walter, P.; Weibel, J.-M.; Aubertin, A.-M.; Pale, P.; Dumas, P.; Marquet, R. J. N. a. r., Targeting the dimerization initiation site of HIV-1 RNA with aminoglycosides: from crystal to cell. **2006**, *34* (8), 2328-2339.
88. Ennifar, E.; Walter, P.; Ehresmann, B.; Ehresmann, C.; Dumas, P. J. N. s. b., Crystal structures of coaxially stacked kissing complexes of the HIV-1 RNA dimerization initiation site. **2001**, *8* (12), 1064-1068.
89. Kulisch, S. E., Teresa; Whitman-Guliaev, Christina; Rubio, Teresa., Dicer-Substrate siRNA Technology.
90. Halman, J. R.; Kim, K.-T.; Gwak, S.-J.; Pace, R.; Johnson, M. B.; Chandler, M. R.; Rackley, L.; Viard, M.; Marriott, I.; Lee, J. S. J. N. N., Biology; Medicine, A cationic amphiphilic co-polymer as a carrier of nucleic acid nanoparticles (Nanps) for controlled gene silencing, immunostimulation, and biodistribution. **2020**, *23*, 102094.
91. Cruz-Acuña, M.; Halman, J. R.; Afonin, K. A.; Dobson, J.; Rinaldi, C. J. N., Magnetic nanoparticles loaded with functional RNA nanoparticles. **2018**, *10* (37), 17761-17770.
92. Stewart, J. M.; Viard, M.; Subramanian, H. K.; Roark, B. K.; Afonin, K. A.; Franco, E. J. N., Programmable RNA microstructures for coordinated delivery of siRNAs. **2016**, *8* (40), 17542-17550.
93. Afonin, K. A.; Viard, M.; Tedbury, P.; Bindewald, E.; Parlea, L.; Howington, M.; Valdman, M.; Johns-Boehme, A.; Brainerd, C.; Freed, E. O. J. N. I., The use of minimal RNA toeholds to trigger the activation of multiple functionalities. **2016**, *16* (3), 1746-1753.
94. Afonin, K. A.; Viard, M.; Koyfman, A. Y.; Martins, A. N.; Kasprzak, W. K.; Panigaj, M.; Desai, R.; Santhanam, A.; Grabow, W. W.; Jaeger, L. J. N. I., Multifunctional RNA nanoparticles. **2014**, *14* (10), 5662-5671.
95. Afonin, K. A.; Kireeva, M.; Grabow, W. W.; Kashlev, M.; Jaeger, L.; Shapiro, B. A. J. N. I., Co-transcriptional assembly of chemically modified RNA nanoparticles functionalized with siRNAs. **2012**, *12* (10), 5192-5195.
96. Erkekoglu, P.; Baydar, T. J. N. n., Acrylamide neurotoxicity. **2014**, *17* (2), 49-57.

97. Cheetham, G. M.; Jeruzalmi, D.; Steitz, T. A. J. N., Structural basis for initiation of transcription from an RNA polymerase–promoter complex. **1999**, *399* (6731), 80-83.
98. Sousa, R.; Chung, Y. J.; Rose, J. P.; Wang, B.-C. J. N., Crystal structure of bacteriophage T7 RNA polymerase at 3.3 Å resolution. **1993**, *364* (6438), 593-599.
99. Afonin, K. A.; Viard, M.; Martins, A. N.; Lockett, S. J.; Maciag, A. E.; Freed, E. O.; Heldman, E.; Jaeger, L.; Blumenthal, R.; Shapiro, B. A. J. N. n., Activation of different split functionalities on re-association of RNA–DNA hybrids. **2013**, *8* (4), 296-304.
100. Kim, T.; Afonin, K. A.; Viard, M.; Koyfman, A. Y.; Sparks, S.; Heldman, E.; Grinberg, S.; Linder, C.; Blumenthal, R. P.; Shapiro, B. A. J. M. T.-N. A., In silico, in vitro, and in vivo studies indicate the potential use of bolaamphiphiles for therapeutic siRNAs delivery. **2013**, *2*, e80.
101. Afonin, K. A.; Desai, R.; Viard, M.; Kireeva, M. L.; Bindewald, E.; Case, C. L.; Maciag, A. E.; Kasprzak, W. K.; Kim, T.; Sappe, A. J. N. a. r., Co-transcriptional production of RNA–DNA hybrids for simultaneous release of multiple split functionalities. **2014**, *42* (3), 2085-2097.
102. Gupta, K.; Afonin, K. A.; Viard, M.; Herrero, V.; Kasprzak, W.; Kagiampakis, I.; Kim, T.; Koyfman, A. Y.; Puri, A.; Stepler, M. J. J. o. C. R., Bolaamphiphiles as carriers for siRNA delivery: From chemical syntheses to practical applications. **2015**, *213*, 142-151.
103. Gupta, K.; Mattingly, S. J.; Knipp, R. J.; Afonin, K. A.; Viard, M.; Bergman, J. T.; Stepler, M.; Nantz, M. H.; Puri, A.; Shapiro, B. A. J. N., Oxime ether lipids containing hydroxylated head groups are more superior siRNA delivery agents than their nonhydroxylated counterparts. **2015**, *10* (18), 2805-2818.
104. Juneja, R.; Lyles, Z.; Vadarevu, H.; Afonin, K. A.; Vivero-Escoto, J. L. J. A. a. m.; interfaces, Multimodal polysilsesquioxane nanoparticles for combinatorial therapy and gene delivery in triple-negative breast cancer. **2019**, *11* (13), 12308-12320.
105. Kim, T.; Viard, M.; Afonin, K. A.; Gupta, K.; Popov, M.; Salotti, J.; Johnson, P. F.; Linder, C.; Heldman, E.; Shapiro, B. A. J. M. T.-N. A., Characterization of Cationic Bolaamphiphile Vesicles for siRNA Delivery into Tumors and Brain. **2020**.
106. Zadeh, J. N.; Steenberg, C. D.; Bois, J. S.; Wolfe, B. R.; Pierce, M. B.; Khan, A. R.; Dirks, R. M.; Pierce, N. A. J. J. o. c. c., NUPACK: analysis and design of nucleic acid systems. **2011**, *32* (1), 170-173.
107. Swinehart, D. F. J. J. o. c. e., The beer-lambert law. **1962**, *39* (7), 333.

108. Quillardet, P.; Hofnung, M. J. T. i. G., Ethidium bromide and safety—readers suggest alternative solutions. **1988**, 4 (4), 89-90.
109. Afonin, K. A.; Grabow, W. W.; Walker, F. M.; Bindewald, E.; Dobrovolskaia, M. A.; Shapiro, B. A.; Jaeger, L. J. N. p., Design and self-assembly of siRNA-functionalized RNA nanoparticles for use in automated nanomedicine. **2011**, 6 (12), 2022.
110. Surratt, C. K.; Milan, S. C.; Chamberlin, M. J. J. P. o. t. N. A. o. S., Spontaneous cleavage of RNA in ternary complexes of Escherichia coli RNA polymerase and its significance for the mechanism of transcription. **1991**, 88 (18), 7983-7987.
111. Binzel, D. W.; Khisamutdinov, E. F.; Guo, P. J. B., Entropy-driven one-step formation of Phi29 pRNA 3WJ from three RNA fragments. **2014**, 53 (14), 2221-2231.

APPENDIX: SUPPLEMENTARY DETAILS

A.1 Hairpin KL full sequences

Table A.1: Full-length sequences of KL hairpins and displacement strands. All KL hairpin sequences are identical for the exception of the active mutated KL regions. Color coding is to help track details of complementarity and possible strand self-interactions.

Sequence Description	Sequence	GC %	T _m °C
GFPasDisplace27_5'A488	(27nt) 5'- /5ATTO488N/ Cgg Tgg TgC AgA TgA ACT TCA ggg TCA -3'	55.6	64.1
GFPasDisplace27	(27nt) 5'- Cgg Tgg TgC AgA TgA ACT TCA ggg TCA -3'	55.6	64.1
KL_0_GFPs	(54nt) 5'- GGGAAUCCAAAAAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	46.3	65.8
KL_1_GFPs	(54nt) 5'- GGGAAUCCAAAGGAGCAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	70
KL_2_GFPs	(54nt) 5'- GGGAAUCCAAAGCCUCCAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	68.8
KL_1'_GFPs	(54nt) 5'- GGGAAUCCAAAGGAGGAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	69.5
KL_2'_GFPs	(54nt) 5'- GGGAAUCCAAAGCCUCCAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	68.3
KL_3_GFPs	(54nt) 5'- GGGAAUCCAAAGCGGACAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	69.2
KL_4_GFPs	(54nt) 5'- GGGAAUCCAAAGCCGCAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	67.7
KL_3'_GFPs	(54nt) 5'- GGGAAUCCAAAGCGGCAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	69
KL_4'_GFPs	(54nt) 5'- GGGAAUCCAAAGCCUCCAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	68.8
KL_5_GFPs	(54nt) 5'- GGGAAUCCAAAGCGCUGAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	68.9
KL_6_GFPs	(54nt) 5'- GGGAAUCCAAAGCGGCAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	68.8
KL_5'_GFPs	(54nt) 5'- GGGAAUCCAAAGCGCGAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	67.3
KL_6'_GFPs	(54nt) 5'- GGGAAUCCAAAGCGGACAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	68.4
KL_7_GFPs	(54nt) 5'- GGGAAUCCAAAGCCAGGAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	69.7
KL_8_GFPs	(54nt) 5'- GGGAAUCCAAAGCCUGCAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	69.5
KL_7'_GFPs	(54nt) 5'- GGGAAUCCAAAGACCCGAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	69.3
KL_8'_GFPs	(54nt) 5'- GGGAAUCCAAAGCCUCCAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	67.9
KL_9_GFPs	(54nt) 5'- GGGAAUCCAAAGCGGGAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	68.8
KL_10_GFPs	(54nt) 5'- GGGAAUCCAAAGCCGCAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	68.6
KL_9'_GFPs	(54nt) 5'- GGGAAUCCAAAGGGCGUAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	68.5
KL_10'_GFPs	(54nt) 5'- GGGAAUCCAAAGCCGCAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	68.6
KL_11_GFPs	(54nt) 5'- GGGAAUCCAAAGCGGGAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	68.5
KL_12_GFPs	(54nt) 5'- GGGAAUCCAAAGCCGGAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	68.7
KL_11'_GFPs	(54nt) 5'- GGGAAUCCAAAGGGGCUAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	69.8
KL_12'_GFPs	(54nt) 5'- GGGAAUCCAAAGCCGCAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	69.2
KL_13_GFPs	(54nt) 5'- GGGAAUCCAAAGCCGCAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	68.7
KL_14_GFPs	(54nt) 5'- GGGAAUCCAAAGGGGCAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	69.9
KL_13'_GFPs	(54nt) 5'- GGGAAUCCAAAGCCGCUAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	68
KL_14'_GFPs	(54nt) 5'- GGGAAUCCAAAGCGGGAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	69.5
KL_15_GFPs	(54nt) 5'- GGGAAUCCAAAGCCGCAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	67.7
KL_16_GFPs	(54nt) 5'- GGGAAUCCAAAGGGCGAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	69.3
KL_15'_GFPs	(54nt) 5'- GGGAAUCCAAAGCCGCUAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	68.5
KL_16'_GFPs	(54nt) 5'- GGGAAUCCAAAGCGGGAAGGAUUCUUUGACCCUGAAGUUCACUCGCACCACCG -3'	55.6	69.3

A.2 Recipes for general laboratory buffer preparation

Table A.2: Recipes for various laboratory buffers and solutions. Amounts of different reagents to add to each mixture are in red.

10X TB (10x Tris-borate) buffer				
	MW (g/mol)	amount (g)	C2 (mM)	V2 (mL)
Tris	121.14	107.81	890	1000
Boric acid	61.83	55.03		
ddiH ₂ O up to 1 L final				
pH adjust to 8.2 with 2 N HCl				
filter solution				

8M Urea				
	MW (g/mol)	amount (g)	C2 (M)	V2 (mL)
Urea	60.06	960.96	8.00	2000
ddiH ₂ O up to 2 L final				
filter solution				

10X TBE (10x Tris-borate-EDTA) buffer				
	MW (g/mol)	amount (g)	C2 (mM)	V2 (mL)
Tris	121.14	107.81	890	1000
Boric acid	61.83	55.03		
EDTA	292.24	5.84	20	
ddiH ₂ O up to 1 L final				
pH adjust to 8.2 with 2 N HCl				
filter solution				

20% A:bA / 8M Urea					
	MW (g/mol)	amount (g)	C2 (M)	V2 (mL)	
Urea	60.06	960.96	8.00	2000	
		C1 (%)	V1 (mL)	C2 (%)	2000
A:bA (19:1)	40	1000	20		
ddiH ₂ O up to 1.8 L, stir & heat at 80°C for 2 - 4 h					
once dissolved, add ddiH ₂ O up to 2 L					
filter solution					

5M NaCl				
	MW (g/mol)	amount (g)	C2 (M)	V2 (mL)
NaCl	58.44	58.44	5.00	200
ddiH ₂ O				
filter solution				

Crush and Soak buffer (300 mM Na ⁺)				
	C1	amount	C2	V2 (mL)
5M NaCl	5M	12 mL	300 mM	200
10x TBE	10x	20 mL	1x	
ddiH ₂ O				

1M MgCl ₂				
	MW (g/mol)	amount (g)	C2 (M)	V2 (mL)
MgCl ₂	95.21	23.80	1.00	250
ddiH ₂ O				
filter solution				

1M KCl				
	MW (g/mol)	amount (g)	C2 (M)	V2 (mL)
KCl	74.55	11.18	1.00	150
ddiH ₂ O				
filter solution				

A.3 Alternative view of crossTalk results matrices

Table A.3.1: Results from crossTalk #1 experiment; alternative view. Each row represents one gel from the experiment. Gel number can be seen in the very left column. Nomenclature is greatly simplified, where each hairpin structure fully named KL_1_GFPs is simplified to **1**, respectively. Additionally, each hairpin structure is also bound with a GFPasDisplaceStrand. See Figure 5.3 in main text for understanding color coding. This Table represents the same data as **Table 5.3** in the main text.

cT1		1	1'	2	2'	3	3'	4	4'	5	5'	6	6'	7	7'	8	8'
34		1	1'	2	2'	3	3'	4	4'	5	5'	6	6'	7	7'	8	8'
35		9	9'	10	10'	11	11'	12	12'	13	13'	14	14'	15	15'	16	16'
36	1		1+2	1+3	1+4	1+5	1+6	1+7	1+8	1+9	1+10	1+11	1+12	1+13	1+14	1+15	1+16
37	1'	1'+1	1'+2'	1'+3'	1'+4'	1'+5'	1'+6'	1'+7'	1'+8'	1'+9'	1'+10'	1'+11'	1'+12'	1'+13'	1'+14'	1'+15'	1'+16'
38	2	2+1'		2+3	2+4	2+5	2+6	2+7	2+8	2+9	2+10	2+11	2+12	2+13	2+14	2+15	2+16
39	2'	2'+1	2'+2	2'+3'	2'+4'	2'+5'	2'+6'	2'+7'	2'+8'	2'+9'	2'+10'	2'+11'	2'+12'	2'+13'	2'+14'	2'+15'	2'+16'
40	3	3+1'	3+2'		3+4	3+5	3+6	3+7	3+8	3+9	3+10	3+11	3+12	3+13	3+14	3+15	3+16
41	3'	3'+1	3'+2	3'+3	3'+4'	3'+5'	3'+6'	3'+7'	3'+8'	3'+9'	3'+10'	3'+11'	3'+12'	3'+13'	3'+14'	3'+15'	3'+16'
42	4	4+1'	4+2'	4+3'		4+5	4+6	4+7	4+8	4+9	4+10	4+11	4+12	4+13	4+14	4+15	4+16
43	4'	4'+1	4'+2	4'+3'	4'+4'	4'+5'	4'+6'	4'+7'	4'+8'	4'+9'	4'+10'	4'+11'	4'+12'	4'+13'	4'+14'	4'+15'	4'+16'
44	5	5+1'	5+2'	5+3'	5+4'		5+6	5+7	5+8	5+9	5+10	5+11	5+12	5+13	5+14	5+15	5+16
45	5'	5'+1	5'+2	5'+3'	5'+4'	5'+5'	5'+6'	5'+7'	5'+8'	5'+9'	5'+10'	5'+11'	5'+12'	5'+13'	5'+14'	5'+15'	5'+16'
46	6	6+1'	6+2'	6+3'	6+4'	6+5'		6+7	6+8	6+9	6+10	6+11	6+12	6+13	6+14	6+15	6+16
47	6'	6'+1	6'+2	6'+3'	6'+4'	6'+5'	6'+6'	6'+7'	6'+8'	6'+9'	6'+10'	6'+11'	6'+12'	6'+13'	6'+14'	6'+15'	6'+16'
48	7	7+1'	7+2'	7+3'	7+4'	7+5'	7+6'		7+8	7+9	7+10	7+11	7+12	7+13	7+14	7+15	7+16
49	7'	7'+1	7'+2	7'+3'	7'+4'	7'+5'	7'+6'	7'+7'	7'+8'	7'+9'	7'+10'	7'+11'	7'+12'	7'+13'	7'+14'	7'+15'	7'+16'
50	8	8+1'	8+2'	8+3'	8+4'	8+5'	8+6'	8+7'		8+9	8+10	8+11	8+12	8+13	8+14	8+15	8+16
51	8'	8'+1	8'+2	8'+3'	8'+4'	8'+5'	8'+6'	8'+7'	8'+8'	8'+9'	8'+10'	8'+11'	8'+12'	8'+13'	8'+14'	8'+15'	8'+16'
52	9	9+1'	9+2'	9+3'	9+4'	9+5'	9+6'	9+7'	9+8'		9+10	9+11	9+12	9+13	9+14	9+15	9+16
53	9'	9'+1	9'+2	9'+3'	9'+4'	9'+5'	9'+6'	9'+7'	9'+8'	9'+9'	9'+10'	9'+11'	9'+12'	9'+13'	9'+14'	9'+15'	9'+16'
54	10	10+1'	10+2'	10+3'	10+4'	10+5'	10+6'	10+7'	10+8'	10+9'		10+11	10+12	10+13	10+14	10+15	10+16
55	10'	10'+1	10'+2	10'+3'	10'+4'	10'+5'	10'+6'	10'+7'	10'+8'	10'+9'	10'+10'	10'+11'	10'+12'	10'+13'	10'+14'	10'+15'	10'+16'
56	11	11+1'	11+2'	11+3'	11+4'	11+5'	11+6'	11+7'	11+8'	11+9'	11+10'		11+12	11+13	11+14	11+15	11+16
57	11'	11'+1	11'+2	11'+3'	11'+4'	11'+5'	11'+6'	11'+7'	11'+8'	11'+9'	11'+10'	11'+11'	11'+12'	11'+13'	11'+14'	11'+15'	11'+16'
58	12	12+1'	12+2'	12+3'	12+4'	12+5'	12+6'	12+7'	12+8'	12+9'	12+10'	12+11'		12+13	12+14	12+15	12+16
59	12'	12'+1	12'+2	12'+3'	12'+4'	12'+5'	12'+6'	12'+7'	12'+8'	12'+9'	12'+10'	12'+11'	12'+12'	12'+13'	12'+14'	12'+15'	12'+16'
60	13	13+1'	13+2'	13+3'	13+4'	13+5'	13+6'	13+7'	13+8'	13+9'	13+10'	13+11'	13+12'		13+14	13+15	13+16
61	13'	13'+1	13'+2	13'+3'	13'+4'	13'+5'	13'+6'	13'+7'	13'+8'	13'+9'	13'+10'	13'+11'	13'+12'	13'+13'	13'+14'	13'+15'	13'+16'
62	14	14+1'	14+2'	14+3'	14+4'	14+5'	14+6'	14+7'	14+8'	14+9'	14+10'	14+11'	14+12'	14+13'		14+15	14+16
63	14'	14'+1	14'+2	14'+3'	14'+4'	14'+5'	14'+6'	14'+7'	14'+8'	14'+9'	14'+10'	14'+11'	14'+12'	14'+13'	14'+14'	14'+15'	14'+16'
64	15	15+1'	15+2'	15+3'	15+4'	15+5'	15+6'	15+7'	15+8'	15+9'	15+10'	15+11'	15+12'	15+13'	15+14'		15+16
65	15'	15'+1	15'+2	15'+3'	15'+4'	15'+5'	15'+6'	15'+7'	15'+8'	15'+9'	15'+10'	15'+11'	15'+12'	15'+13'	15'+14'	15'+15'	15'+16'
66	16	16+1'	16+2'	16+3'	16+4'	16+5'	16+6'	16+7'	16+8'	16+9'	16+10'	16+11'	16+12'	16+13'	16+14'	16+15'	
67	16'	16'+1	16'+2	16'+3'	16'+4'	16'+5'	16'+6'	16'+7'	16'+8'	16'+9'	16'+10'	16'+11'	16'+12'	16'+13'	16'+14'	16'+15'	16'+16'

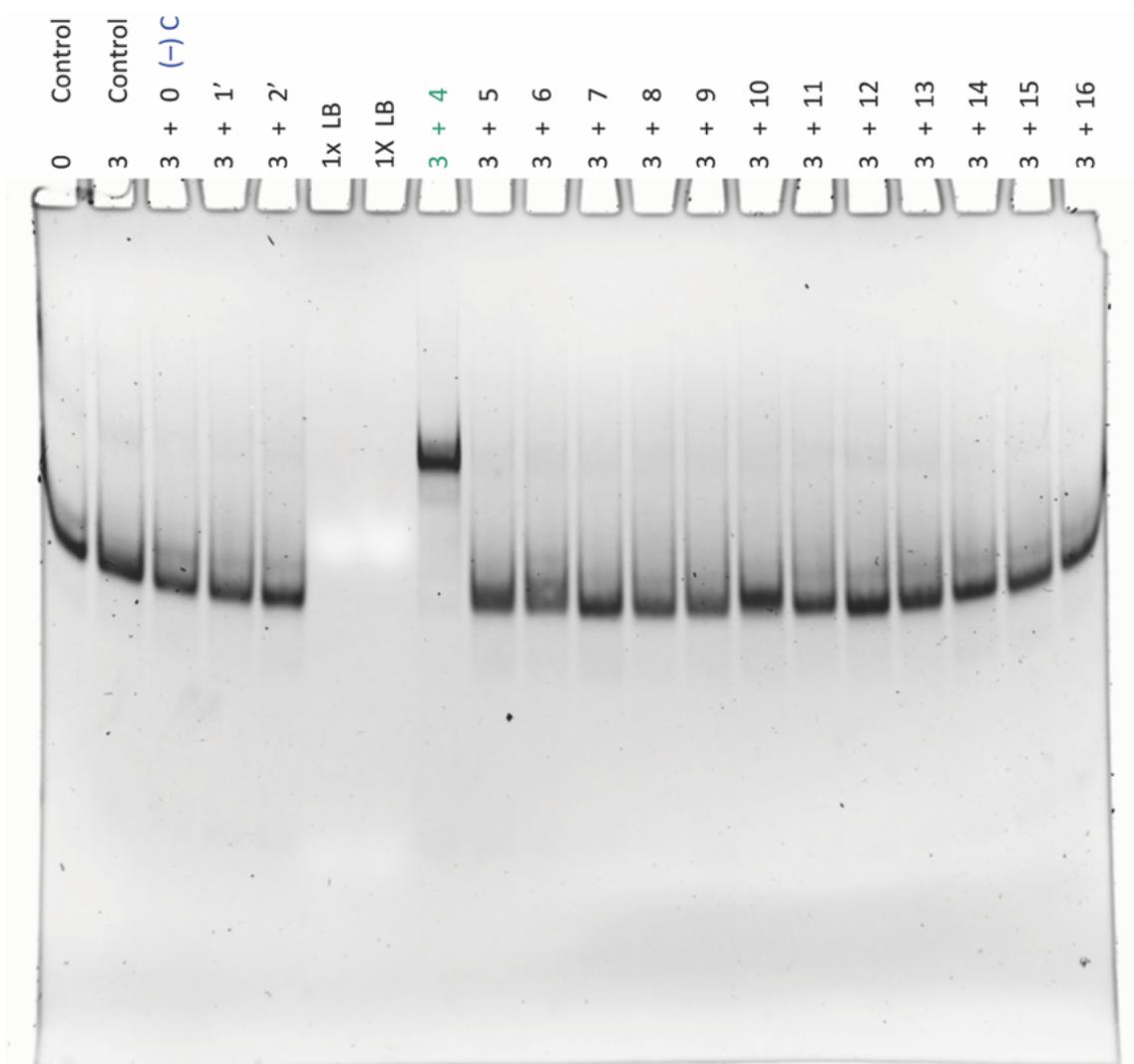
Table A.3.2: Results from crossTalk #2 experiment; alternative view. Each row represents two gels from the experiment, where separation between the two gels is the bold line running down the whole data matrix. Gel numbers can be seen in the very left and very right columns. Nomenclature is greatly simplified, where each hairpin structure fully named KL_1_GFPs is simplified to **1**, respectively. Additionally, each hairpin structure is also bound with a GFPasDisplaceStrand. The top four gels (96, 97, 98, and 99) only have the hairpin with displacement strand structures. See Figure 5.3 for understanding color coding. This Table represents the same data as **Table 5.4** in main text.

		cT2																	
96		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	97	
98		1'	2'	3'	4'	5'	6'	7'	8'	9'	10'	11'	12'	13'	14'	15'	16'	99	
100	1		1+2	1+3	1+4	1+5	1+6	1+7	1+8	1+9	1+10	1+11	1+12	1+13	1+14	1+15	1+16	101	
102	1'	1'+1	1'+2	1'+3	1'+4	1'+5	1'+6	1'+7	1'+8	1'+9	1'+10	1'+11	1'+12	1'+13	1'+14	1'+15	1'+16	103	
104	2	2+1		2+3	2+4	2+5	2+6	2+7	2+8	2+9	2+10	2+11	2+12	2+13	2+14	2+15	2+16	105	
106	2'	2'+1	2'+2	2'+3	2'+4	2'+5	2'+6	2'+7	2'+8	2'+9	2'+10	2'+11	2'+12	2'+13	2'+14	2'+15	2'+16	107	
108	3	3+1'	3+2'		3+4	3+5	3+6	3+7	3+8	3+9	3+10	3+11	3+12	3+13	3+14	3+15	3+16	109	
110	3'	3'+1	3'+2	3'+3	3'+4	3'+5	3'+6	3'+7	3'+8	3'+9	3'+10	3'+11	3'+12	3'+13	3'+14	3'+15	3'+16	111	
112	4	4+1'	4+2'	4+3'		4+5	4+6	4+7	4+8	4+9	4+10	4+11	4+12	4+13	4+14	4+15	4+16	113	
114	4'	4'+1	4'+2	4'+3	4'+4	4'+5	4'+6	4'+7	4'+8	4'+9	4'+10	4'+11	4'+12	4'+13	4'+14	4'+15	4'+16	115	
116	5	5+1'	5+2'	5+3'	5+4'		5+6	5+7	5+8	5+9	5+10	5+11	5+12	5+13	5+14	5+15	5+16	117	
118	5'	5'+1	5'+2	5'+3	5'+4	5'+5	5'+6	5'+7	5'+8	5'+9	5'+10	5'+11	5'+12	5'+13	5'+14	5'+15	5'+16	119	
120	6	6+1'	6+2'	6+3'	6+4'	6+5'		6+7	6+8	6+9	6+10	6+11	6+12	6+13	6+14	6+15	6+16	121	
122	6'	6'+1	6'+2	6'+3	6'+4	6'+5	6'+6	6'+7	6'+8	6'+9	6'+10	6'+11	6'+12	6'+13	6'+14	6'+15	6'+16	123	
124	7	7+1'	7+2'	7+3'	7+4'	7+5'		7+8	7+9	7+10	7+11	7+12	7+13	7+14	7+15	7+16	125		
126	7'	7'+1	7'+2	7'+3	7'+4	7'+5	7'+6	7'+7	7'+8	7'+9	7'+10	7'+11	7'+12	7'+13	7'+14	7'+15	7'+16	127	
128	8	8+1'	8+2'	8+3'	8+4'	8+5'	8+6'	8+7'		8+9	8+10	8+11	8+12	8+13	8+14	8+15	8+16	129	
130	8'	8'+1	8'+2	8'+3	8'+4	8'+5	8'+6	8'+7	8'+8	8'+9	8'+10	8'+11	8'+12	8'+13	8'+14	8'+15	8'+16	131	
132	9	9+1'	9+2'	9+3'	9+4'	9+5'	9+6'	9+7'	9+8'		9+10	9+11	9+12	9+13	9+14	9+15	9+16	133	
134	9'	9'+1	9'+2	9'+3	9'+4	9'+5	9'+6	9'+7	9'+8	9'+9	9'+10	9'+11	9'+12	9'+13	9'+14	9'+15	9'+16	135	
136	10	10+1'	10+2'	10+3'	10+4'	10+5'	10+6'	10+7'	10+8'	10+9'		10+11	10+12	10+13	10+14	10+15	10+16	137	
138	10'	10'+1	10'+2	10'+3	10'+4	10'+5	10'+6	10'+7	10'+8	10'+9	10'+10	10'+11	10'+12	10'+13	10'+14	10'+15	10'+16	139	
140	11	11+1'	11+2'	11+3'	11+4'	11+5'	11+6'	11+7'	11+8'	11+9'	11+10'		11+12	11+13	11+14	11+15	11+16	141	
142	11'	11'+1	11'+2	11'+3	11'+4	11'+5	11'+6	11'+7	11'+8	11'+9	11'+10	11'+11	11'+12	11'+13	11'+14	11'+15	11'+16	143	
144	12	12+1'	12+2'	12+3'	12+4'	12+5'	12+6'	12+7'	12+8'	12+9'	12+10'	12+11'		12+13	12+14	12+15	12+16	145	
146	12'	12'+1	12'+2	12'+3	12'+4	12'+5	12'+6	12'+7	12'+8	12'+9	12'+10	12'+11	12'+12	12'+13	12'+14	12'+15	12'+16	147	
148	13	13+1'	13+2'	13+3'	13+4'	13+5'	13+6'	13+7'	13+8'	13+9'	13+10'	13+11'	13+12'		13+14	13+15	13+16	149	
150	13'	13'+1	13'+2	13'+3	13'+4	13'+5	13'+6	13'+7	13'+8	13'+9	13'+10	13'+11	13'+12	13'+13	13'+14	13'+15	13'+16	151	
152	14	14+1'	14+2'	14+3'	14+4'	14+5'	14+6'	14+7'	14+8'	14+9'	14+10'	14+11'	14+12'	14+13'		14+15	14+16	153	
154	14'	14'+1	14'+2	14'+3	14'+4	14'+5	14'+6	14'+7	14'+8	14'+9	14'+10	14'+11	14'+12	14'+13	14'+14	14'+15	14'+16	155	
156	15	15+1'	15+2'	15+3'	15+4'	15+5'	15+6'	15+7'	15+8'	15+9'	15+10'	15+11'	15+12'	15+13'	15+14'		15+16	157	
158	15'	15'+1	15'+2	15'+3	15'+4	15'+5	15'+6	15'+7	15'+8	15'+9	15'+10	15'+11	15'+12	15'+13	15'+14	15'+15	15'+16	159	
160	16	16+1'	16+2'	16+3'	16+4'	16+5'	16+6'	16+7'	16+8'	16+9'	16+10'	16+11'	16+12'	16+13'	16+14'	16+15'		161	
162	16'	16'+1	16'+2	16'+3	16'+4	16'+5	16'+6	16'+7	16'+8	16'+9	16'+10	16'+11	16'+12	16'+13	16'+14	16'+15	16'+16	163	

A.4 Sample gels from crossTalk #1 experiment

CrossTalk #1 experiment consisted of 34 medium-sized native-PAGs.

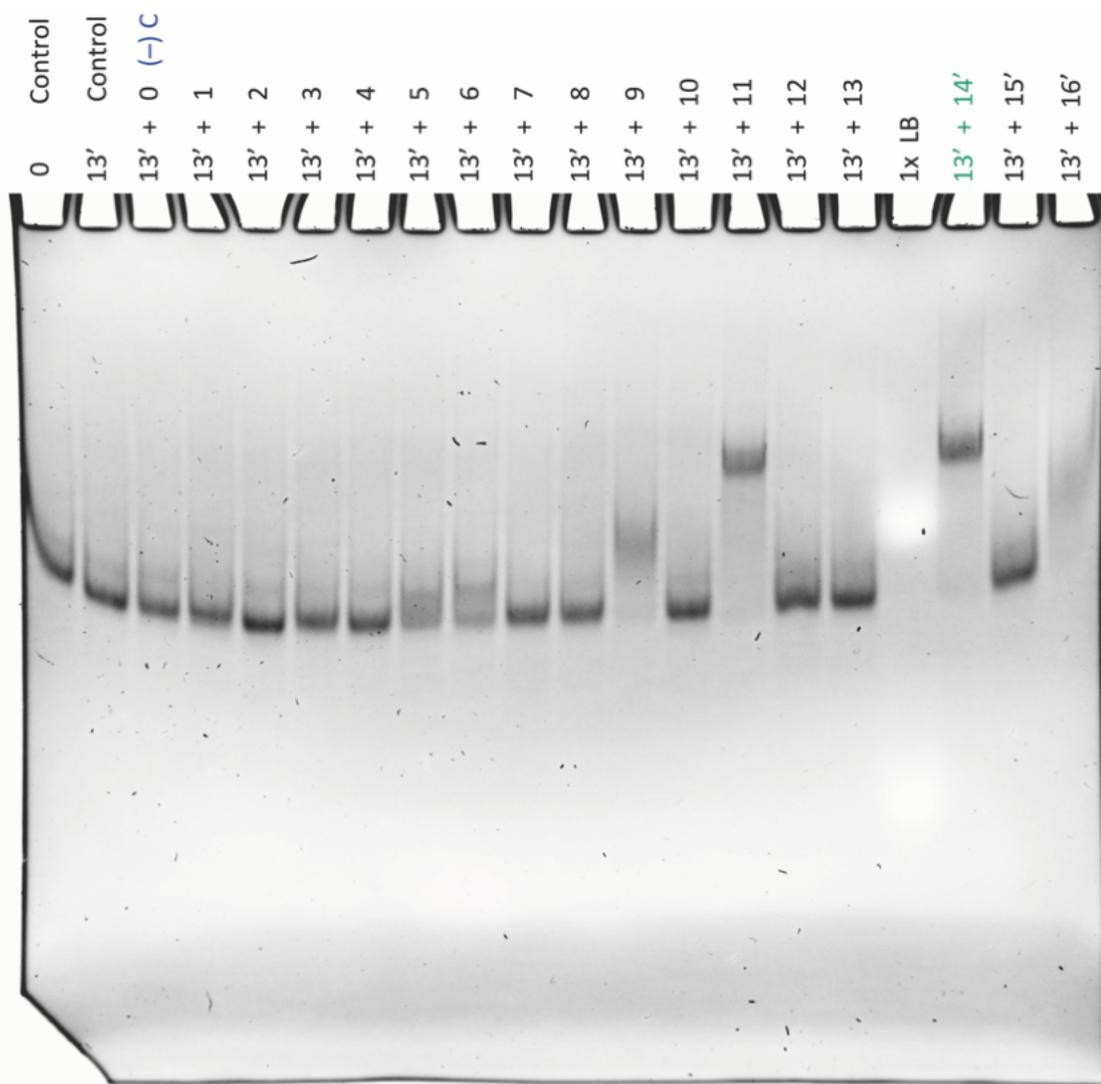
Gels 040 and 061 are shown below.



Gel 040, KL_cT1_OAS,
(19:1 A:bA) 8% native 2mM Mg⁺⁺, medium gel, C.B.S. Scientific Gel Stand
bioRad 300V, 150mA, 60min, total EtBr staining 5min

File: Gel040_hivKL_cT1_OAS_01.23.19_bioRad300V150mA60min_EtBr5min

* 0 = (KL_hiv0) + (GFPas_DS27nt)
...



Gel 061, KL_cT1_OAS,
 (19:1 A:bA) 8% native 2mM Mg⁺⁺, medium gel, C.B.S. Scientific Gel Stand
 bioRad 300V, 150mA, 60min, total EtBr staining 5min

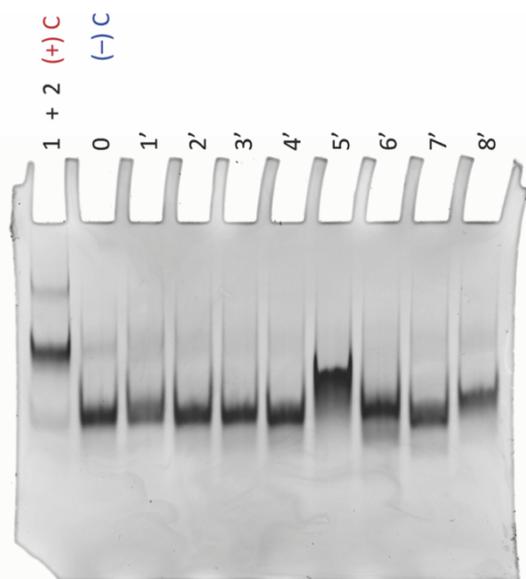
File: Gel061_hivKL_cT1_OAS_02.06.19_bioRad300V150mA60min_EtBr5min

* 0 = (KL_hiv0) + (GFPas_DS27nt)
 ...

A.5 Sample gels from crossTalk #2 experiment

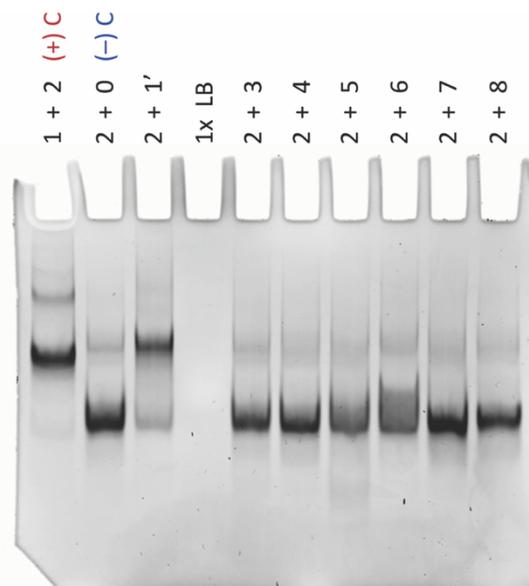
CrossTalk #2 experiment consisted of 68 Bio-Rad Mini-PROTEAN® native-PAGs.

Gels 098, 104, 124, and 153 are shown below.



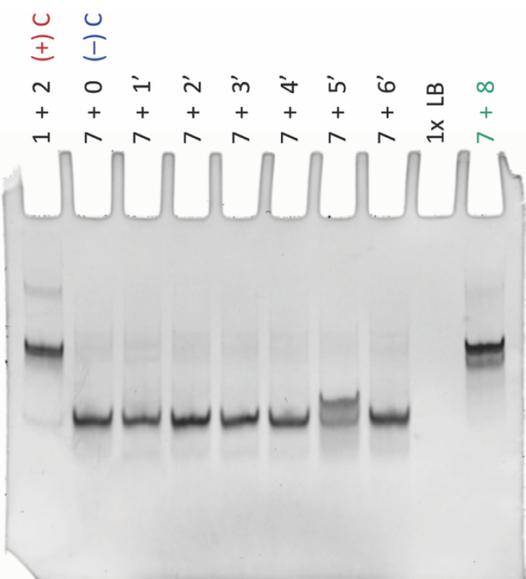
Gel 098, KL_ct2_OAS,
(19:1 A:bA) 8% native 2mM Mg⁺⁺,
bioRad 300V, 150mA, 30min, total EtBr staining 5min

File: Gel098_hivKL_ct2_OAS_11.06.19_bioRad300V150mA30min_EtBr5min



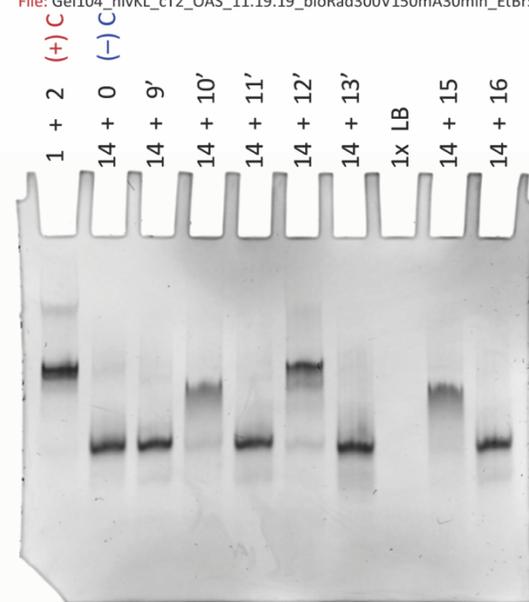
Gel 104, KL_ct2_OAS,
(19:1 A:bA) 8% native 2mM Mg⁺⁺,
bioRad 300V, 150mA, 30min, total EtBr staining 5min

File: Gel104_hivKL_ct2_OAS_11.19.19_bioRad300V150mA30min_EtBr5min



Gel 124, KL_ct2_OAS,
(19:1 A:bA) 8% native 2mM Mg⁺⁺,
bioRad 300V, 150mA, 30min, total EtBr staining 5min

File: Gel124_hivKL_ct2_OAS_12.16.19_bioRad300V150mA30min_EtBr5min



Gel 153, KL_ct2_OAS,
(19:1 A:bA) 8% native 2mM Mg⁺⁺,
bioRad 300V, 150mA, 30min, total EtBr staining 5min

File: Gel153_hivKL_ct2_OAS_01.14.20_bioRad300V150mA30min_EtBr5min

A.6 Sample gels from Tm #1 experiment

Tm #2 experiment consisted of 18 Bio-Rad Mini-PROTEAN® native-PAGs.

Gels 176, 179, 188, and 192 are shown below.

