Sequence- and structure-specific targeting of RNAs by short nucleobase-modified dsRNA-Binding PNAs Incorporating A-U pair-recognizing fluorescent light-up benzothiophene uracil and G-C pair-recognizing guanidinium

Shivakumar Krishna Manchugondanahalli

2020

Shivakumar Krishna Manchugondanahalli. (2020). Sequence- and structure-specific targeting of RNAs by short nucleobase-modified dsRNA-Binding PNAs Incorporating A-U pair-recognizing fluorescent light-up benzothiophene uracil and G-C pair-recognizing guanidinium. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/137354>

<https://doi.org/10.32657/10356/137354>

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Downloaded on 26 Apr 2023 17:54:21 SGT

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**SEQUENCE- AND STRUCTURE-SPECIFIC
TARGETING OF RNAs BY SHORT NUCLEOBASE-
MODIFIED DSRNA-BINDING PNAs
INCORPORATING A-U PAIR-RECOGNIZING
FLUORESCENT LIGHT-UP BENZOTHIOPHENE
URACIL AND G-C PAIR-RECOGNIZING
GUANIDINIUM**

SHIVAKUMAR KRISHNA MANCHUGONDANAHALLI

**SCHOOL OF PHYSICAL AND MATHEMATICAL
SCIENCES**

2020

**SEQUENCE- AND STRUCTURE-SPECIFIC
TARGETING OF RNAs BY SHORT NUCLEOBASE-
MODIFIED DSRNA-BINDING PNAs
INCORPORATING A-U PAIR-RECOGNIZING
FLUORESCENT LIGHT-UP BENZOTHIOPHENE
URACIL AND G-C PAIR-RECOGNIZING
GUANIDINIUM**

SHIVAKUMAR KRISHNA MANCHUGONDANAHALLI

SCHOOL OF PHYSICAL AND MATHEMATICAL SCIENCES

A thesis submitted to the Nanyang Technological
University in partial fulfilment of the requirement for
the degree of Doctor of Philosophy

2020

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research done by me except where otherwise stated in this thesis. The thesis work has not been submitted for a degree or professional qualification to any other university or institution. I declare that this thesis is written by myself and is free of plagiarism and of sufficient grammatical clarity to be examined. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

25 Nov 2019

.....

Date

.....

Shivakumar Krishna Manchugondanahalli

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it of sufficient grammatical clarity to be examined. To the best of my knowledge, the thesis is free of plagiarism and the research and writing are those of the candidate's except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

25 Nov 2019

.....

Date

.....

Chen Gang

Authorship Attribution Statement

This thesis contains material from two papers published in the following peer-reviewed journals in which I am listed as an author.

Chapter 2 is published as M. S. Krishna, D.-F. K. Toh, Z. Meng, A. A. L. Ong, Z. Wang, Y. Lu, K. Xia, M. Prabakaran, G. Chen. Sequence- And Structure-Specific Probing of RNAs by Short Nucleobase-Modified dsRNA-Binding PNAs Incorporating a Fluorescent Light-up Uracil Analog. *Analytical Chemistry* **91**, 5331–5338 (2019). DOI: 10.1021/acs.analchem.9b00280.

The contributions of the co-authors are as follows:

- Dr Chen provided the initial project direction.
- I carried out all the laboratory work which involves the synthesis of the ^{bt}U, L, and Q PNA monomers, PNA oligomers and characterized all the samples.
- I carried out all the fluorescence titration experiments, circular dichroism experiments and analyzed all the data.
- Ms D.-F. K. Toh performed PAGE experiments and analyzed the data.
- Mr. Ong assisted in the synthesis of Q monomer intermediates used in IR-1X PNA.
- Dr Meng, Dr Lu, and Dr Xia did computer modelling studies.
- Mr Wang and Dr Prabakaran provide us the total RNA extract from mock-infected and influenza virus-infected cells.
- I wrote the manuscript drafts. The manuscript was revised together with Dr Chen, Dr Meng, and Dr Prabakaran.

Chapter 3 is published as M. S. Krishna, Z. Wang, L. Zheng, J. Bowry, A. A. L. Ong, Y. Mu, M. Prabakaran, and G. Chen. Incorporating G-C Pair-Recognizing Guanidinium into PNAs for Sequence and Structure Specific Recognition of dsRNAs over dsDNAs and ssRNAs. *Biochemistry* **58**, 3777–3788 (2019). DOI: 10.1021/acs.biochem.9b00608.

The contributions of the co-authors are as follows:

- Dr Chen provided the initial project direction.
- I carried out all the laboratory work which involves the synthesis of the R, L, and Q PNA monomers, PNA oligomers and characterized all the samples.
- I carried out all the PAGE, fluorescence titration, thermal melting experiments and analyzed all the data
- Mr. Ong assisted in the synthesis of Q monomer intermediates used to target influenza panhandle viral RNA duplex.
- Mr Wang and Dr Prabakaran did the confocal microscopic experiments.
- Dr Zheng and Dr Mu did the computer modelling studies.
- I wrote the manuscript drafts. The manuscript was revised together with Dr Chen, Dr Zheng, and Dr Prabakaran.

25 Nov 2019

.....

.....

Date

Shivakumar Krishna Manchugondanahalli

Abstract

Ribonucleic acids (RNAs) play prominent roles in many biological activities. The structure of RNA determines its function not only in translation but also in catalysis, gene regulation, immunomodulation and so on. RNA consists of nitrogenous bases, a ribose sugar and a phosphate group extending from 5' to 3' end through a phosphodiester bond. RNA secondary structures contain loops and stems. The stem region often has a duplex structure formed through Watson–Crick hydrogen bonding and base stacking. RNA tertiary structures involve interactions among secondary structure units. A triplex is a tertiary structure motif and is formed when an RNA duplex accepts one more strand of RNA either in its major or minor groove. Major groove RNA triplexes are comparatively more stable than minor groove triplexes.

Peptide nucleic acids (PNAs) are a structural imitate of RNA with an acyclic pseudo-peptide skeleton made of repeating aminoethylglycine (aeg) units. PNAs have several advantages over naturally occurring nucleic acids and the analogues containing sugar phosphate backbones. Due to the presence of a neutral backbone, ability to withstand enzymatic degradation, easy functionalization at both backbone and nucleobases, PNAs have a great potential in applications in therapeutics and in diagnosis. A relatively short PNA pyrimidine strand can bind to the major-groove of a complementary RNA duplex containing a purine-rich strand through Hoogsteen hydrogen bonding resulting in the formation of a PNA•RNA₂ triplex. Numerous modifications have been developed to improve the binding properties of PNAs. For example, several chemically modified PNA backbones and bases were reported to (i) achieve the recognition of Watson–

Crick A-U and G-C as well as inverted Watson–Crick U-A and C-G base pairs at physiological pH and salt conditions and to (ii) improve cellular uptake of PNAs.

Project 1. PNA bases with improved binding and photophysical properties can be achieved through the chemical modifications. Recently, RNAs have been emerging as an important target for therapeutics and diagnosis. Compared to targeting RNAs through duplex formation, targeting the pre-formed dsRNA regions through structure-specific triplex formation provides a complementary RNA probing strategy. The method is more advantageous in terms of binding specificity with fast binding kinetics as compared to traditional antisense strategy which can be used only to detect RNA sequences rather than structures. In this context, we have developed a second-generation modified ^{bt}U PNA monomer derived from uracil which recognizes the Watson–Crick A-U base pair. The ^{bt}U monomer shows fluorescence enhancement upon triplex formation with the complementary duplex RNA under physiological pH conditions. Incorporating ^{bt}U monomer along with thio-pseudoisocytosine (L) and guanidine-modified 5-methylcytosine (Q) PNA monomer facilitates a sequence specific and selective fluorescence enhancement for duplex RNA but not duplex DNA or single-stranded RNA. We also extended our method for the direct detection of influenza viral panhandle structure and HIV-1 –1 ribosomal frameshift inducing RNA hairpins by designing complementary PNA strands incorporated with ^{bt}U monomer. Therefore, dsRNA-binding peptide nucleic acids (dbPNAs) with chemically modified nucleobases will be useful biochemical tools in the understanding and mapping of complex secondary structures of RNA through the fluorescence enhancement in a programmable fashion.

Project 2. Selective binding of PNAs to RNA duplex through the triplex-formation provides a useful tool for the understanding of RNA secondary structures and thus its functions. Compared to duplex formation which involves Watson–Crick hydrogen bonding, the triplex formation through Hoogsteen hydrogen bonding is relatively challenging at near-physiological conditions. Natural bases such as cytosine require low pH conditions for the recognition of Watson–Crick G-C base pairs. Guanidine has unique hydrogen bonding and stacking properties recognizes G bases in nucleic acids and is often present in many proteins and natural products. Due to its amphiphilic nature, guanidine groups help deliver cargo in the cellular delivery of many bio active compounds. As evidenced by our computer modelling study, we hypothesized that guanidium (protonated guanidine) with a suitable linker length attach to the PNA backbone can be useful in recognizing Watson–Crick G-C base pair in an RNA duplex. Synthesis of guanidium PNA monomer (designated R monomer) is simple, facile and involves only few well reported steps. Subsequently, we grafted the R monomer in PNA oligomer sequences and studied the binding and selectivity through PAGE, 2-aminopurine fluorescence, and thermal melting experiments. The data obtained confirmed that, R monomer shows remarkable sequence specificity and recognizes Watson–Crick G-C base pair in an RNA duplex over DNA duplex and single-stranded RNA at near physiological conditions. PNA R monomer is compatible with C, T, L, and Q PNA monomers and binds to biological RNAs such as influenza A viral panhandle duplex structure and HIV–1 frameshift site RNA hairpin. The presence of R residue in PNA sequences improves the solubility and the cell uptake properties.

Acknowledgements

I have spent 3.5 years as a graduate student to fulfil my dream at Nanyang Technological University, Singapore. There are many individuals who contributed to this journey both directly and indirectly. I will take an opportunity to express my sincere gratitude and heartfelt thanks to every one of them for providing the indispensable support, valuable guidance, and timely help.

First of all, it is my supervisor Professor Chen Gang, without his kind support and encouragement, it would have not been possible to accomplish this. His knowledge and wisdom helped me to shape my research career throughout my candidature. It's always a joy and pleasure to work with him. He gave me a free hand to design and develop the research projects and helped me to overcome the challenging situations by sharing his thoughts and ideas. I am here very grateful and obliged to him for this opportunity.

Meanwhile, I will also thank my thesis advisory committee members Prof. Phan Anh Tuan and Prof. Zhang Qichun for their fruitful suggestions during the meetings.

I am greatly thankful to all the lab members of our group. Apart from teaching all the basic lab skills, they created a suitable friendly environment, which is always very joyful and fun.

I would like to express my gratitude to NTU technical staff for granting me the access to several instruments required to carry out my research and the office people for the paperwork.

I would like to express a special thanks to my parents and grandparents, sister, mama, all the other family members and friends who nurtured my dream to grow big and stood beside me during my difficult time. There is one person who could be the reason that I have wandered different places to achieve my dream, thanks to you for coming to my life at the right time.

I am very thankful and obliged for the support and encouragement received from my village people. I am very overwhelmed for the blessings received from them. It helped me to stay more focused and responsible.

Finally, I am grateful to Nanyang Technological University, Singapore, for providing me a space to carry out my research work with its great infrastructure and for the financial assistantship through the MOE Tier 2 research scholarship.

Table of Contents

Abstract	1
Acknowledgements	4
Table of Contents	6
List of Abbreviations	9
Chapter 1 Introduction	12
1.1 Ribonucleic acids (RNAs)	12
1.2 RNA triplexes	14
1.3 Peptide nucleic acids (PNAs)	16
1.3.1 Modifications of the peptide nucleic acids (PNAs) at the backbone	17
1.3.2 Modifications of the peptide nucleic acids (PNAs) at the nucleobases	19
1.4 Developments on dsRNA-binding fluorescent light-up PNA probes	23
References	27
Chapter 2 Sequence- and structure-specific probing of RNAs by short nucleobase-modified dsRNA-binding PNAs incorporating a fluorescent light-up uracil analog	30
2.1 Introduction	30
2.2 Materials and Methods	34
2.2.1 Computational modeling	34
2.2.2 General experimental methods	35
2.2.2.1 Synthesis of fluorescent light-up ^{bt} U PNA monomer.....	35

2.2.3	Solid phase synthesis of PNA oligomers.....	40
2.2.4	Determination of extinction co-efficient.....	41
2.2.5	Fluorescence binding study	41
2.2.6	Estimation of the Limit of Detection (LOD)	42
2.2.7	Nondenaturing polyacrylamide gel electrophoresis.....	43
2.2.8	Circular dichroism spectroscopic characterization.....	43
2.2.9	Cell line and virus.....	44
2.2.10	Viral infection and extraction of total RNA.....	44
2.3	Results and Discussion	45
2.3.1	Modelling of PNA·RNA ₂ triplexes containing ^{bt} U·A-U triples...45	
2.3.2	Fluorescence of ^{bt} U monomer and ^{bt} U-containing PNA oligomers.....	46
2.3.3	Fluorescence detection of duplex RNAs through triplex formation.....	51
2.3.4	Fluorescence detection of dsRNA over ssRNA and dsDNA.....	61
2.3.5	Developing dsRNA-binding PNAs for the direct detection of influenza viral panhandle structure and HIV-1 –1 ribosomal frameshift inducing RNA hairpin.....	63
2.4	Conclusions	70
	References	72

Chapter 3 Incorporating G-C Pair-Recognizing Guanidinium into PNAs for Sequence and Structure Specific Recognition of dsRNAs over dsDNAs and

ssRNA	75
3.1 Introduction	75
3.2 Materials and Methods	79
3.2.1 Computational modeling	79
3.2.2 General experimental methods.....	79
3.2.2.1 Synthesis of PNA monomer R	80
3.2.3 Solid phase synthesis of PNA oligomers	84
3.2.4 Nondenaturing polyacrylamide gel electrophoresis	85
3.2.5 Fluorescence binding study	86

3.2.6	Thermal melting	87
3.2.7	Cellular uptake and confocal microscopy study	87
3.3	Results and Discussion	88
3.3.1	Recognition of an internal G-C pair in an RNA duplex (but not DNA duplex) by R-modified PNAs	88
3.3.2	Effects of sequence context, pH, and salt on PNA·RNA ₂ triplex formation	99
3.3.3	Incorporation of PNA R residues stabilizes and destabilizes PNA·RNA ₂ triplex and PNA-RNA duplex, respectively	100
3.3.4	Binding to HIV-1 minus-one ribosomal frameshift inducing RNA hairpin and influenza viral panhandle structure by dbPNAs containing multiple R residues	102
3.3.5	Cellular uptake of an R-modified PNA	106
3.4	Conclusions	108
	References	109
Chapter 4	Conclusion and Future Work	113
	References	119
Appendixes		120
Appendix 1.1	¹ H and ¹³ C NMR Spectra for the synthesized compounds studied in chapter 2	120
Appendix 1.2	MALDI-TOF characterization of PNA oligomers in chapter 2.	124
Appendix 2.1	¹ H and ¹³ C NMR Spectra for the synthesized compounds studied in chapter 3	131
Appendix 2.2	MALDI-TOF characterization of PNA oligomers in chapter 3.	135
Curriculum Vitae / List of Publications		144

List of Abbreviations

A	Adenine
2-AP	2-Aminopurine
Boc	tert-Butyloxycarbonyl
^{bt} U	5-benzothiopheneuracil
δ	Chemical shift
Cy3	Cyanine-3
C	Cytosine
CHCA	α -cyano-4-hydroxycinnamic acid
°C	degrees Celsius
CD	circular dichroism
DCM	Dichloromethane
DMF	Dimethylformamide
DIPEA	N,N-Diisopropylethylamine
dHP	DNA hairpin
dsDNA	Double-stranded Deoxyribonucleic acid
dsRNA	Double-stranded Ribonucleic acid
dbPNA	dsRNA binding peptide nucleic acid
E	3-oxo-2,3-dihydropyridazine
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetraacetic acid
Fmoc	Fluorenylmethyloxycarbonyl
G	Guanine
HBTU	3-[Bis(dimethylamino)methylumyl]-3H-benzotriazol-1-oxide Hexafluorophosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIV	Human immunodeficiency virus
HRMS-EI ionization	High-resolution mass spectrometry with electron
iC	isocytosine
J	pseudoisocytosine
L	thio-pseudoisocytosine

LOD	Limit of detection
LCMS-ESI	Liquid chromatography-mass spectrometry with electrospray ionization
M	2-Aminopyridine
MALDI-TOF	Matrix-assisted laser desorption/ionization-Time of flight
MBHA·HCl	4-Methylbenzhydrylamine hydrochloride
MES	2-(N-morpholino)ethanesulfonic acid
mHz	MegaHertz
μ	Micro
MDCK	Madin-Darby Canine Kidney
NIR	Near infrared
nM	Nanomolar
NMM	N-methyl morpholine
NMR	Nuclear Magnetic Resonance
PAGE	Polyacrylamide gel electrophoresis
P	2-pyrimidinone base with a carbonyl methylene linker
P _{ex}	2-pyrimidinone base with a carbonyl ethylene linker
PBS	Phosphate Buffer Saline
Pd/C	Palladium on carbon
ppm	Part per million
PyBop	(benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate)
Q	guanidine-modified 5-methylcytosine
rHP	RNA hairpin
ROI	RNAs of interest
RP-HPLC	Reverse phase high-performance liquid chromatography
ssRNA	single-stranded RNA
TBE	Tris-borate-EDTA
TFA	Trifluoroacetic acid
TFMSA	Trifluoromethanesulfonic acid
TFO	Triplex forming oligonucleotide
T	Thymine
TICT	Twisted Intramolecular Charge Transfer
THF	Tetrahydrofuran

T_m	Melting temperature
U	Uracil
UV	Ultra-violent

Chapter 1

Introduction

1.1 Ribonucleic acids (RNAs)

Ribonucleic acids (RNAs) are macromolecules and play a prominent role in many biological activities. Recent advances in understanding the structures, energetics, dynamics, and functions of RNAs ratify that RNAs are not only limited to protein biosynthesis (such as ribosomal, transfer and messenger RNAs), but also involved in catalysis, gene regulation, immunomodulation, etc.¹ Small nuclear RNAs (snRNAs) are involved in pre-mRNA splicing.² Gene editing is done in a programmable manner by guide RNAs.³ Telomerase RNA serves as a template for the replications of chromosome ends.⁴ Signal recognition particle (SRP) RNA is responsible for protein translocation.⁵ Small nucleolar RNAs (snoRNAs) are necessary for the modification of ribosomal RNAs and other RNAs.⁶

RNAs exist in many simple and complex forms, which determine their biological activities.⁷ The primary structure of RNA is the sequence of nucleotides constituted by nitrogenous bases like adenine, guanine, cytosine and uracil attached to the five-carbon ribose sugar with carbon atoms numbered from 1' to 5', which are in turn connected by phosphate. Each nucleotide is linked with each other through a phosphodiester bond extending from 5' to 3' end (**Figure 1.1a**). RNA secondary structures contain both single-stranded loop regions and double-

stranded stem regions formed by Watson–Crick base pairing between two complementary RNAs (**Figure 1.1b**).⁸ RNA stem regions exist as a helical duplex referred to as double-stranded RNA (dsRNA) and are found in many biological systems. RNA tertiary structures are formed by the interactions between secondary structural units and include a loop-loop, stem-stem or stem-loop interactions.⁹ Further, the RNA tertiary structures also include triplex RNAs which are discussed in details in the following sections. In addition, RNAs like snRNPs and metabolite sensing riboswitches bind to proteins and many other ions or molecules forming complex structural features.^{10, 11}

Figure 1.1. RNA Chemical structure. **(a)** Primary structure of an RNA strand containing two nucleotides with guanine (G) and uracil nucleobases (U). **(b)** Watson–Crick A-U and G-C base pairs with hydrogen bonds shown in dashed lines.

1.2 RNA triplexes

RNA triplexes are formed when an oligonucleotide binds to the duplex RNA without disrupting the already existing Watson–Crick geometry.^{12, 13} In the duplex RNAs, there is still enough space in the relatively deep and narrow major groove to accommodate one more strand of oligonucleotide.^{14, 15} For example, U·A-U or T·A-U triplex formed when the third strand of the pyrimidine bases (Uracil or Thymine) recognize the purine sequences (Adenine) in the RNA duplex by Hoogsteen hydrogen bonding (**Figure 1.2**). The duplex is formed via Watson-Crick hydrogen bonding between poly(U) and poly(A), while the third poly(U) strand binds to the RNA duplex via Hoogsteen hydrogen bonding resulting in a triplex (**Figure 1.2**). The poly(U) strands, also known as pyrimidine strand, is found parallel to the purine strand as seen in most of the major groove triplex structures. In these cases, triplexes are known as the parallel motif triplexes. However, the interruptions of purine strand at one or more places by pyrimidine bases resulting in inverted Watson–Crick U-A or C-G sites are commonly found in many naturally occurring RNA duplexes.^{16, 17} As for minor groove triplexes, which are unstable in most of the cases when isolated, are usually formed within large RNA and protein bound RNA complexes.¹⁸ In nature, RNA triplexes seen in many biologically important RNAs such as –1 ribosomal frameshift-inducing mRNA pseudoknots,¹⁸ telomerase RNAs,¹⁹ metabolite-sensing riboswitches,^{10, 20} group I and group II introns,^{21, 22} ribosomal and long noncoding RNAs (lncRNAs).^{23, 24} The triplex structures are responsible for molding RNAs into complex three-dimensional architectures for diverse biological functions.²⁵

Figure 1.2. Major groove triples of $C^+ \cdot G-C$ and $U \cdot A-U$.

In the literature, several nucleic acid analogues which specifically target any sequence of DNA/RNA duplexes were reported (**Figure 1.3**).^{15, 26} However, due to the differences between DNA and RNA duplexes, the triplex binding properties differ greatly.^{14, 27} For example, 2'-OMe RNA (**Figure 1.3c**) has a greater tendency for binding DNA duplexes rather than RNA duplexes as confirmed by thermal melting and EMSA studies.^{15, 28} The possible reason could be due to the lack of the hydrogen bonding between 2'-OH in a 2-OMe RNA with nonbridging phosphate oxygen in the purine strand. Thus, the backbone-backbone hydrogen bonding between the third strand and the RNA duplex is critical for triplex formation. Further, the presence of a methoxy group in 2'-OMe RNA creates a steric clash and reduces its binding affinity toward duplex RNA over duplex DNA.^{15, 24, 27} On the other hand, locked nucleic acids (LNAs) (**Figure 1.3d**) shows a similar binding properties and forms triplexes with both DNA and RNA duplexes despite the lack of backbone-backbone hydrogen bonding in RNAs.¹⁵ This is presumably due to the steric constraint minimized by bridged methylene group formed between C_2' -O and C_3 atoms at the ribose sugar.

Figure 1.3. Chemical structures of natural and synthetic nucleic acid analogues.

(a and b) chemical structures of naturally occurring DNA and RNA. (c-e) Chemical structures of synthetic nucleic acid analogues 2'-OMe RNA, locked nucleic acid (LNA), and peptide nucleic acid (PNA).

1.3 Peptide nucleic acids (PNAs)

Unlike naturally occurring triplex forming RNA strands and triplex forming oligonucleotides (TFOs), PNAs (**Figure 1.3e**) have been proved to have enhanced binding affinity and sequence-specific recognition of RNAs over DNAs.^{12, 29, 30} PNAs have better geometrical compatibility with the duplex RNAs than duplex DNAs and show greater affinity for RNAs through triplex formation.^{29, 30} In addition, PNAs could prevent the unnecessary or undesirable binding of proteins and other tertiary interaction due to its distinct structural features over other natural nucleic acids and its structural analogues with traditional sugar phosphate backbone.¹²

PNAs were introduced by Nielsen et al. in 1991.³¹ PNAs have been one of the central platforms in nucleic acid chemistry to overcome the drawbacks confronted by traditional sugar-phosphate backbones.³² Currently, PNAs have

garnered a tremendous research interest mainly because of their potency to serve as a promising agents for altering and regulating the gene expressions and for establishing structural insights into the functions of DNA and RNA.^{33, 34} Also, their unique binding properties enable the detection of RNA/DNA structural polymorphism, a key feature of nucleic acid structures and functions in the living organisms.^{35, 36} Since its discovery, several modifications have been made to improve the binding and selectivity of PNAs with target nucleic acid sites.³⁷ These chemical modifications can be broadly classified into two types and are discussed in detail below.

1.3.1 Modifications of the peptide nucleic acids (PNAs) at the backbone

A large variety of PNAs with backbone modification have been reported in the literature and some of them are shown in **Figure 1.4**. However, a simple PNA is constituted by the nucleic acid structural imitate of an acyclic pseudo-peptide skeleton made of repeating aminoethylglycine (aeg) units (**Figure 1.4b**).^{31, 32} Introducing substituents at the γ -position at the ethyl part of the backbone results in γ -PNAs which helps in imposing a specific chirality to the PNA strand (**Figure 1.4c**).³⁸⁻⁴⁰ In order to reduce the flexibility of the backbone, many PNAs comprising rigid cyclic structures have also been reported but, all these modifications result in complex sequence context effects (**Figure 1.4d,e**).^{41, 42} Therefore, PNAs made of simple aeg units are still often used for the recognition of nucleic acids.³⁴ PNAs containing positively charged lysine (**Figure 1.4f**)^{43, 44} and arginine (**Figure 1.4g**) show improved binding properties and biological activities.^{45, 46} In these cases, the presence of positive charge in the backbone results in the pronounced PNA-RNA binding and also increased cellular uptake.

The less explored α PNAs (**Figure 1.4h**) entails a nucleobase attached at every second or third α -amino acids in a peptide chain and shows weak recognition with any DNA or RNA structures.^{47, 48} The tPNAs (**Figure 1.4i**) are modified α PNAs with thio-ester linkages often shows improved binding and recognition of complementary DNA strand over its parent form.⁴⁹ Overall, a common rationale behind the design of these backbones are summarized as follows.

- The presence of the neutral peptide backbone has been exploited in achieving greater affinity against the complementary oligonucleotides without requiring cationic shielding of the phosphate backbone. In other words, the binding affinity of PNAs with that of naturally occurring nucleic acids are unaltered due to a small change in the ionic concentration.⁵⁰
- The presence of unnatural peptide backbone causes PNAs to withstand enzymatic degradation and serves as a promising entity in many in vivo studies.^{12, 37}
- Unlike the naturally occurring nucleic acids and the analogues built with traditional sugar-phosphate backbone, a PNA with expanded letters of nucleobases in the oligomer sequence can show a remarkable binding with the complementary DNA/RNA structures.^{29, 30, 32, 51}
- Easy functionalization at the backbone and/or at the nucleobases manifest the potential in developing numerous novel PNA monomers and oligomers with immense biological importance.¹²

Figure 1.4. Molecular structures of DNA backbone (a) and PNA backbones with various modifications (b-i).

1.3.2 Modifications of the peptide nucleic acids (PNAs) at the nucleobases

In addition to the chemical modifications at the backbone, modifications of nucleobases are another important aspect which allow an improved binding and selectivity over natural nucleic acid bases.¹² A list of all the reported duplex binding PNA bases were shown in **Figure 1.5**. Naturally occurring bases thymine and cytosine recognize the Watson-Crick A-U and G-C base pairs, respectively, (**Figure 1.5a,b**) in an RNA duplex.⁴³ Unlike thymine base, PNAs containing cytosine residue bind to the duplex DNA or RNA only at low pH conditions.¹² The reason behind this observation is that the N₃ nitrogen atom has a pK_a around 4.5, which is not basic enough to be protonated at physiological pH.⁵² Many attempts aiming for the increase of basicity of nitrogen atom for enhancing Hoogsteen bond formation have been documented. One such finding was the nucleobase 2-amino pyridine which is abbreviated as M monomer

(**Figure 1.5c**) reported by Rozners and co-workers.⁵¹ The pK_a of M monomer was found to be 6.7 and can assist the formation of triplexes under near-physiological conditions.

In contrast to using the charged protonated bases for the duplex binding, an alternative approach is to design neutral nucleobases. The popular examples include pseudoisocytosine abbreviated as J monomer (**Figure 1.5d**), which was originally developed by Kan et. al., for TFOs⁵³ and extended to PNAs by Nielsen group and Rozners group.^{51, 54} The use of J monomer allows the nucleic acid chemists to achieve the duplex binding by a PNA oligomer under physiological conditions which manifest pH independent binding and thus useful for applications for in vivo. In 2014, Chen and his research group reported thio-pseudoisocytosine (**Figure 1.5e**) abbreviated as L monomer.²⁹ Apart from achieving duplex binding at physiological pH conditions, L monomer known for its selective binding to the RNA duplex over complementary single-stranded RNA by stabilizing Hoogsteen L-G-C pairing and destabilizing Watson-Crick L-G pairing. Recently, the same group also reported R monomer (**Figure 1.5f**), which has a simple guanidine attached to the PNA backbone through a two-carbon linker. Unlike the L monomer, the synthesis of R monomer is facile and R shows a pH independent sequence specific binding to Watson-Crick G-C base pair.⁵⁵

The inverted Watson-Crick C-G and U-A base pairs are known to be present in many biological RNA duplex structures. The realization of the applications of PNAs in therapeutics and diagnosis is facilitated by designing and developing

suitable PNA monomers to recognize the inverted base pairs in RNA sequences. 5-Methylisocytosine (designated as iC) (**Figure 1.5g**)⁵⁶⁻⁵⁸ and 2-pyrimidinone base with carbonyl methylene and carbonyl ethylene linkages (P and P_{ex} monomers) (**Figure 1.5h,i**),⁵⁹⁻⁶¹ were incorporated into TFOs and PNAs to form Hoogsteen pairs with inverted Watson–Crick C-G base pair. A derivative of 5-methyl cytosine bearing ethylene guanidine (Q monomer) for the effective recognition of an inverted C-G base pair was reported by Chen et. al (**Figure 1.5j**).³⁰ Similar to L and R monomer, Q monomer also shows a selective recognition of duplex RNAs over single-stranded RNA and duplex DNA under physiological pH conditions. A modified nucleobase based on 3-oxo-2,3-dihydropyridazine (E monomer) by Rozner et. al., (**Figure 1.5k**)^{46, 61, 62} and PNA S-monomer (**Figure 1.5l**) by Chen and his co-workers were developed to recognize inverted Watson–Crick U-A base pair in an RNA duplex.⁶³ As a result, in the literature, a number of natural and modified nucleobases with a greater affinity for the selective binding to duplex RNAs through the Hoogsteen bonds were reported. In addition, a number of modified second-generation PNA bases have been reported, with improved binding, selectivity and photo physical properties.⁶⁴⁻⁶⁶

Figure 1.5. Base triples of PNA nucleobases to RNA base pairs through major groove triple formation (a-i). T: thymidine, C: cytidine, A: Adenine, G-Guanidine, U-uridine, J: pseudoisocytosine, L: thio-pseudoisocytosine, M⁺: Protonated 2-aminopyridine, iC: 5-methylisocytosine, P and P_{ex} are 2-pyrimidinone base with carbonyl methylene and carbonyl ethylene linkages, E: 3-oxo-2,3-dihydropyridazine, Q: 5-methylcytosine with ethylene guanidine.

1.4 Developments on dsRNA-binding fluorescent light-up PNA probes

In the literature, several methods have been developed to better understand the RNA structures, interactions, biogenesis, and functions, and to precisely detect and control functions of RNA in cells and viruses.⁶⁷ Most of these techniques involve the design of molecular probes which are promising in targeting unstructured regions of RNA. However, these probes are not specific for structured RNAs and may suffer from binding to off-target sites.⁶⁷ Many small molecule ligands, which display a specific binding, are often limited to non-Watson–Crick loops; it is challenging to target small molecules to the Watson–Crick structured regions.⁶⁸

In this context, duplex-binding PNAs provide an alternative tool with many advantages over conventional antisense methods and small molecules. Since duplex-binding involve the binding to pre-existing secondary structures, it is expected to involve fast binding kinetics without the need for unfolding the preformed structures. It also facilitates the design of oligomer sequences for directly detecting the RNAs of interest at any specific site in a programmable fashion. The RNA duplex-specific targeting strategy is relatively less explored, mainly because of the weak Hoogsteen bonding with the complementary duplex RNA under physiological pH conditions. Till date, PNA bases like thiazole orange (TO), benzo[c,d]indole-quinoline (BIQ) and derivatives 2-Aminopyridine (M monomer) have been the only reported probes in the literature for the direct detection of duplex RNAs (**Figure 1.6**) using a PNA platform.⁶⁹⁻⁷¹

Figure 1.6. Chemical structures of fluorescent turn-on nucleobases. **(a)** Thiazole orange base, **(b)** benzo[c,d]indole-quinoline, **(c-d)** PhEthM⁺·G-C and PhpM⁺·G-C base triples.

Fluorescence enhancement through the forced intercalation (FIT) method provides direct approach for the detection of any nucleic acid sequence/structure of interest. Seitz and his co-workers reported thiazole orange (TO) (**Figure 1.6a**) as an intercalating dye for the fluorescent detection of complementary DNA strand.⁷² The method involves a intercalating dye acting as a base surrogate and shows fluorescence enhancement upon binding to complementary nucleic acid strand and avoids the use of an accessory dye which can create instability due to its bulkiness while binding.^{67, 73, 74} The fluorescence enhancement arises due to the restricted bond rotation by intercalation upon duplex formation with the complementary nucleic acid strand. Later, several modified FIT PNA probes with remarkable photophysical properties were chemically synthesized and used extensively in an antisense PNA technology to study many biologically important RNAs.^{71, 75-77}

Meanwhile, TO has also been explored as a universal base surrogate in the real time detection of RNA duplexes. PNA oligomers incorporated with TO is very sensitive for the detection of RNA duplexes and shows significant decrease in

the fluorescence intensity for the mismatch sequences over the complementary duplex.⁶⁹ However, the fluorescence turn-on effect is observed only at pH 5.5 and fails to display any significant intensity at neutral pH conditions. Further, the lack of specific hydrogen bonding interactions for TO makes it a probe with no sequence- and structure specificity, and thus provides no detailed information of local base pair present opposite to TO. The fluorescence enhancement is mainly because of restricted bond rotation through the intercalation upon triplex formation.

In a recent report benzo[c,d]indole-quinoline (BIQ) (**Figure 1.6b**), was used as an intercalating dye which shows deep red fluorescent light-up signal upon binding to complementary single stranded RNA and duplex RNA. The fluorescent probe is sensitive to the presence of mismatch sequences and/or structures and thus was employed successfully in imaging nucleoli RNAs in the living cells.⁷¹ Similar to TO, benzo[c,d]indole-quinoline is also incapable to show any hydrogen bonding array with the opposite nucleobase and/or Watson–Crick base pairs in an RNA duplex.⁷¹

The fluorescence detection of duplex RNA under physiological pH conditions were reported by Rozner's and his co-workers using a PNA monomers derived from M monomer (**Figure 1.5c and 1.6c,d**). Out of two modified PNAs, oligomers incorporated with 3-(1-phenylethynyl)-M (**Figure 1.6c**) serves better than phenylpyrrolo-M (**Figure 1.6d**), as a fluorescent probe due to its high selectivity for the complementary Watson–Crick G-C base pair and has higher quantum yield.⁷⁰ The base stacking of 3-(1-phenylethynyl)-M upon triplex

formation results in quenching of fluorescence intensity instead of enhancement as observed for TO and BIQ base. However, in reports involving TO and modified M monomers, there is no experimental evidences for the selective detection of duplex RNAs over duplex DNAs and single-stranded RNAs. In addition, the study was limited for only model RNA hairpins, without the application in the detection of biologically relevant RNA duplexes.

Therefore, developing PNA probes, which selectively detect duplex RNAs over duplex DNA and single-stranded RNA under physiological pH conditions through the fluorescence enhancement has been challenging and underdeveloped. Hence, it is timely and significant to develop PNA-based probes, which address these issues of sequence- and structure-specific recognition of RNA duplexes for not only to understand the structures but also to regulate the functions.

References

1. Cech, T. R.; Steitz, J. A. *Cell* **2014**, 157, 77-94.
2. Nag, A.; Steitz, J. A. *RNA Biol.* **2012**, 9, 334-342.
3. Raper, A. T.; Stephenson, A. A.; Suo, Z. *J. Am. Chem. Soc.* **2018**, 140, 2971-2984.
4. Jiang, J.; Chan, H.; Cash, D. D.; Miracco, E. J.; Ogorzalek Loo, R. R.; Upton, H. E.; Cascio, D.; O'Brien Johnson, R.; Collins, K.; Loo, J. A.; Zhou, Z. H.; Feigon, J. *Science* **2015**, 350, aab4070.
5. Beckert, B.; Kedrov, A.; Sohmen, D.; Kempf, G.; Wild, K.; Sinning, I.; Stahlberg, H.; Wilson, D. N.; Beckmann, R. *Nat. Struct. Mol. Biol.* **2015**, 22, 767-773.
6. Tollervey, D. *Science* **1996**, 273, 1056-1057.
7. Doudna, J. A. *Nat. Struct. Biol.* **2000**, 7 Suppl, 954-956.
8. Tinoco, I., Jr.; Bustamante, C. *J. Mol. Biol.* **1999**, 293, 271-281.
9. Butcher, S. E.; Pyle, A. M. *Acc. Chem. Res.* **2011**, 44, 1302-1311.
10. Serganov, A.; Patel, D. J. *Nat. Rev. Genet.* **2007**, 8, 776-790.
11. Berglund, H.; Rak, A.; Serganov, A.; Garber, M.; Hard, T. *Nat. Struct. Biol.* **1997**, 4, 20-23.
12. Devi, G.; Zhou, Y.; Zhong, Z.; Toh, D. F.; Chen, G. *Wiley Interdiscip Rev. RNA* **2015**, 6, 111-128.
13. Bacolla, A.; Wang, G.; Vasquez, K. M. *PLoS Genet.* **2015**, 11, e1005696.
14. Roberts, R. W.; Crothers, D. M. *Science* **1992**, 258, 1463-1466.
15. Zhou, Y.; Kierzek, E.; Loo, Z. P.; Antonio, M.; Yau, Y. H.; Chuah, Y. W.; Geifman-Shochat, S.; Kierzek, R.; Chen, G. *Nucleic Acids Res.* **2013**, 41, 6664-6673.
16. Velagapudi, S. P.; Gallo, S. M.; Disney, M. D. *Nat. Chem. Biol.* **2014**, 10, 291-297.
17. Brakier-Gingras, L.; Charbonneau, J.; Butcher, S. E. *Expert Opin. Ther. Targets* **2012**, 16, 249-258.
18. Su, L.; Chen, L.; Egli, M.; Berger, J. M.; Rich, A. *Nat. Struct. Biol.* **1999**, 6, 285-292.
19. Qiao, F.; Cech, T. R. *Nat. Struct. Mol. Biol.* **2008**, 15, 634-40.
20. Liberman, J. A.; Salim, M.; Krucinska, J.; Wedekind, J. E. *Nat. Chem. Biol.* **2013**, 9, 353-355.
21. Adams, P. L.; Stahley, M. R.; Kosek, A. B.; Wang, J.; Strobel, S. A. *Nature* **2004**, 430, 45-50.
22. Toor, N.; Keating, K. S.; Taylor, S. D.; Pyle, A. M. *Science* **2008**, 320, 77-82.
23. Brown, J. A.; Valenstein, M. L.; Yario, T. A.; Tycowski, K. T.; Steitz, J. A. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, 109, 19202-19207.
24. Mitton-Fry, R. M.; DeGregorio, S. J.; Wang, J.; Steitz, T. A.; Steitz, J. A. *Science* **2010**, 330, 1244-1247.
25. Buske, F. A.; Mattick, J. S.; Bailey, T. L. *RNA Biol.* **2011**, 8, 427-439.
26. Fox, K. R.; Brown, T. *Biochem. Soc. Trans.* **2011**, 39, 629-634.
27. Escude, C.; Francois, J. C.; Sun, J. S.; Ott, G.; Sprinzl, M.; Garestier, T.; Helene, C. *Nucleic Acids Res.* **1993**, 21, 5547-5553.
28. Wang, S.; Kool, E. T. *Nucleic Acids Res.* **1995**, 23, 1157-1164.
29. Devi, G.; Yuan, Z.; Lu, Y.; Zhao, Y.; Chen, G. *Nucleic Acids Res.* **2014**, 42, 4008-4018.
30. Toh, D. K.; Devi, G.; Patil, K. M.; Qu, Q.; Maraswami, M.; Xiao, Y.; Loh, T. P.; Zhao, Y.; Chen, G. *Nucleic Acids Res.* **2016**, 44, 9071-9082.
31. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, 254, 1497-1500.
32. Nielsen, P. E. *Acc. Chem. Res.* **1999**, 32, 624-630.
33. Meister, G.; Tuschl, T. *Nature* **2004**, 431, 343-349.

34. Nielsen, P. E. *Chem. Biodivers.* **2010**, 7, 786-804.
35. Keene, M. A.; Elgin, S. C. *Cell* **1984**, 36, 121-129.
36. Nielsen, P. E. *Curr. Opin. Mol. Ther.* **2010**, 12, 184-191.
37. Rozners, E. *J. Nucleic Acids* **2012**, 2012, 518162.
38. Rapireddy, S.; He, G.; Roy, S.; Armitage, B. A.; Ly, D. H. *J. Am. Chem. Soc.* **2007**, 129, 15596-15600.
39. Sahu, B.; Sacui, I.; Rapireddy, S.; Zanolotti, K. J.; Bahal, R.; Armitage, B. A.; Ly, D. H. *J. Org. Chem.* **2011**, 76, 5614-5627.
40. Sacui, I.; Hsieh, W. C.; Manna, A.; Sahu, B.; Ly, D. H. *J. Am. Chem. Soc.* **2015**, 137, 8603-8610.
41. Ganesh, K. N.; Nielsen, P. E. *Curr. Org. Chem.* **2000**, 4, 931-943.
42. Pokorski, J. K.; Witschi, M. A.; Purnell, B. L.; Appella, D. H. *J. Am. Chem. Soc.* **2004**, 126, 15067-15073.
43. Nielsen, P. E.; Haaima, G.; Lohse, A.; Buchardt, O. *Angew. Chem. Int. Ed.* **1996**, 35, 1939-1942.
44. Tedeschi, T.; Sforza, S.; Dossena, A.; Corradini, R.; Marchelli, R. *Chirality* **2005**, 17, S196-S204.
45. Zhou, P.; Dragulescu-Andrasi, A.; Bhattacharya, B.; O'Keefe, H.; Vatta, P.; Hyldig-Nielsen, J. J.; Ly, D. H. *Bioorg. Med. Chem. Lett.* **2006**, 16, 4931-4935.
46. Gupta, P.; Muse, O.; Rozners, E. *Biochemistry* **2012**, 51, 63-73.
47. Howarth, N. M.; Wakelin, L. P. G. *J. Org. Chem.* **1997**, 62, 5441-5450.
48. Garner, P.; Dey, S.; Huang, Y. *J. Am. Chem. Soc.* **2000**, 122, 2405-2406.
49. Ura, Y.; Beierle, J. M.; Leman, L. J.; Orgel, L. E.; Ghadiri, M. R. *Science* **2009**, 325, 73-77.
50. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. *Nature* **1993**, 365, 566-568.
51. Zengeya, T.; Gupta, P.; Rozners, E. *Angew. Chem. Int. Ed.* **2012**, 51, 12593-6.
52. Li, M.; Zengeya, T.; Rozners, E. *J. Am. Chem. Soc.* **2010**, 132, 8676-8681.
53. Ono, A.; Ts'o, P. O. P.; Kan, L. S. *J. Am. Chem. Soc.* **1991**, 113, 4032-4033.
54. Egholm, M.; Christensen, L.; Dueholm, K. L.; Buchardt, O.; Coull, J.; Nielsen, P. E. *Nucleic Acids Res.* **1995**, 23, 217-222.
55. Krishna, M. S.; Wang, Z.; Zheng, L.; Bowry, J.; Ong, A. A. L.; Mu, Y.; Prabakaran, M.; Chen, G. *Biochemistry* **2019**, 58, 3777-3788.
56. Switzer, C.; Moroney, S. E.; Benner, S. A. *J. Am. Chem. Soc.* **1989**, 111, 8322-8323.
57. Chen, G.; Kierzek, R.; Yildirim, I.; Krugh, T. R.; Turner, D. H.; Kennedy, S. D. *J. Phys. Chem. B* **2007**, 111, 6718-6727.
58. Zengeya, T.; Li, M.; Rozners, E. *Bioorg. Med. Chem. Lett.* **2011**, 21, 2121-2124.
59. Buchini, S.; Leumann, C. J. *Angew. Chem. Int. Ed.* **2004**, 43, 3925-3928.
60. Gildea, B.; McLaughlin, L. W. *Nucleic Acids Res.* **1989**, 17, 2261-2281.
61. Gupta, P.; Zengeya, T.; Rozners, E. *Chem. Commun.* **2011**, 47, 11125-11127.
62. Eldrup, A. B.; Dahl, O.; Nielsen, P. E. *J. Am. Chem. Soc.* **1997**, 119, 11116-11117.
63. Ong, A. A. L.; Toh, D. K.; Patil, K. M.; Meng, Z.; Yuan, Z.; Krishna, M. S.; Devi, G.; Haruehanroengra, P.; Lu, Y.; Xia, K.; Okamura, K.; Sheng, J.; Chen, G. *Biochemistry* **2019**, 58, 1319-1331.
64. Patil, K. M.; Toh, D. K.; Yuan, Z.; Meng, Z.; Shu, Z.; Zhang, H.; Ong, A. A. L.; Krishna, M. S.; Lu, L.; Lu, Y.; Chen, G. *Nucleic Acids Res.* **2018**, 46, 7506-7521.
65. Ong, A. A. L.; Toh, D. K.; Krishna, M. S.; Patil, K. M.; Okamura, K.; Chen, G. *Biochemistry* **2019**, 58, 3444-3453.
66. Krishna, M. S.; Toh, D. K.; Meng, Z.; Ong, A. A. L.; Wang, Z.; Lu, Y.; Xia, K.; Prabakaran, M.; Chen, G. *Anal. Chem.* **2019**, 91, 5331-5338.

67. Vilaivan, T. *Beilstein J. Org. Chem.* **2018**, 14, 253-281.
68. Paige, J. S.; Wu, K. Y.; Jaffrey, S. R. *Science* **2011**, 333, 642-646.
69. Sato, T.; Sato, Y.; Nishizawa, S. *J. Am. Chem. Soc.* **2016**, 138, 9397-9400.
70. Cheruiyot, S. K.; Rozners, E. *ChemBioChem* **2016**, 17, 1558-1562.
71. Yoshino, Y.; Sato, Y.; Nishizawa, S. *Anal. Chem.* **2019**, 91, 14254-14260.
72. Seitz, O.; Bergmann, F.; Heindl, D. *Angew. Chem. Int. Ed.* **1999**, 38, 2203-2206.
73. Hovelmann, F.; Seitz, O. *Acc. Chem. Res.* **2016**, 49, 714-723.
74. Armitage, B. A. *Curr. Opin. Chem. Biol.* **2011**, 15, 806-812.
75. Robertson, K. L.; Yu, L.; Armitage, B. A.; Lopez, A. J.; Peteanu, L. A. *Biochemistry* **2006**, 45, 6066-6074.
76. Kolevzon, N.; Hashoul, D.; Naik, S.; Rubinstein, A.; Yavin, E. *Chem. Commun.* **2016**, 52, 2405-2407.
77. Peled, I.; Yavin, E. *ACS Omega* **2018**, 3, 3813-3818.

Chapter 2

Sequence- and structure-specific probing of RNAs by short nucleobase-modified dsRNA-binding PNAs (dbPNAs) incorporating a fluorescent light-up uracil analog

2.1 Introduction

Recent advances in the discovery and understanding of complex RNA structures, dynamics, and functions have dramatically changed our view of RNA's role in biology. RNA has been found to play increasingly important and diverse regulatory and catalytic roles in many biological processes, ranging from replication, transcription, splicing, to translation.¹ RNAs form complex structures containing both single-stranded (ss) and double-stranded (ds) regions critical for their diverse regulatory and catalytic functions and biological applications. The emerging RNA sequence and structure databases provide the foundation for developing RNA-binding ligands through the recognition of RNA sequence and structure.^{2,3}

Many attempts have been reported on the development of small molecule ligands for probing and targeting non-Watson–Crick regions in nucleic acids.⁴⁻⁸ Targeting RNA structures with small molecules, however, are often limited for the probing of the ssRNA or non-Watson–Crick regions. Insertions of artificial non-Watson–Crick structures into the RNAs of interest (ROI) facilitates indirect

detection of RNA sequences and structures.^{5-7, 9} For example, known protein binding RNA structure motifs (e.g., MS2 binding hairpins) have been attached to the ROI for the indirect detection of ROIs.⁹ Recently, tRNA-Ile2 has been inserted to ROIs for subsequent indirect chemical labelling via the covalent attachment of an agmatine analogue on tRNA-Ile2.¹⁰

Short oligonucleotides have seen a wide range of applications, including the traditional antisense-based probes and therapeutics that act through the formation of Watson–Crick duplexes with the RNA targets.¹¹⁻¹⁸ Chemically modified DNA/RNA oligonucleotides and artificial nucleic acids have been developed to enhance the binding affinity and sequence specificity of duplex formation. The strategy of using oligonucleotides to target RNA through duplex formation, however, may be limited to the unstructured or partially structured regions of RNA. The structured regions in the RNA targets may significantly slow down the binding kinetics and weaken the binding strength of oligonucleotides.¹⁹ In addition, traditional antisense compounds can be used to detect RNA sequences but not the structures of RNAs, and thus may be disadvantageous compared to the structure-based targeting strategies.

Chemical probes for the direct and selective detection of dsRNA regions of any desired sequence are underdeveloped. The strategy of directly targeting dsRNA by triplex-formation is relatively less explored mainly due to the weak binding at physiological conditions as well as preferential binding to dsDNAs over dsRNAs for the traditional triplex-forming oligonucleotides (TFOs).²⁰⁻²³ We and others have successfully developed nucleobase-modified dsRNA-binding

peptide nucleic acids (dbPNAs) to facilitate sequence- and structure-specific recognition of dsRNAs through triplex formation without disrupting the pre-formed dsRNA structures at near-physiological conditions.²⁴⁻³¹ For example, we have incorporated base T with or without modifications, modified base L, and modified base Q into dbPNAs for the recognition of A-U, G-C, and C-G pairs, respectively, in dsRNAs (**Figure 2.1**).³²⁻³⁴ Here, L and Q are dbPNA monomers carrying modified bases, which are grafted on the peptide-like neutrally-charged PNA backbone.^{24, 32-37} Remarkably, our dbPNAs can specifically bind to dsRNA over ssRNA and dsDNA regions. Hence, the nucleobase-modified dsRNA-specific dbPNAs are useful for applications in biochemistry and biotechnology, with the potential of overcoming the limitations of reported probing methodologies, *e.g.*, pH and salt dependence, preferential binding to dsDNAs over dsRNAs, and nondiscriminatory binding to dsRNAs and ssRNAs.^{20-23, 26, 38-}

40

Figure 2.1. Chemical structures of base triples and potential Watson–Crick and Watson–Crick-like base pairs formed between PNA (blue) and RNA (black). The hydrogen atoms shown in red in the base triples may be involved in van der Waals

interactions in stabilizing base triple formation. **(a-e)** Hydrogen bonding patterns for base triples formed between PNA bases (T, C, L, Q, and ^{bt}U) and RNA Watson–Crick base pairs. **(f-h)** Watson–Crick C-G pair formed between PNA C and RNA G and unstable Watson–Crick-like L-G and Q-C pairs. Our L- and Q-modified dbPNAs show selective binding to dsRNAs over ssRNAs because potential PNA-RNA duplexes are significantly destabilized with unstable Watson–Crick-like L-G and Q-G pairs as well as a weak Q-C pair.^{24, 32-36}

Figure 2.2. RNA, DNA, and PNA molecules studied in this chapter. **(a-e)** Model RNA and DNA hairpins. **(f)** Model ssRNA rHP1-ss. **(g-q)** PNAs studied in this chapter. The letter “X” shown in red represents the modified residue ^{bt}U. **(r-s)** PNA·RNA₂ triplexes formed by PNAs P5-TXT and P5t-TXTT targeting rHP1 and rHP3, respectively.

Here, we incorporated into a series of dbPNAs an A-U pair-recognizing chemically-modified nucleobase, 5-benzothiophene uracil (^{bt}U) (**Figures 2.1e and 2.2**),⁴¹⁻⁴⁴ followed by binding assays including fluorescence titration and

non-denaturing polyacrylamide gel electrophoresis (PAGE). The novel ^{bt}U-modified dsRNA-specific dbPNAs show fluorescent light-up response upon binding to targeted dsRNAs through PNA·RNA₂ triplex formation in a sequence-specific manner at 200 mM NaCl, pH 7.5. The non-denaturing PAGE data suggest that the incorporation of ^{bt}U into dbPNAs does not affect the binding affinity toward dsRNA significantly in most cases. The blue fluorescence emission of ^{bt}U-modified dbPNAs is sequence- and structure-specifically enhanced by dsRNAs, but not ssRNAs or DNAs, at a near-physiological condition (200 mM NaCl, pH 7.5).

2.2 Materials and Methods

2.2.1 Computational modeling

The PNA·RNA₂ triplexes containing ^{bt}U were modelled based on a previously modelled PNA-RNA₂ triplex structure by Molecular Dynamics (MD) simulation.³⁴ ^{bt}U base was optimized using Gaussian09 (Gaussian, Inc., Wallingford CT, USA) under B3LYP/6-31G* scheme. The benzothiophene ring and uracil ring are not completely coplanar in the optimized structure, consistent with the previously reported result for a thiophene uracil analog.⁴⁵ The optimized ^{bt}U structure was then incorporated at various positions into the original triplex structure modelled by us previously.³⁴ Specifically, a T base in the parental triplex was replaced with a ^{bt}U, with the uracil ring of ^{bt}U superimposed with T. The stacking patterns were generated using Discovery studio (Dassault Systèmes BIOVIA, San Diego CA, USA).

2.2.2 General experimental methods

The starting materials, required reagents, and all the solvents were purchased from the available commercial sources and used without further purification. Unless otherwise mentioned, the reactions were performed using the HPLC grade solvents at room temperature. Thin layer chromatography (TLC) on aluminum sheets silica gel 60 F254 (Merck) and liquid chromatography-mass spectrometry with electrospray ionization source (LCMS-ESI) were used to monitor the progress of the reactions. Column chromatographic purification was done for all backbone ester compounds using DCM/methanol or ethyl acetate/hexane solvent systems on flash silica gel 230–400 mesh size. The final monomer and the intermediates were characterized by high-resolution mass spectrometry (electron ionization) (HRMS-EI), and ^1H and ^{13}C NMR spectra on a 300 MHz, 400 MHz (100 MHz, ^{13}C), or 500 MHz Bruker spectrometer. The NMR chemical shifts (δ) are expressed in parts per million (ppm). The residual solvent peaks were used as references for the ^1H (chloroform- d : 7.26; dimethyl sulfoxide- d_6 : 2.50) and ^{13}C (chloroform- d : 77.0; dimethyl sulfoxide- d_6 : 39.5) NMR spectra.

2.2.2.1 Synthesis of fluorescent turn-on $^{\text{bt}}\text{U}$ PNA monomer

A detailed synthetic protocol was shown in **Scheme 2.1** and recorded in **Appendix Figure 1.1**. A slightly modified synthetic procedure was followed to obtain the required $^{\text{bt}}\text{U}$ PNA monomer (**6**).^{34, 44}

Synthesis of *ethyl 2-(5-iodo-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetate* (**2**). A dry 100 ml round-bottom flask (RBF) was charged with 5-iodouracil (12.6

mmol) and K_2CO_3 (11.34 mmol) in dry DMF (15 ml). The reaction mixture was cooled to 0 °C and a solution of ethyl bromoacetate (11.34 mmol) in dry DMF (15 ml) was added dropwise for about 2h to prevent the formation of di-substituted by-product at N3 and N1. After the dropwise addition was completed, the reaction mixture was slowly brought to room temperature and left stirring overnight. The excess DMF

Scheme 2.1. Reagents and conditions for PNA ⁵U monomer synthesis: **(a)** Ethyl bromoacetate, K_2CO_3 , DMF, 0°C- room temperature (rt), 8h, 74%. **(b)** LiOH, THF:H₂O, 2h, rt, 90%. **(c)** ethyl (2-((tert-butoxycarbonyl)amino)ethyl)glycinate, HBTU, NMM, DMF, 6h, 76%. **(d)** Stille coupling, Benzo[b]thiophen-2-yl tributylstannane, $Pd(PPh_3)_2Cl_2$, Dioxane, 24h, 90 °C, 78%. **(e)** LiOH, THF:H₂O, 2h, 88%. **(f)** 50% TFA in DCM, rt, 1h, 86%.

was removed under reduced pressure. The crude product was then extracted with ethyl acetate and dried over anhydrous Na_2SO_4 . The crude product was then column purified using DCM/MeOH (10:0.1, v/v) eluent system to yield

compound **2** in 74%. ¹H NMR (500 MHz, CDCl₃): δH 1.29-1.33 (t, 3H), 4.24-4.29 (q, 2H), 4.46 (s, 2H), 7.57 (s, 1H), 8.79 (br, 1H). HRMS (EI): *m/z*: calculated for C₈H₁₀IN₂O₄ [M-H]⁻ 324.9685, found: 324.9681.

Synthesis of *(5-Iodo-2,4-dioxo-3,4-dihydro-1(2H)-pyrimidinyl)acetic acid (3)*. Ethyl 2-(5-iodo-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetate (6.30 mmol) was dissolved in THF (20 ml) and 1 M lithium hydroxide solution was added at room temperature. The reaction mixture was stirred for 2h and acidified to pH 4-5 with Dowex H⁺ resin followed by the evaporation of THF. The compound **3** was obtained as a white solid with 90% of isolated yield. The ¹H NMR spectra were found to be identical as those reported.³⁴ ¹H NMR (500 MHz, DMSO-d₆): δH 4.41 (s, 2H), 8.21 (s, 1H), 11.77 (s, 1H), 13.22 (s, 1H); HRMS (EI): *m/z*: calculated for C₆H₄IN₂O₄ [M-H]⁻ 294.9216, found: 294.9211.

Synthesis of *Ethyl N-(2-((tert-Boc)aminoethyl)-N-(2-(5-iodo-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetyl)glycinate (4)*. Compound **3** (4.0 mmol) and HBTU (4.1 mmol) were added into a dry 100 mL RBF with a magnetic stir bar and kept under vacuum for 0.5 h. The contents were dissolved in anhydrous DMF (5 mL) followed by the addition of NMM (*N*-methylmorpholine) (10.0 mmol) under N₂ atmosphere. The mixture was stirred at room temperature for 15 mins to activate the acid. Ethyl *N*-(2-Boc-aminoethyl)glycinate (4.1 mmol) dissolved in anhydrous DMF (2 mL) was added dropwise to the above reaction mixture. The reaction was left stirring at room temperature for 6 h and the reaction progress was monitored by TLC and liquid chromatography-mass spectrometry with electrospray ionization source (LCMS-ESI). After the completion of

reaction, the product was extracted using ethyl acetate, washed with ice cold water, and saturated aqueous NaHCO₃ and KHSO₄ solution. The combined organic layer was treated with brine and dried over Na₂SO₄. The crude was subjected to flash chromatography using MeOH/DCM (v/v;10/0.1) to yield compound **4** in 76%. The compound exists as two rotamers. The chemical shifts for the peaks of the minor rotamer are given in brackets. ¹H NMR (500 MHz, CDCl₃): δH 1.26-1.33 (m, 3H), 1.45 (1.43) (s, 9H), 3.26-3.34 (m, 2H), 3.53-3.54 (m, 2H), 4.08 (4.11) (s, 2H), 4.13-4.28 (m, 2H), 4.67 (4.52) (s, 2H), 5.71 (br, 1H), 7.64 (7.69) (s, 1H), 9.95 (br, 1H); ¹³C NMR (100 MHz, CDCl₃): δC 14.15, 28.46 (28.51), 38.75, 48.36 (48.16), 48.73, 49.14, 61.84 (62.45), 68.19, 80.12, 149.73, 150.96, 156.28, 160.95 (162.82), 167.11 (167.49), 169.68 (169.45); HRMS (EI): *m/z*: calculated for C₁₇H₂₆IN₄O₇ [M+H]⁺ 525.0846, found: 525.0853.

Synthesis of *ethyl N-(2-(5-(benzo[b]thiophen-2-yl)-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetyl)-N-(2-((tert-butoxycarbonyl)amino)ethyl)glycinate (5)*. Stille coupling of benzo[b]thiophen-2-yltributylstannane (4.5 mmol) with **4** (1.5 mmol) in the presence of Pd(PPh₃)₂Cl₂ catalyst (0.3 mmol) in dioxane (30 ml) for 24h at 90 °C under N₂ results in precursor benzothiophene ester **5**. The reaction was confirmed by TLC and liquid chromatography-mass spectrometry with electrospray ionization source (LCMS-ESI). The dioxane was removed using rotavapor, followed by the addition of 30 ml of ethyl acetate filtration through celite to remove inorganic impurities, and washing twice with the ethyl acetate solvent (15 ml each time). The organic layer was treated with water (45 ml ×2), half saturated NaHCO₃ solution (45ml ×2), and half saturated KHSO₄ solution (45 ml ×2) followed by brine (45 ml ×2) and dried over

anhydrous Na₂SO₄. The crude was column purified using MeOH/DCM (v/v;10/0.1) to obtain compound **5** in 78% yield. The compound exists as two rotamers. The chemical shifts for the peaks of the minor rotamer are given in brackets. ¹H NMR (500 MHz, [D₆]DMSO): δ = 11.81 (br, 1H), 8.28(ma.) and 8.23 (mi.) (s, 1H), 7.92 (d, *J*=7.6 Hz, 1H), 7.79 (d, *J*=6.4 Hz, 2H), 7.36–7.29 (m,2H), 6.95 (ma.) and 6.74 (mi.) (t, *J*=5.2 Hz, 1H), 4.86 (ma.) and 4.68 (mi.) (s, 2H), 4.34–4.04(m, 4H), 3.45–3.35 (m, 2H), 3.24–3.19 (ma.) and 3.07–3.03 (mi.) (m, 2H), 1.39 (ma.) and 1.36 (mi.) (s, 9H), 1.26 (mi.) and 1.18 (ma.) (t, *J*=7.0 Hz, 3H) ppm; ¹³C NMR (100 MHz, [D₆]DMSO): δ = 169.2 (mi.) and 168.9 (ma.), 167.4 (mi.) and 167.1 (ma.), 161.7, 155.8 (ma.) and 155.6 (mi.), 149.9, 143.5, 139.1, 138.6 and 134.9, 124.4, 124.0, 123.0, 122.0, 119.2, 107.1 (ma.) and 107.0(mi.), 78.0 (ma.) and 77.7 (mi.), 61.2 (mi.) and 60.5 (ma.), 48.2, 47.9, 47.0, 38.2, 28.1, 14.0 ppm; HRMS: Calcd. for C₂₅H₃₀N₄O₇SNa: 553.1733 [M+Na]⁺; found: 553.1727.

Synthesis of *N*-(2-(5-(benzo[*b*]thiophen-2-yl)-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)acetyl)-*N*-(2-((*tert*-butoxycarbonyl)amino)ethyl)glycine (**6**). Ethyl *N*-(2-(5-(benzo[*b*]thiophen-2-yl)-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)acetyl)-*N*-(2-((*tert*-butoxycarbonyl)amino)ethyl)glycinate **5** (1.00 mmol) was dissolved in THF (10 mL) and kept under 0 °C. 1 M lithium hydroxide was added into the reaction medium and stirred at 0 °C for 2h. Upon evaporation of THF and acidification to pH 2 with Dowex H⁺ resin, the desired monomer **6** was obtained as white solid in 88% yield. The compound exists as two rotamers. The chemical shifts for the peaks of the minor rotamer are given in brackets. ¹H NMR (300 MHz, [D₆]DMSO): δ = 11.82 (ma.) and 11.79 (mi.) (br, 1H), 8.29 (ma.) and

8.24 (mi.) (s, 1 H), 7.93–7.90 (m, 1H), 7.81–7.77 (m, 2H), 7.65–7.55 (mi.) 7.37–7.27 (ma.) (m, 2H), 6.97 (ma.) and 6.79 (mi.) (t, J=5.0 Hz, 1H), 4.85 (ma.) and 4.67 (mi.) (s, 2H), 4.14 (mi.) and 4.00 (ma.) (s, 2H), 3.46–3.41 (ma.) and 3.36–3.32 (mi.) (m, 2H), 3.25–3.21 (ma.) and 3.08–3.04 (mi.) (m, 2H), 1.39 (ma.) and 1.36 (mi.) (s, 9H) ppm; ^{13}C NMR (100 MHz, [D6]DMSO): δ = 170.8 (mi.) 170.4 (ma.), 167.3 (mi.) and 166.9 (ma.), 161.7, 155.8 (ma.) and 155.6 (mi.), 149.9, 143.6, 139.1, 138.6(mi.) and 138.5 (ma.), 134.9, 124.5 (ma.) and 124.4 (mi.), 124.1 (ma.) and 124.0 (mi.), 123.0, 122.0, 119.2 (ma.) and 119.1 (mi.), 107.1 and 107.0, 78.1 (ma.) and 77.7 (mi.), 48.3(mi.) and 48.2 (ma.), 47.8, 47.1, 46.9, 38.2 (ma.) and 37.6 (mi.), 28.2 ppm HRMS: Calcd. for $\text{C}_{23}\text{H}_{26}\text{N}_4\text{O}_7\text{SNa}$: 525.1420 $[\text{M}+\text{Na}]^+$; found: 525.1422.

2.2.3 Solid phase synthesis of PNA oligomers

The unmodified N-(2-aminoethyl)glycine PNA (aegPNA) monomers were ordered from ASM Research Chemicals and used as such. The L and Q monomers were synthesized as per the previously reported procedures.^{32, 33} Methylbenzhydrylamine hydrochloride (MBHA·HCl) polystyrene-based resin was chosen as a solid support for the synthesis of PNA oligomers. Capping was done using acetic anhydride. The initial loading value of 1.5–1.7 mmol/g was brought down to 0.35 mmol/g. The oligomer synthesis was performed using Boc-strategy, in the presence of (Benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyBop) and N,N-Diisopropylethylamine (DIPEA) as the coupling agents. After the coupling was done for all the unmodified and modified PNA monomers, the high-low trifluoroacetic acid (TFA)-

trifluoromethanesulfonic acid (TFMSA) method was used for the final cleavage of the oligomer from the resin. The crude PNA oligomers were obtained by precipitation using ice cold diethyl ether and dissolved in de-ionized water. The PNA oligomers were purified by reverse-phase high performance liquid chromatography (RP-HPLC) using water-CH₃CN-0.1% TFA as the mobile phase. The oligomers were characterized by matrix assisted laser desorption/ionization-time of flight (MALDI-TOF) spectrometry with α -cyano-4-hydroxycinnamic acid (CHCA) as sample crystallization matrix and shown in **Appendix Table A1, Figure A1.2.**

2.2.4 Determination of extinction co-efficient

Using the Shimadzu UV-2550 ultraviolet–visible spectrophotometer the absorbance was recorded for a known concentration of PNA monomer dissolved in de-ionized water at 25 °C. The optical path length of the eight-microcell cuvette was 1 cm. Using Beer-Lambert’s law the extinction coefficients of PNA T, C, L, Q, and ^{bt}U at 260 nm were estimated to be 8.8, 7.3, 7.3, 7.3, and 13.1 mM⁻¹cm⁻¹, respectively.

2.2.5 Fluorescence binding study

The fluorescence excitation and emission spectra were recorded on a Varian Cary Eclipse fluorescence spectrophotometer, with 1 cm square quartz cuvette at room temperature. The required RNA and DNA oligonucleotides purified by RP-HPLC were obtained from Sigma-Aldrich, Singapore. All the fluorescence experiments were carried with the ^{bt}U monomer and PNAs at 1 μ M and RNAs

or DNAs at varied concentrations (0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10, and 20 μM). RNA/DNA alone samples at 10 μM were used as controls. The samples were prepared in 70 μL incubation buffer containing 200 mM NaCl, 0.5 mM EDTA, 20mM HEPES, pH 7.5. The samples were prepared by snap cooling of RNA/DNA hairpins from 95 $^{\circ}\text{C}$ to 0 $^{\circ}\text{C}$, followed by the addition of PNA oligomers. The solutions were then vortexed for 10 min with a subsequent slow cooling from 65 $^{\circ}\text{C}$ to room temperature and incubation at 4 $^{\circ}\text{C}$ overnight. ssRNAs and PNAs were annealed by slow cooling from 95 $^{\circ}\text{C}$ to room temperature, followed by overnight incubation at 4 $^{\circ}\text{C}$. All the emission spectra were recorded at room temperature over a wavelength range of 350–600 nm, at a fixed excitation wavelength of 319 nm with a slit width 10 nm.

2.2.6 Estimation of the Limit of Detection (LOD)

Using Varian Cary Eclipse fluorescence spectrophotometer, with 1 cm square quartz cuvette at room temperature the limit of detection (LOD) was obtained using PNA IR-1X designed to target influenza panhandle hairpin. Three independent experiments were performed. The plots were drawn for the fluorescence intensity of 1 μM IR-1X at 466 nm as a function of PH-v and PH-v-ss at concentrations of 0, 0.01, 0.02, 0.05, 0.1, 0.2, and 0.5 μM in 70 μL buffer containing 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5. The standard deviation (σ) was attained using the equation $\sigma = \sqrt{\sum(x_i - \mu)^2/N}$. Where, x_i is each value from the population, μ is population mean and N is size of the population. The LOD was calculated based on the standard deviation (σ) of the IR-1X alone sample and the slope of the calibration plot: $\text{LOD} = 3\sigma/\text{slope}$.

2.2.7 Nondenaturing polyacrylamide gel electrophoresis

The nondenaturing (12%) polyacrylamide gel electrophoresis (PAGE) experiments were performed using an incubation buffer containing 200mM NaCl, 0.5mM EDTA, 20mM HEPES (pH 7.5). All samples were prepared with a loading volume of 25 μ L. The RNA/DNA hairpins were snap cooled from 95 °C by directly dipping the sample tubes into ice bath. The annealing with PNAs was done by slow cooling from 65 °C to room temperature followed by incubation at 4 °C overnight. Right before the loading of the samples into the wells, 35% glycerol (20% of the total loaded volume) was added followed by vortexing for a few seconds. The gels were run at 4 °C for 6 h using a running buffer 1 $\times$ TBE (Tris–Borate–EDTA), pH 8.3. The voltage was set to be 250V for all the experiments. Finally, the gels were imaged using a Typhoon imager after ethidium bromide staining for 30 min.

2.2.8 Circular dichroism spectroscopic characterization

The circular dichroism (CD) spectra were recorded using a JASCO model J-1500-150 spectropolarimeter (Japan Spectroscopic Co. Ltd., Tokyo, Japan). A quartz cell having a path length 10 mm was used to load the samples. The RNA and PNA concentrations were maintained at 2 and 4 μ M, respectively. RNA hairpins were snap cooled from 95 °C to 4 °C in an incubation buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0). PNAs were annealed with the RNAs by slow cooling from 65 °C to room temperature, followed by incubation overnight at 4 °C. The data were collected at room temperature with a wavelength

range of 200-320 nm. The data were averaged from six scans with a scan rate at 20 nm/minute.

2.2.9 Cell line and virus

Madin-Darby Canine Kidney (MDCK) cell line was obtained from American Type Culture Collection. MDCK cells were grown in Dulbecco's Modified Eagle's Medium (DMEM, Gibco, Grand Island, NY, USA) containing antibiotics (penicillin 100U/mL and streptomycin 100µg/mL) and 10% fetal bovine serum at 37 °C in 5% CO₂. Highly pathogenic influenza A human H5N1 (A/Indonesia/CDC1031/2007) was used for the experiment.

2.2.10 Viral infection and extraction of total RNA

MDCK cells were seeded in 12-well plate at a cell density of 1×10^5 cells/mL for 24 hrs. Prior to infection, sub-confluent monolayers of MDCK cells were gently washed two times with Phosphate Buffer Saline (PBS) and infected with H5N1 virus at a multiplicity of infection (MOI) of 1 for 1h at 37 °C. Cells were washed twice with PBS and 1 mL of DMEM medium containing 2% FBS was added. Mock-infected cells were used as the negative control. The culture plates were incubated at 37 °C for 24 h. Total RNA was extracted from H5N1 virus-infected or mock-infected MDCK cells using mirVana miRNA isolation kit (Thermo Fisher Scientific, MA, USA), according to the manufacturer's protocol. The concentration and quality of total RNA were determined by using the NanoDrop 2000 UV-Vis Spectrophotometer (Thermo Fisher Scientific).

2.3 Results and Discussion

2.3.1 Modelling of PNA·RNA₂ triplexes containing ^{bt}U·A-U triples

The properties of a fluorescent molecular probe may be significantly affected by the molecular interactions such as hydrogen bonding, stacking, intramolecular and intermolecular electron transfers, and solvent properties.⁴⁶⁻⁵² We modelled the PNA·RNA₂ triplex structures (**Figure 2.3**) based on the model triplexes we previously generated.³⁴ A favorable interaction between the sulfur atom in benzothiophene and 4 carbonyl oxygen in uracil⁵³⁻⁵⁷ may restrict the free rotation of the single bond connecting benzothiophene and uracil to have a close to planar but not completely planar structure with a dihedral angle between the two rings of about 18°. ⁴⁵ Through replacing a T residue with a ^{bt}U residue, we obtained the base stacking patterns of ^{bt}U with various nearest neighbors (**Figures 2.2 and Figure 2.3**). The modelled structures show that, except for the ^{bt}U^{bt}U and T^{bt}U (from N-terminus to C-terminus) steps, there is essentially no stacking interactions involving the benzothiophene group in ^{bt}U. Thus, PNA·RNA₂ triplex formation results in relatively unstacked microenvironments for ^{bt}U positioned in varied sequence contexts, which may allow ^{bt}U to serve as a structure-sensing fluorescent probe.

Figure 2.3. Potential stacking patterns of PNA·RNA-RNA base triples involving a ^{bt}U·A-U triple based on the modeled triplex structures. (a) ^{bt}U·A-U/L·G-C. (b) L·G-C/^{bt}U·A-U. (c) C·G-C/^{bt}U·A-U. (d) T·A-U/^{bt}U·A-U. (e) ^{bt}U·A-U/T·A-U. (f) ^{bt}U·A-U/^{bt}U·A-U. The numbers beside each triple are the positions (from N-terminus to C-terminus) of the PNA (NH₂-Lys-TLTLTTTL-CONH₂) with T and L replaced by ^{bt}U and C, respectively, at various positions. The presence of an L^{bt}U (panel b) or a C^{bt}U (panel c) step in a dbPNA may cause nonradiative electron transfer, resulting in no fluorescence change or fluorescence quenching, respectively, upon binding to dsRNA.

2.3.2 Fluorescence of ^{bt}U monomer and ^{bt}U-containing PNA oligomers

The PNA ^{bt}U monomer was synthesized based on a reported protocol with minor modifications (Scheme 2.1).^{34, 44} A series of PNAs without or with ^{bt}U incorporated at various positions were synthesized (Figure 2.2g-q, Table A1),

followed by fluorescence characterization and other binding assays. Interestingly, incorporation of ^{bt}U into PNA oligomers causes a significant reduction of fluorescence intensity (compared to ^{bt}U monomer, compound 6 in **Scheme 2.1**) with no significant shift in fluorescence emission peak (~468 nm) (**Figure 2.4a,b**).⁴⁴ Notably, deprotected ^{bt}U monomer (dp-^{bt}U, compound 7 in **Scheme 2.1**) also shows significantly reduced fluorescence (**Figure 2.4a**), indicating that the Boc protecting group on ^{bt}U monomer may interact with ^{bt}U base resulting in stabilizing a more emissive conformation of ^{bt}U. Restriction of the rotation of the C–C single bond connecting benzothiophene and uracil rings may also reduce the non-radiative decay resulting fluorescence enhancement.^{40, 41, 52} Indeed, we observed that increasing solvent viscosity by increasing the percentage of glycerol in methanol/glycerol mixtures results in enhanced fluorescence of ^{bt}U monomer (**Figure 2.5**).

In general, the reduction of fluorescence intensity is relatively more significant for the sequences (see **Figure 2.2**) with ^{bt}U positioned next to L (P5-LXT, P5-LXL, P5-XL, P5-LXXT and P5-LXLXT), compared to those with ^{bt}U non-adjacent to L (P5-TXT, P5t-TXTT, P5-TTXTT, and P4-CXT) (**Figures 2.4b and 2.6a**). Interestingly, the fluorescence intensity reduction is most significant for the sequences with two ^{bt}U residues incorporated adjacent to each other or separated by an L residue (P5-LXXT and P5-LXLXT) (**Figure 2.6a**). It is thus likely that stacking of ^{bt}U with L or another ^{bt}U within the PNA oligomers causes the fluorescence quenching.

Figure 2.4. Fluorescence emission of monomers, oligomers, and triplex titration experiments for various PNAs binding to rHP1. **(a)** Excitation and emission spectra of ^{bt}U , deprotected ^{bt}U ($dp-^{bt}U$), and P5-TXT. **(b)** Fluorescence emission spectra for ^{bt}U , $dp-^{bt}U$, and various oligomers. **(c, f and i)** Fluorescence intensity enhances upon the formation of rHP1·P5-XL, rHP1·P5-TXT and rHP1·P5-LXLT triplexes, respectively. **(d, e and h)** Fluorescence intensity remains unchanged for rHP1·P5-LXL, rHP1·P5-LXT and rHP1·P5-LXLXT triplexes. **(g and l)** Fluorescence quenching upon the titration of rHP1 into P4-CXT, indicating ^{bt}U

fluorescence is sensitive to triplex structure formation and protonation of a C base in an adjacent C⁺·G-C triple. (j) Determination of K_d by fluorescence titrations for P5-TXT, P5-XL, and P5-LXXT binding to rHP1. Fluorescence intensities at 474 nm were used for P5-LXXT. (k) ^{bt}U fluorescence intensity of P5-LXT, P5-LXL and P5-LXLXT remains relatively constant upon the titrations of rHP1. The monomers and oligomers were at 1 μ M. The RNA hairpin alone samples were at 10 μ M. All the samples were prepared with 1 μ M PNA with 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10, and 20 μ M rHP1 in 70 μ L of buffer of 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5. rHP1 binds to P5-XL and P5-LXXT (see **Figure 2.8**) and induces fluorescence enhancement of ^{bt}U. rHP1 binds to P5-LXL (see **Figure 2.8**) with no significant change in ^{bt}U fluorescence. rHP1 shows weak binding to P5-LXLXT (see **Figure 2.8**) and induces no significant change in ^{bt}U fluorescence. The K_d values were fit to the equation: $Y = Y_0 + (B/2/R_0) (R_0 + X + K_d - ((R_0 + X + K_d)^2 - 4R_0X)^{1/2})$. R_0 is the PNA concentration (1 μ M). Y_0 and B are the minimum and maximum fluorescence intensities, respectively. X is the total RNA concentration and K_d is the dissociation constant.

Figure 2.5. Fluorescence emission spectra of ^{bt}U monomer in methanol-glycerol solvent system.⁵⁸ ^{bt}U monomer was at 1 μ M in 70 μ L with 100:0, 70:30, 50:50, 20:80, 10:90, 5:95, 0:100 volume ratios of methanol: glycerol, respectively. Increasing viscosity results in an increase of the fluorescence intensity of ^{bt}U monomer, indicating molecular rotor behavior of benzothiophene attached to uracil base.

Figure 2.6. Comparison of ^{bt}U fluorescence in varied sequence and structural environment. **(a)** Fluorescence emission intensities (at 466 nm) for PNA monomers ^{bt}U (protected) and dp-^{bt}U (deprotected), RNA hairpins, and various PNA oligomers without and with 10 μ M RNA hairpins. Incorporating ^{bt}U non-adjacent to L and C may be ideal for fluorescent light-up upon binding target RNA hairpins. **(b)** Sequence- and structure-specific fluorescence light-up of ^{bt}U-containing dbPNAs upon binding to dsRNAs. The monomers and oligomers were at 1 μ M. The RNA hairpin alone samples were at 10 μ M. for the potential triplexes and duplexes, the RNA/DNA hairpins and rHP1-ss are at 10 μ M. The buffer contains 200 mM NaCl, 0.5 mM EDTA, 20mM HEPES, pH 7.5 with a volume of 70 μ L.

We hypothesize that intramolecular twisting between benzothiophene and uracil rings may favor charge transfer from the benzothiophene ring to uracil ring resulting in an enhanced fluorescence emission due to a Twisted Intramolecular Charge Transfer (TICT) effect.^{59, 60} However, the twisting angle may be altered with ^{bt}U incorporated into the PNA oligomers due to the stacking with adjacent and non-adjacent bases within a flexible single-stranded PNA.^{45, 61} Having two ^{bt}U residues within single-stranded PNAs (P5-LXXT and P5-LXLXT) may severely quench the fluorescence by disfavoring emissive twisted conformations through extensive stacking interactions as evident in the ^{bt}U absorption spectrum. Other mechanisms such as self-quenching of fluorophores may also contribute to the relatively low fluorescence for the two PNAs incorporating two ^{bt}U residues.⁶² The fluorescence emission wavelength shows no observable shift for the PNAs containing two ^{bt}U residues compared to ^{bt}U monomer and the PNAs containing a single ^{bt}U residue (**Figures 2.4b and 2.6a**), suggesting that there is no excimer formation.⁶³⁻⁶⁵

2.3.3 Fluorescence detection of duplex RNAs through triplex formation

Our previous non-denaturing PAGE study shows that PNA P5 (see **Figure 2.2g**) can bind to rHP1 ($K_d = 0.2 \mu\text{M}$ at 200 mM NaCl, pH 7.5, see **Table 2.1**).^{32, 33} Incorporating ^{bt}U residues into P5 is not expected to cause significant changes in the base-base hydrogen bonding interactions, but may alter stacking and hydration environment (**Figure 2.1a,e and Figure 2.3**). Our circular dichroism (CD) spectroscopic data confirm the triplex formation for P5-TXT with target RNA duplex rHP1 (**Figure 2.7**). We further quantified the triplex formation stabilities of the PNAs derived from P5 and other PNAs by non-denaturing

PAGE (**Figure 2.8**). The PAGE data show that replacing T with ^btU only slightly destabilizes the triplex formation in most cases (**Table 2.1, Figures 2.8 and 2.9**). PNA P5-LXXT with two ^btU units shows slightly more weakened binding to rHP1 with a K_d of 1.3 μ M. Surprisingly, we observe a significantly weakened binding of P5-LXLXT ($K_d > 20$ μ M), which might be due to the disruption of the hydration spine in the major groove by two non-adjacent relatively hydrophobic and bulky benzothiophene rings.^{66, 67}

Table 2.1. K_d values (in μ M) for binding of PNAs to model RNA hairpin rHP1, obtained by fluorescence titration and non-denaturing PAGE experiments at pH 7.5.^a

PNAs	rHP1	
	Fluorescence Titration	PAGE
P5 ^b	-	0.2±0.1
P5-XL	1.1±0.6	0.6±0.2
P5-LXL	NFE	0.6±0.2
P5-LXT	NFE	0.9±0.3
P5-TXT	0.9±0.1	0.3±0.1
P4 ^b	-	3.2±0.7
P4-CXT	0.4±0.1 (FQ)	1.2±0.2
P5-LXLXT	NFE	>20
P5-LXXT	0.6±0.3	1.3±0.3
P5t-TXTTT	NFE	>20
P5t-TTXXT	NFE	-

^aIncubation buffer used: 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES at pH 7.5. “-” indicates that the data were not obtained in this study. “NFE” indicates no fluorescence enhancement observed. “FQ” indicates fluorescence quenching observed. The fluorescence and PAGE experiments were carried out at room temperature and 4 °C, respectively. ^bThe PNA oligomer sequences were previously studied.^{32, 33}

Figure 2.7. CD spectroscopic characterization of triplex formation. rHP1·P5-TXT triplex formation is evidenced by a red shift in CD upon the addition of P5-TXT into rHP1. Adding P5-TXT into rHP2 shows no change in CD. The concentrations of the rHPs and P5-TXT are 2 and 4 μ M, respectively. The spectra were taken in a buffer containing 100 mM NaCl, 10 mM sodium phosphate, pH 7.0, at room temperature.

Figure 2.8. 12wt% nondenaturing polyacrylamide gel electrophoresis (PAGE). (a-g) PAGE characterization for the formation of rHP1·P5-XL, rHP1·P5-LXL, rHP1·P5-LXT, rHP1·P5-TXT, rHP1·P5-CXT, rHP1·P5-LXLXT, and rHP1·P5-LXXT triplexes. (h-k) PAGE characterization for the binding of P5t-TXTTT to RNA hairpins. rHP3 shows the strongest binding, confirming specific recognition an A-U pair in rHP3 by ^{bt}U. (l) PAGE characterization for the binding of dbPNA L4-LXT to HIV-HP. All the samples were prepared using a buffer containing 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5. RNA hairpins loaded were at 0.25 μ M in 25 μ L buffer. The PNA concentrations in lanes from left to right are 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 1, 1.6, 2, 4, 10, 16, and 20 μ M, respectively. The running buffer used was 1 $\times$ TBE at 250 V. The K_d values obtained from PAGE are consistent with those obtained from fluorescence titrations (**Table 2.1**).

Figure 2.9. K_d determination by non-denaturing PAGE for various PNAs binding to rHP1 and rHP3 at 200 mM NaCl, pH 7.5. (a-i) K_d determination for the formation of rHP1·P5-XL, rHP1·P5-LXL, rHP1·P5-LXT, rHP1·P5-TXT, rHP1·P4-CXT, rHP1·P5-LXLXT, rHP1·P5-LXXT, rHP3·P5t-TXTT, and rHP4·P5t-TXTTT triplexes. The fraction of triplex formation as calculated based on equation $Y = I_{\text{triplex}} / (I_{\text{triplex}} + aI_{\text{duplex}})$. The band intensities were normalized according to $a = I_{\text{triplex max}} / I_{\text{duplex max}}$. $I_{\text{duplex max}}$ is the band intensity for hairpin alone without the addition of PNA. $I_{\text{triplex max}}$ is the triplex band intensity with the highest band intensity. Whereas, the normalized duplex intensity (see panels f and g) was obtained by normalizing the duplex bands: $I_{\text{duplex}} / I_{\text{duplex max}}$. The data were fit to the equation: $Y = Y_0 + (B/2/R_0)(R_0 + X + K_d - ((R_0 + X + K_d)^2 - 4R_0X)^{1/2})$ where R_0 is the RNA hairpin concentration (0.25 μ M). Y_0 and B are

the minimum and maximum fractions of triplex formation, respectively. X is the total PNA concentration and K_d is the dissociation constant.

Our fluorescence titration data (**Figures 2.4c-l and 2.6**) reveal that the ^{bt}U-modified PNAs show distinct responses to the titration of rHP1 at a near-physiological condition (200 mM NaCl, pH 7.5). For the PNA oligomers P5-TXT, P5-XL and P5-LXXT, ^{bt}U fluorescence is enhanced by 3-5 fold (**Figures 2.4 and 2.6**). However, no significant change in fluorescence was observed for P5-LXT, P5-LXL, and P5-LXLXT. We speculate that an L^{bt}U step may have non-radiative decay through electron transfer from ^{bt}U to L upon triplex formation, resulting no net change in fluorescence. Surprisingly, fluorescence quenching was observed for P4-CXT upon binding to rHP1 (**Figure 2.4g,l**). It is probable that C⁺·G-C triple formation adjacent to the ^{bt}U·A-U triple facilitates non-radiative electron transfer from ^{bt}U to protonated C⁺ in the C⁺·G-C triple. Thus, ^{bt}U maybe useful for probing the protonation states of adjacent residues. Thus, triplex formation of the PNA strands results in changes in the stacking, hydrogen bonding, and local environment for ^{bt}U and other PNA residues, and incorporating ^{bt}U non-adjacent to L or C residues may be useful for the detection of triplex formation through dsRNA binding-induced fluorescence enhancement.

Interestingly, ^{bt}U fluorescence for P5-LXXT increases by 6 fold upon the titration of rHP1, with a slightly red shifted emission peak at 474 nm (**Figure 2.4i,j**). The second ^{bt}U upstream of T may contribute to the fluorescence enhancement upon triplex formation. The stacking geometry of the two adjacent ^{bt}U residues in single-stranded PNA is expected to deviate from that in a PNA·RNA₂ triplex.

The result further suggests that ^{bt}U fluorescence is sensitive to stacking environment. However, no significant change in ^{bt}U fluorescence was observed for P5-LXLXT upon the addition of rHP1 (**Figures 2.4h**), which may be due to the quenching of neighboring L residues as well as a significantly weakened binding ($K_d > 20 \mu\text{M}$) to rHP1 (**Figure 2.9f**).

The modelled stacking geometries reveal that benzothiophene ring is not strongly involved in stacking interactions in most cases (**Figure 2.3**), suggesting that ^{bt}U can be in a non-planar conformation within a triplex allowing ^{bt}U fluorescence enhancement by a TICT effect.^{59, 60} In addition, fluorescence enhancement may result from a restricted rotation^{40, 41, 52} of the single bond connecting benzothiophene and uracil rings due to non-covalent sulfur-oxygen interaction between the sulfur atom in benzothiophene and 4 carbonyl oxygen in uracil⁵³⁻⁵⁷ and a hydration spine^{66, 67} potentially present in the groove of a triplex. Clearly, incorporating ^{bt}U into dbPNAs can be useful for monitoring the triplex formation.

P5t-TXTTT shows ^{bt}U fluorescence enhancement upon binding to rHP3 ($K_d = 1.8 \mu\text{M}$, at 200 mM NaCl, pH 7.5) with ^{bt}U recognizing a Watson–Crick A-U base pair (**Figures 2.2s**), but no fluorescence response upon the titration of the hairpins with the A-U pair in rHP3 replaced with G-C (rHP1), C-G (rHP2), or U-A (rHP4) (**Figures 2.6 and 2.10, Table 2.2**). The K_d values obtained by fluorescence titrations show a similar trend as those by non-denaturing PAGE (**Table 2.1, Figures 2.8-2.10**).

We further employed ^{bt}U-containing PNAs P5-TTXXTT and P5-TXT to study the molecular recognition specificity of T and L bases, respectively. We observe high sequence specificity for T in P5-TTXXTT (the T residue right upstream of ^{bt}U) in recognizing A-U (rHP3) over G-C (rHP1), C-G (rHP2), and U-A (rHP4) (**Figure 2.11**). Similarly, the middle L residue in P5-TXT recognizes a G-C pair in rHP1 with high sequence specificity (**Figures 2.4f,j, 2.6 and 2.12**). Thus, our data suggest that incorporation of a ^{bt}U residue does not affect the sequence specificity of T and L bases in recognizing A-U and G-C pairs, respectively.

Table 2.2. Sequence- and structure-specific fluorescence light-up of ^{bt}U-containing dbPNAs upon binding to dsRNAs allows for the quantification of K_d values (in μM).^a

	P5-XL	P5-TXT	P5t-TXTTT	P5t-TTXXTT
rHP1	1.1±0.6	0.9±0.1	NFE	NFE
rHP2	-	NFE	NFE	NFE
rHP3	-	NFE	1.8±0.1	2.3±1.3
rHP4	-	NFE	NFE	NFE
dHP1	NFE	NFE	-	-
rHP1-ss	NFE	NFE	-	-

^aIncubation buffer used: 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES at pH 7.5. “-” indicates that the data were not obtained in this study. “NFE” indicates that no significant fluorescence enhancement was observed.

Figure 2.10. Sequence-specific binding of rHP3 to P5t-TXTTT results in an enhancement of ^{bt}U fluorescence at 200 mM NaCl, pH 7.5. (a) RNA structures used to study the sequence specificity of ^{bt}U in P5t-TXTTT. (b-f) ^{bt}U fluorescence in P5t-TXTTT is specifically enhanced by rHP3 but not by rHP1 (G-C), rHP2 (C-G), or rHP4 (U-A). All the samples were prepared with 1 μM PNA with 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10, and 20 μM RNAs in 70 μL of buffer of 200 mM NaCl, 0.5 mM EDTA, 20mM HEPES, pH 7.5.

Figure 2.11. Sequence-specific binding of rHP3 to P5t-TTXXTT results in an enhancement of ^{bt}U fluorescence at 200 mM NaCl, pH 7.5. (a) RNA structures used to study the sequence specificity of ^{bt}U in P5t-TTXXTT. (b-f) ^{bt}U fluorescence in P5t-TTXXTT is specifically enhanced by rHP3 but not by rHP1 (G-C), rHP2 (C-G), or rHP4 (U-A). The data suggest that T base in P5t-TTXXTT specifically recognizes an A-U pair. All the samples were prepared with 1 μ M PNA with 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10, and 20 μ M RNAs in 70 μ L of buffer of 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5.

Figure 2.12. Sequence-specific binding of rHP1 to P5-TXT results in an enhancement of ^{bt}U fluorescence at 200 mM NaCl, pH 7.5 (see **Figures 2.4f and 2.6b**). (a) RNA structures used to study the sequence specificity of ^{bt}U in P5-TXT. (b-e) ^{bt}U fluorescence in P5-TXT is specifically enhanced by rHP1 (see **Figures 2.4f and 2.6b**) but not by rHP2 (C-G), rHP3 (A-U), or rHP4 (U-A). The data suggest that L base in P5-TXT specifically recognizes a G-C pair. All the samples were prepared with 1 μ M PNA with 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1,

2, 5, 10, and 20 μ M RNAs in 70 μ L of buffer of 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5.

2.3.4. Fluorescence detection of dsRNA over ssRNA and dsDNA

We further employed 32 P-labeled PNA for characterizing the structural specificity. Significantly, P5-TXT shows fluorescence enhancement upon binding to rHP1, but no appreciable fluorescence changes upon adding dHP1 and ssRNA (**Figures 2.4f,j, 2.6b and 2.13**). Similarly, P5-XL with 32 P positioned at N-terminus shows fluorescence enhancement upon the titration of rHP1, but no response toward dHP1 and ssRNA (**Figures 2.4c,j, 2.6b and 2.14**). The binding selectivity for dsRNAs over dsDNAs is probably because a PNA has a geometrical compatibility of with dsRNA with a relatively deep and narrow major groove and reduced helical twisting as compare to dsDNA. Incorporation of L residues results in enhanced recognition of G-C pairs with minimal pH dependence in dsRNA through L-G-C triple formation, but destabilized Watson-Crick-like G-L pair formation due to a steric clash between the sulfur atom and the amine group of guanine (**Figure 2.1c,g**).³² The structural specificity observed here is consistent with our previous PAGE and thermal melting studies for dbPNAs.^{32, 33}

Figure 2.13. Structure-specific binding of rHP1 to P5-TXT results in an enhancement of ^{bt}U fluorescence at 200 mM NaCl, pH 7.5 (see **Figures 2.4f and 2.6b**). (a) RNA and DNA structures used to study the structure specificity of ^{bt}U in P5-TXT. (b-e) ^{bt}U fluorescence in P5-TXT is specifically enhanced by rHP1 (see **2.4f and 2.6b**) but not by dHP1 and rHP1-ss. All the samples were prepared with 1 μ M PNA with 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10, and 20 μ M RNAs in 70 μ L of buffer of 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5.

Figure 2.14. Structure-specific binding of rHP1 to P5-XL results in an enhancement of ^{bt}U fluorescence at 200 mM NaCl, pH 7.5 (see **Figures 2.4c and 2.6b**). (a) RNA and DNA structures used to study the structure specificity of ^{bt}U in P5-XL. (b-e) ^{bt}U fluorescence in P5-XL is specifically enhanced by rHP1 (see **2.4c and 2.6b**) but not by dHP1 or rHP1-ss. All the samples were prepared with 1 μM PNA with 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10, and 20 μM RNAs in 70 μL of buffer of 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5.

2.3.5 Developing dsRNA-binding PNAs for the direct detection of influenza viral panhandle structure and HIV-1 –1 ribosomal frameshift inducing RNA hairpin

Influenza A virions contain eight segments of negative-strand viral RNA (vRNA) with highly conserved panhandle duplex structure formed between the first 13 nucleotides at the 5' end and the last 12 nucleotides at the 3' end.⁶⁸⁻⁷¹ To probe and detect the influenza viral panhandle structure, we designed a 10-mer dbPNA IR-1X containing T, L, Q, and ^{bt}U (**Figure 2.15**) by replacing a T with a ^{bt}U in a

previously developed influenza A panhandle structure-binding dbPNA IR-1.⁷² The fluorescence intensity of dbPNA IR-1X increases by 4-fold upon triplex formation with a K_d of 0.7 μ M at 200 mM NaCl, pH 7.5, consistent with the PAGE result (**Figure 2.16a,d**) and our previously reported data for IR-1.⁷² We estimated the limit of detection of IR-1X for the purified influenza panhandle hairpin to be submicromolar (126 nM, **Figure 2.17**). Notably, a ssRNA fragment, PH-v-ss RNA (**Figure 2.15b**), however, does not induce fluorescence change in IR-1X (**Figure 2.16b**), presumably because IR-1X does not bind strongly to PH-v-ss due to the presence of unstable Watson–Crick-like L-G and Q-C pairs (**Figure 2.1g,h**). We further tested if dbPNA IR-1X can be used to detect the presence of influenza A viral RNAs. We obtained total RNA extracts (with the removal of proteins and cell debris) from influenza H5N1-infected and mock-infected cells (approximately 0.4×10^6 cells/well) for the test (**Figure 2.18**). We indeed observed an enhancement ($\sim 30\%$) of fluorescence intensity of 2 μ M of IR-1X at 466 nm upon the addition of ~ 750 pg/ μ L of total RNA extract from infected cells at 200 mM NaCl, pH 7.5. Importantly, the total RNA extract from mock-infected cells does not affect the fluorescence intensity of IR-1X, with the concentration up to ~ 3600 pg/ μ L. The fluorescence of IR-1X is quenched with the total RNA above ~ 3600 pg/ μ L for both the total RNAs extracted from infected and mock-infected cells, probably due to a nonspecific effect. The detection sensitivity of IR-1X is not as sensitive as the reported PCR-based methods.^{73, 74} However, through further modification of ^{bt}U and improvement in RNA binding, we expect that highly sensitive and user-friendly probes may be developed in the future. Taken together, our data show that dbPNA IR-1X would

be potentially useful for structure-specifically detecting the influenza panhandle structure.

We next made a 6-mer dbPNA L4-TXL for the targeting of HIV-1 –1 ribosomal frameshift-inducing RNA hairpin (HIV-HP, **Figure 2.15e**).⁷⁵⁻⁷⁸ The peak fluorescence intensity of PNA L4-TXL increases by about 4 fold upon triplex formation with HIV-HP (**Figure 2.16c**). The K_d value at 200 mM NaCl, pH 7.5 obtained by fluorescence titration ($1.8 \pm 0.9 \mu\text{M}$) is in good agreement with that obtained by PAGE ($2.3 \pm 0.4 \mu\text{M}$) (**Figures 2.16, 2.19-2.20**). Interestingly, HIV-HP binds to L4-LXT (**Figure 2.8l, 2.15g, and 2.20**) with no significant change in ^{bt}U fluorescence, which is in contrast to L4-TXL. The result further confirms that an L^{bt}U step in PNA may cause ^{bt}U fluorescence to be insensitive to triplex formation. Clearly, one may avoid incorporating ^{bt}U downstream of L in a fluorescent light-up probe for probing dsRNAs.

Figure 2.15. Viral RNAs targeted by ^{bt}U-labelled dbPNAs. **(a-d)** Influenza panhandle viral RNA duplex (PH-v). PH-v but not the ssRNA segment of PH-v (PH-v-ss) is targeted by ^{bt}U-labelled dbPNA IR-1X. **(e-h)** HIV-1 frameshift inducing RNA (HIV-HP). HIV-HP binds to dbPNAs L4-TXL and L4-LXT.

Figure 2.16. Fluorescence and non-denaturing PAGE studies for PH-v·IR-1X triplex and HIV-HP·L4-TXL triplex. **(a,b)** Fluorescence titration of IR-1X with PH-v and PH-v-ss. ^{bt}U fluorescence of IR-1X is enhanced by PH-v but not by PH-v-ss. **(c)** Fluorescence titration of L4-TXL with HIV-HP. The concentration of PNA was fixed at 1 μ M in 70 μ L buffer. The measurement was done with an excitation wavelength at 319 nm at room temperature. **(d,e)** PAGE characterization of the formation of PH-v·IR-1X triplex and HIV-HP·L4-TXL triplex. The PNA concentrations in lanes from left to right are 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 1, 1.6, 2, 4, 10, 16, and 20 μ M, respectively. RNA concentration

was fixed at 0.25 μM in 25 μL in a buffer containing 200 mM NaCl, 0.5 mM EDTA, 20mM HEPES, pH 7.5. The running buffer used was 1 $\times$ TBE.

Figure 2.17. Estimation of the Limit of Detection (LOD). The plots shown are the fluorescence intensity of 1 μM IR-1X at 466 nm as a function of PH-v and PH-v-ss at concentrations of 0, 0.01, 0.02, 0.05, 0.1, 0.2, and 0.5 μM in 70 μL buffer containing 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5. By putting the values from three individual experiments in the equation $[\sigma = \sqrt{\sum(x_i - \mu)^2/N}]$ yields standard deviation (σ) and is found to be 1.0685. The obtained slope for the calibration plot of IR-1X binding to PH-v is 25.458. The LOD for PH-v is estimated to be 0.126 μM . The LOD was obtained based on the standard deviation (σ) of the IR-1X alone sample and the slope of the calibration plot: $\text{LOD} = 3\sigma/\text{slope}$.

Figure 2.18. Test of IR-1X in detecting viral RNA extracted from infected cells.

The stock concentration of the total RNAs from mock-infected and infected cells are 374 and 360 ng/μL, respectively. The samples were prepared by diluting the stocks in 70 μL buffer containing 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5, followed by mixing with 2 μM IR-1X at room temperature. The samples were incubated at room temperature for 3 hours before the fluorescence measurement. The excitation wavelength is 319 nm. The fluorescence intensity of IR-1X is not affected by the addition of the total RNA (up to ~3600 pg/μL) from mock-infected cells (**a,c**). The fluorescence intensity of IR-1X is enhanced upon the addition of ~750 pg/μL total RNA from infected cells (**b,d**). The fluorescence of IR-1X is quenched with the total RNA above ~3600 pg/μL for both infected and mock-infected cells, probably due to a nonspecific effect. Representative data are shown.

Figure 2.19. Fluorescence binding studies for influenza viral panhandle structure (PH-v) and HIV-1 -1 ribosomal frameshift inducing RNA hairpin (HIV-HP). (a,b) ^{bt}U fluorescence of IR-1X is enhanced by PH-v but not by PH-v-ss. Sequence- and structure-specific fluorescence enhancement allows for the determination of the K_d value for IR-1X binding to PH-v. (c) Sequence- and structure-specific fluorescence enhancement allows for the determination of the K_d value for L4-TXL binding to HIV-HP. (d,e) HIV-HP binds to L4-LXT (see Figure 2.8I) with no significant change in ^{bt}U fluorescence. The result further confirms that the presence of an L^{bt}U step in PNA may cause ^{bt}U fluorescence to be insensitive to triplex formation.

Figure 2.20. Determination of K_d by non-denaturing PAGE for PNAs binding to viral RNA hairpins. **(a)** K_d determination for IR-1X binding influenza viral panhandle structure (PH-v). **(b-c)** K_d determination for PNAs L4-TXL and L4-LXT binding to HIV-1 -1 ribosomal frameshift inducing RNA hairpin (HIV-HP).

2.4 Conclusions

In summary, we have developed a convenient biosensing strategy for probing and detecting RNA duplex structures in a buffer mimicking physiological conditions in a sequence- and structure-specific manner, by incorporating a ^{bt}U residue into our nucleobase-modified dbPNAs. The incorporation of our previously developed PNA L base allows enhanced triplex formation and reduced duplex formation at physiological pH,^{24, 32, 33} without the need to reduce the pH.²⁶ The combination of the microenvironment-sensitive fluorescent light-up ^{bt}U probe and our unique dbPNAs allows one to probe and detect sequence specifically dsRNAs over dsDNAs and ssRNAs. Our data suggest that one may have ^{bt}U positioned non-adjacent to L and C residues for achieving optimal fluorescence light-up response. We further demonstrate the application of the ^{bt}U -incorporated dbPNAs for the recognition of influenza viral panhandle RNA duplex and HIV-1 frameshift-inducing hairpin at a near-physical condition (200 mM NaCl, pH 7.5).

To the best of our knowledge, this is the first report involving the fluorescent detection of biological RNAs using dbPNAs at physiological pH conditions.

Compared to targeting RNAs through duplex formation, targeting the preformed dsRNA regions through RNA structure-specific triplex formation provides a complementary RNA probing strategy and is more advantageous for the structural specificity. Although there is an impressively large Stokes shift (~ 147 nm, **Figure 2.4a**) probably due to a TICT effect, a blue fluorescent emission and a relatively moderate fluorescence enhancement of our probes may limit the practical application in the detection of real biological samples. Our work clearly shows that it is possible to develop fluorescent light-up dbPNAs with further improved sensitivity and red-shifted fluorescence signals as user-friendly tools. For example, fluorescent light-up dbPNAs incorporated with a cross-linking moiety facilitates enhanced binding and biosensing sensitivity, allowing for the development of useful biochemical, diagnostic, and therapeutic agents through the direct recognition of the sequence and secondary structure of RNAs^{74, 79, 80}. Our probes can also be used for mapping complex secondary structures of newly discovered RNAs.^{77, 81-83}

References

1. Cech, T. R.; Steitz, J. A. *Cell* **2014**, 157, 77-94.
2. Cannone, J. J.; Subramanian, S.; Schnare, M. N.; Collett, J. R.; D'Souza, L. M.; Du, Y.; Feng, B.; Lin, N.; Madabusi, L. V.; Muller, K. M.; Pande, N.; Shang, Z.; Yu, N.; Gutell, R. R. *BMC Bioinf.* **2002**, 3, 2.
3. Kozomara, A.; Griffiths-Jones, S. *Nucleic Acids Res.* **2014**, 42, D68-73.
4. Disney, M. D.; Yildirim, I.; Childs-Disney, J. L. *Org. Biomol. Chem.* **2014**, 12, 1029-1039.
5. Warner, K. D.; Chen, M. C.; Song, W.; Strack, R. L.; Thorn, A.; Jaffrey, S. R.; Ferre-D'Amare, A. R. *Nat. Struct. Mol. Biol.* **2014**, 21, 658-663.
6. Huang, H.; Suslov, N. B.; Li, N. S.; Shelke, S. A.; Evans, M. E.; Koldobskaya, Y.; Rice, P. A.; Piccirilli, J. A. *Nat. Chem. Biol.* **2014**, 10, 686-691.
7. Braselmann, E.; Wierzba, A. J.; Polaski, J. T.; Chrominski, M.; Holmes, Z. E.; Hung, S. T.; Batan, D.; Wheeler, J. R.; Parker, R.; Jimenez, R.; Gryko, D.; Batey, R. T.; Palmer, A. E. *Nat. Chem. Biol.* **2018**, 14, 964-971.
8. Zhang, L.; Liu, H.; Shao, Y.; Lin, C.; Jia, H.; Chen, G.; Yang, D.; Wang, Y. *Anal. Chem.* **2015**, 87, 730-737.
9. Katz, Z. B.; English, B. P.; Lionnet, T.; Yoon, Y. J.; Monnier, N.; Ovryn, B.; Bathe, M.; Singer, R. H. *eLife* **2016**, 5, e10415.
10. Li, F.; Dong, J.; Hu, X.; Gong, W.; Li, J.; Shen, J.; Tian, H.; Wang, J. *Angew. Chem. Int. Ed.* **2015**, 54, 4597-4602.
11. Testa, S. M.; Disney, M. D.; Turner, D. H.; Kierzek, R. *Biochemistry* **1999**, 38, 16655-16662.
12. Kierzek, E.; Kierzek, R.; Moss, W. N.; Christensen, S. M.; Eickbush, T. H.; Turner, D. H. *Nucleic Acids Res.* **2008**, 36, 1770-1782.
13. Gude, L.; Berkovitch, S. S.; Santos, W. L.; Kutchukian, P. S.; Pawloski, A. R.; Kuimelis, R.; McGall, G.; Verdine, G. L. *J. Biol. Chem.* **2012**, 287, 18843-18853.
14. Kole, R.; Krainer, A. R.; Altman, S. *Nat. rev. Drug discovery* **2012**, 11, 125-140.
15. Moffitt, J. R.; Bambah-Mukku, D.; Eichhorn, S. W.; Vaughn, E.; Shekhar, K.; Perez, J. D.; Rubinstein, N. D.; Hao, J.; Regev, A.; Dulac, C.; Zhuang, X. *Science* **2018**, 362, eaau5324.
16. Shen, X.; Corey, D. R. *Nucleic Acids Res.* **2018**, 46, 1584-1600.
17. Crooke, S. T.; Witztum, J. L.; Bennett, C. F.; Baker, B. F. *Cell Metab.* **2018**, 27, 714-739.
18. Torres, A. G.; Fabani, M. M.; Vigorito, E.; Williams, D.; Al-Obaidi, N.; Wojciechowski, F.; Hudson, R. H.; Seitz, O.; Gait, M. J. *Nucleic Acids Res.* **2012**, 40, 2152-2167.
19. Armitage, B. A. *Drug Discovery Today* **2003**, 8, 222-228.
20. Han, H.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, 90, 3806-3810.
21. Roberts, R. W.; Crothers, D. M. *Science* **1992**, 258, 1463-6.
22. Escude, C.; Francois, J. C.; Sun, J. S.; Ott, G.; Sprinzl, M.; Garestier, T.; Helene, C. *Nucleic Acids Res.* **1993**, 21, 5547-5553.
23. Zhou, Y.; Kierzek, E.; Loo, Z. P.; Antonio, M.; Yau, Y. H.; Chuah, Y. W.; Geifman-Shochat, S.; Kierzek, R.; Chen, G. *Nucleic Acids Res.* **2013**, 41, 6664-6673.
24. Devi, G.; Zhou, Y.; Zhong, Z.; Toh, D.-F. K.; Chen, G. *Wiley Interdiscip. Rev. RNA* **2015**, 6, 111-128.
25. Virta, P. M.; Tahtinen, V.; Granqvist, L.; Murtola, M.; Stromberg, R. *Chem. Eur. J.* **2017**, 23, 7113-7124.
26. Sato, T.; Sato, Y.; Nishizawa, S. *J. Am. Chem. Soc.* **2016**, 138, 9397-9400.

27. Endoh, T.; Hnedzko, D.; Rozners, E.; Sugimoto, N. *Angew. Chem. Int. Ed.* **2016**, 55, 899-903.
28. Li, M.; Zengeya, T.; Rozners, E. *J. Am. Chem. Soc.* **2010**, 132, 8676-8681.
29. Kim, K. T.; Chang, D. L.; Winssinger, N. *Helv. Chim. Acta* **2018**, 101, e1700295.
30. Hnedzko, D.; McGee, D. W.; Karamitas, Y. A.; Rozners, E. *RNA* **2017**, 23, 58-69.
31. Zengeya, T.; Gupta, P.; Rozners, E. *Angew. Chem. Int. Ed.* **2012**, 51, 12593-12596.
32. Devi, G.; Yuan, Z.; Lu, Y.; Zhao, Y.; Chen, G. *Nucleic Acids Res.* **2014**, 42, 4008-4018.
33. Toh, D.-F. K.; Devi, G.; Patil, K. M.; Qu, Q.; Maraswami, M.; Xiao, Y.; Loh, T.-P.; Zhao, Y.; Chen, G. *Nucleic Acids Res.* **2016**, 44, 9071-9082.
34. Patil, K. M.; Toh, D.-F. K.; Yuan, Z.; Meng, Z.; Shu, Z.; Zhang, H.; Ong, A. A. L.; Krishna, M. S.; Lu, L.; Lu, Y.; Chen, G. *Nucleic Acids Res.* **2018**, 46, 7506-7521.
35. Patil, K. M.; Chen, G. *Modified Nucleic Acids in Biology and Medicine*, Erdmann, V. A.; Jurga, S.; Barciszewski, J., Eds. Springer: 2016; pp 299-317.
36. Toh, D.-F. K.; Patil, K. M.; Chen, G. *J Visualized Exp.* **2017**, 127, e56221.
37. Puah, R. Y.; Jia, H.; Maraswami, M.; Kaixin Toh, D. F.; Ero, R.; Yang, L.; Patil, K. M.; Lerk Ong, A. A.; Krishna, M. S.; Sun, R.; Tong, C.; Huang, M.; Chen, X.; Loh, T. P.; Gao, Y. G.; Liu, D. X.; Chen, G. *Biochemistry* **2018**, 57, 149-159.
38. Cheruiyot, S. K.; Rozners, E. *ChemBioChem* **2016**, 17, 1558-1562.
39. D'Agata, R.; Giuffrida, M. C.; Spoto, G. *Molecules* **2017**, 22, 1951.
40. Vilaivan, T. *Beilstein J. Org. Chem.* **2018**, 14, 253-281.
41. Srivatsan, S. G.; Tor, Y. *Nat. Protoc.* **2007**, 2, 1547-1555.
42. Sinkeldam, R. W.; Wheat, A. J.; Boyaci, H.; Tor, Y. *ChemPhysChem* **2011**, 12, 567-570.
43. Cservenyi, T. Z.; Van Riesen, A. J.; Berger, F. D.; Desoky, A.; Manderville, R. A. *ACS Chem. Biol.* **2016**, 11, 2576-2582.
44. Sabale, P. M.; Srivatsan, S. G. *ChemBioChem* **2016**, 17, 1665-1673.
45. Endoh, T.; Annoni, C.; Hnedzko, D.; Rozners, E.; Sugimoto, N. *Phys. Chem. Chem. Phys.* **2016**, 18, 32002-32006.
46. Seidel, C. A. M.; Schulz, A.; Sauer, M. H. M. *J. Phys. Chem.* **1996**, 100, 5541-5553.
47. Torimura, M.; Kurata, S.; Yamada, K.; Yokomaku, T.; Kamagata, Y.; Kanagawa, T.; Kurane, R. *Anal. Sci.* **2001**, 17, 155-160.
48. Rachofsky, E. L.; Osman, R.; Ross, J. B. *Biochemistry* **2001**, 40, 946-956.
49. Jean, J. M.; Hall, K. B. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, 98, 37-41.
50. Lee, B. J.; Barch, M.; Castner, E. W., Jr.; Volker, J.; Breslauer, K. J. *Biochemistry* **2007**, 46, 10756-10766.
51. Xu, W.; Chan, K. M.; Kool, E. T. *Nat. Chem.* **2017**, 9, 1043-1055.
52. Sinkeldam, R. W.; Greco, N. J.; Tor, Y. *Chem. Rev.* **2010**, 110, 2579-2619.
53. Hansell, D. P.; Robinson, K. J.; Wallis, J. D.; Povey, D. C. *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.* **1996**, 52, 136-139.
54. Peacock, H.; Luo, J. H.; Yamashita, T.; Luccarelli, J.; Thompson, S.; Hamilton, A. D. *Chem. Sci.* **2016**, 7, 6435-6439.
55. Creuven, I.; Norberg, B.; Olivier, A.; Evrard, C.; Evrard, G.; Wigerinck, P.; Herdewijn, P.; Durant, F. *J. Chem. Crystallogr.* **1996**, 26, 777-789.
56. Beno, B. R.; Yeung, K. S.; Bartberger, M. D.; Pennington, L. D.; Meanwell, N. A. *J. Med. Chem.* **2015**, 58, 4383-4438.
57. Collins, I.; Moyes, C.; Davey, W. B.; Rowley, M.; Bromidge, F. A.; Quirk, K.; Atack, J. R.; McKernan, R. M.; Thompson, S. A.; Wafford, K.; Dawson, G. R.; Pike, A.;

- Sohal, B.; Tsou, N. N.; Ball, R. G.; Castro, J. L. *J. Med. Chem.* **2002**, 45, 1887-1900.
58. Burns, D. D.; Teppang, K. L.; Lee, R. W.; Lokensgard, M. E.; Purse, B. W. *J. Am. Chem. Soc.* **2017**, 139, 1372-1375.
 59. Grabowski, Z. R.; Rotkiewicz, K.; Rettig, W. *Chem. Rev.* **2003**, 103, 3899-4032.
 60. Teran, N. B.; He, G. S.; Baev, A.; Shi, Y.; Swihart, M. T.; Prasad, P. N.; Marks, T. J.; Reynolds, J. R. *J. Am. Chem. Soc.* **2016**, 138, 6975-6984.
 61. Nuthanakanti, A.; Srivatsan, S. G. *Nanoscale* **2016**, 8, 3607-3619.
 62. Noe, M. S.; Sinkeldam, R. W.; Tor, Y. *J. Org. Chem.* **2013**, 78, 8123-8128.
 63. Teo, Y. N.; Kool, E. T. *Chem. Rev.* **2012**, 112, 4221-4245.
 64. Wang, S.; Guo, J.; Ono, T.; Kool, E. T. *Angew. Chem. Int. Ed.* **2012**, 51, 7176-7180.
 65. Singh, V.; Wang, S.; Chan, K. M.; Clark, S. A.; Kool, E. T. *J. Am. Chem. Soc.* **2013**, 135, 6184-6191.
 66. Radhakrishnan, I.; Patel, D. J. *Biochemistry* **1994**, 33, 11405-11416.
 67. McDermott, M. L.; Vanselous, H.; Corcelli, S. A.; Petersen, P. B. *ACS Cent. Sci.* **2017**, 3, 708-714.
 68. Lenartowicz, E.; Keszy, J.; Ruskowska, A.; Soszynska-Jozwiak, M.; Michalak, P.; Moss, W. N.; Turner, D. H.; Kierzek, R.; Kierzek, E. *PLoS One* **2016**, 11, e0148281.
 69. Hsu, M. T.; Parvin, J. D.; Gupta, S.; Krystal, M.; Palese, P. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, 84, 8140-8144.
 70. Bae, S. H.; Cheong, H. K.; Lee, J. H.; Cheong, C.; Kainosho, M.; Choi, B. S. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, 98, 10602-10607.
 71. Gultyaev, A. P.; Fouchier, R. A. M.; Olsthoorn, R. C. L. *Int. Rev. Immunol.* **2010**, 29, 533-556.
 72. Keszy, J.; Patil, K. M.; Kumar, S. M.; Shu, Z.; Yee, Y. H.; Zimmermann, L.; Ong, A. A. L.; Toh, D.-F. K.; Krishna, M. S.; Yang, L.; Decout, J. L.; Luo, D.; Prabakaran, M.; Chen, G.; Kierzek, E. *Bioconjugate Chem.* **2019**, 30, 931-943.
 73. Velumani, S.; Du, Q.; Fenner, B. J.; Prabakaran, M.; Wee, L. C.; Nuo, L. Y.; Kwang, J. *J. Virol. Methods* **2008**, 147, 219-225.
 74. Medina, R. A.; Rojas, M.; Tuin, A.; Huff, S.; Ferres, M.; Martinez-Valdebenito, C.; Godoy, P.; Garcia-Sastre, A.; Fofanov, Y.; SantaLucia, J., Jr. *J. Clin. Microbiol.* **2011**, 49, 335-344.
 75. Baril, M.; Dulude, D.; Gendron, K.; Lemay, G.; Brakier-Gingras, L. *RNA* **2003**, 9, 1246-1253.
 76. Garcia-Miranda, P.; Becker, J. T.; Benner, B. E.; Blume, A.; Sherer, N. M.; Butcher, S. E. *J. Virol.* **2016**, 90, 6906-6917.
 77. Watts, J. M.; Dang, K. K.; Gorelick, R. J.; Leonard, C. W.; Bess, J. W., Jr.; Swanstrom, R.; Burch, C. L.; Weeks, K. M. *Nature* **2009**, 460, 711-716.
 78. Gareiss, P. C.; Miller, B. L. *Curr. Opin. Investig. Drugs* **2009**, 10, 121-128.
 79. Barczak, A. K.; Gomez, J. E.; Kaufmann, B. B.; Hinson, E. R.; Cosimi, L.; Borowsky, M. L.; Onderdonk, A. B.; Stanley, S. A.; Kaur, D.; Bryant, K. F.; Knipe, D. M.; Sloutsky, A.; Hung, D. T. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, 109, 6217-6222.
 80. Moss, W. N.; Priore, S. F.; Turner, D. H. *RNA* **2011**, 17, 991-1011.
 81. Mathews, D. H.; Moss, W. N.; Turner, D. H. *Cold Spring Harb Perspect. Biol.* **2010**, 2, a003665.
 82. Kladwang, W.; VanLang, C. C.; Cordero, P.; Das, R. *Nat. Chem.* **2011**, 3, 954-962.
 83. Eubanks, C. S.; Hargrove, A. E. *Biochemistry* **2019**, 58, 199-213.

Chapter 3

Incorporating G-C Pair-Recognizing Guanidinium into PNAs for Sequence and Structure Specific Recognition of dsRNAs over dsDNAs and ssRNAs

3.1 Introduction

RNAs form diverse structures for interacting with metabolites and proteins, for catalysing RNA splicing and protein synthesis, and for regulating gene expression. Thus, developing chemical compounds for targeting desired RNA structures has great potential in biomedical applications.¹⁻⁴ The strategy of directly targeting double-stranded RNA (dsRNA) by triplex-formation is relatively underexplored mainly due to the weak binding at physiological conditions for the traditional triplex-forming oligonucleotides (TFOs).⁵⁻⁸

Nucleobase-modified dsRNA-binding peptide nucleic acids (dbPNAs) are promising in targeting dsRNAs in a sequence and structure specific manner (**Figures 3.1 and 3.2**).⁹⁻¹⁸ Substituting C with a modified nucleobase may alleviate the pH dependence of G-C pair recognition through the formation of $C^+ \cdot G-C$ triples (**Figures 3.1b-d and 3.2a,b**).^{19, 20} For example, 2-aminopyridine (M, **Figure 3.2b**) has a monomer pK_a of 6.7 for the protonation of M to form M^+ and M-modified PNAs show enhanced recognition of G-C pairs in dsRNAs.^{15, 21} M^+ residues also enhance cellular uptake of PNAs.^{16, 17}

Pseudoisocytosine (J, **Figure 3.2a**) is a modified neutral nucleobase for the recognition of G-C pairs.²² J-modified PNAs have reduced pH dependence but relatively weakened recognition of G-C pair through Hoogsteen pairing.^{11, 15} We reported the incorporation of thio-pseudoisocytosine (L, **Figure 3.1c,d**) into dbPNAs for the recognition of G-C pairs in dsRNAs.¹¹ The unique and selective base pairing (Hoogsteen over Watson–Crick, **Figures 3.1c,d and 3.2c**) and stacking properties of the L base result in L-modified dbPNAs to be selective in binding to dsRNAs over ssRNAs and dsDNAs at near-physiological conditions. However, the synthesis of PNA L monomer is nontrivial and incorporation of L into PNAs does not improve cell permeability.

Figure 3.1. Base triples and a potential Watson–Crick-like base pair. The guanidinium moiety serving as hydrogen bond donors is highlighted in red. The hydrogen atoms shown in gray may be involved in van der Waals interactions in enhancing base triple formation (see panels **c** and **d**). (**a–g**) Hydrogen bonding patterns for base triples formed between PNA bases (T, C, L, Q, and R) and RNA

Watson–Crick base pairs. **(h)** Unstable Watson–Crick-like R-G pair. **(i-l)** Modeled base triple structures of L·G-C and R·G-C. Two populations of structures with the formation of at least one hydrogen bond are shown.

Guanidinium (protonated form of guanidine) has unique hydrogen bonding and interesting stacking properties.^{23, 24} Guanidinium (arginine residue) is present in proteins often involved in hydrogen bonding and stacking interactions with nucleic acids.²⁵⁻²⁸ Various compounds incorporating guanidinium moiety show important biological activities.²⁹⁻³⁷ Incorporation of the guanidinium moiety facilitates enhanced cellular uptake of various cargos.^{13, 23, 38-49} Free guanidinium has recently been found to be tightly bound in riboswitches.^{50, 51}

Guanidine modified aromatic moieties have been incorporated into artificial nucleic acids for the recognition of single-stranded G residues as G-clamps^{52, 53} or C-G pairs through Q·C-G triple (**Figure 3.1e**) or other triple formation.^{13, 14, 54-56} Our modeling studies (**Figure 3.1f,g,k,l**) suggest that guanidinium may be directly incorporated into PNAs (without an extra aromatic group) for enhanced recognition of G-C pairs in dsRNAs. We grafted guanidinium moiety into a PNA monomer (designated as R) through a simple two-carbon linker (**Figure 3.1f,g**) followed by incorporating the R residue into various dbPNAs for binding studies. We further tested the compatibility of R in combination with C, T, L, and Q residues for targeting model dsRNAs and dsRNA structures found in influenza and HIV-1 viral RNAs.

Figure 3.2. Chemical and modeled structures of base triples and unstable Watson-Crick-like base pairs formed between PNA and RNA ^{11, 13, 15}. **(a-d)** Chemical structures. **(e,f)** Modeled three-base-triple structures containing residues L and R. **(g,h)** Modeled structures involving an R derivative, R', with an extra carbon atom in the base-backbone linker. No stable coplanar structure is formed between R' and G. **(i)** A major structure during simulation containing R'. Incorporation of R' results in (i) no stable formation of a base triple involving R' and (ii) a distortion of the adjacent base triple structure.

3.2 Materials and Methods

3.2.1 Computational modeling

The PNA·dsRNA triplex containing the PNA sequence H₂N-Lys-TLTRTTTL-CONH₂ was generated based on the model for a PNA·dsRNA triplex containing the PNA sequence H₂N-Lys-TLTLTTTL-CONH₂.^{57, 58} We also replaced the two-carbon base-backbone linker in R with a three-carbon linker (R') to obtain a PNA·dsRNA triplex containing the PNA sequence H₂N-Lys-TLTR'TTTL-CONH₂. We modelled the RNA molecules with Amber force field χ OL3.^{59, 60} and the PNA residues with Amber GAFF force field⁶¹ and AM1-BCC atomic charges.⁶² TIP3P explicit solvent⁶³ and counter ions (Na⁺) were added to the simulation boxes. Next, we performed a 1000-step energy minimization using steepest descent algorithm.⁶⁴ For each model, three independent 20-ns NPT simulations were carried out with random initial velocities. The cutoff for van der Waals and electrostatics was 1.2 nm. All the simulations were performed with OpenMM simulation package.⁶⁵ The three trajectories for each system were merged by removing the first 10% of the trajectories, and then the hydrogen bonds were calculated with Mdtraj,⁶⁶ considering the distance between hydrogen and hydrogen bond acceptor within 0.25 nm and angle between hydrogen bond acceptor, hydrogen and hydrogen bond donor larger than 120 °.

3.2.2 General experimental methods

The required reagents, starting materials, and all the solvents were obtained from the available commercial vendors and used without further purification. All the

reactions were performed using the high-performance liquid chromatography (HPLC) grade solvents at room temperature unless otherwise mentioned. The reaction progress was monitored by thin layer chromatography (TLC) on aluminum sheets silica gel 60 F254 (Merck) and liquid chromatography-mass spectrometry with electrospray ionization source (LCMS-ESI). Column chromatographic purification was done for Fmoc-, t-butyl ester compound **5**, using ethyl acetate/hexane solvent systems on flash silica gel 230–400 mesh size. The intermediate compounds and the final monomer were characterized by high-resolution mass spectrometry (electron ionization) (HRMS-EI), and ^1H and ^{13}C NMR spectra on a 400 MHz (100 MHz, ^{13}C) or 500 MHz Bruker spectrometer. The NMR chemical shifts (δ) are expressed in parts per million (ppm). The residual solvent peaks were used as references for the ^1H (chloroform- d : 7.26; dimethyl sulfoxide- d_6 : 2.50) and ^{13}C (chloroform- d : 77.0; dimethyl sulfoxide- d_6 : 39.5) NMR spectra.

3.2.2.1 Synthesis of PNA monomer R

The synthetic protocol for the PNA R monomer is outlined in **Scheme 3.1** and **Appendix Figure A2.1**. The starting materials **1** and **2** were synthesized using the previously reported procedures.^{67, 68} Compound **4** was obtained by a 4-hour reaction of the di-Cbz protected pyrazole guanidine **1** with crushed 3-aminopropionic acid **3** using DIPEA in DCM at room temperature. The obtained white solid **4** was reacted with **2** in the presence of HBTU and NMM in anhydrous DMF. The reaction was carried out under nitrogen atmosphere at room temperature followed by a column purification using EtOAc/Hexane (v/v; 3/10) solvent system to give precursor t-butyl ester compound **5**. Compound **5** was

treated with 50% TFA-DCM for 1 hour, followed by the removal of the solvent with a subsequent precipitation using cold di-ethyl ether. The product was filtered, washed, and dried in vacuum to yield the required PNA R monomer **6**.

Scheme 3.1. Chemical synthesis of PNA R monomer.

Compounds (**1** and **2**) were synthesized following the reported literature procedure and their ^1H NMR spectra were found to be identical as those reported ^{67, 68}.

benzyl (Z)-[(((benzyloxy)carbonyl)amino)(1H-pyrazol-1-yl)methylene]carbamate (1). ^1H NMR (500 MHz, CDCl_3): δ_{H} 9.33 (s, 1H), 8.29 (s, 1H), 7.63 (s, 1H), 7.44-7.31 (m, 10H), 6.45 (dd, 1H), 5.24 (s, 2H), 5.21 (s, 2H); HRMS (EI): m/z : calculated for $\text{C}_{20}\text{H}_{18}\text{N}_4\text{O}_4\text{Na}$ $[\text{M}+\text{Na}]^+$ 401.1226, found: 401.1228.

tert-butyl [2-(((9H-fluoren-9-yl)methoxy)carbonyl)amino]ethyl]glycinate hydrochloride (2). ^1H NMR (500 MHz, DMSO-d_6): δ_{H} 9.26 (s, 2H), 7.90-7.88 (d, $J = 7$ Hz, 2H), 7.70-7.68 (d, $J = 7$ Hz, 2H), 7.55 (s, 1H), 7.43-7.40 (t, $J = 7$ Hz, 1H), 7.34-7.31 (t, $J = 7$ Hz, 2H), 5.75 (s, 1H), 4.34-4.32 (d, 2H), 4.24-4.21 (t, 1H), 3.87 (s, 2H), 3.37-3.32 (m, 2H), 2.99 (t, 2H), 1.45 (s, 9H); HRMS (EI): m/z : calculated for $\text{C}_{23}\text{H}_{29}\text{N}_2\text{O}_4$ $[\text{M}+\text{H}]^+$ 397.2128, found: 397.2130.

(E)-3-(2,3-bis((benzyloxy)carbonyl)guanidino)propanoic acid (**4**). Compound **1** (0.414 g, 1.09 mmol) and crushed β – alanine **3** (0.094 g, 1.06 mmol) were dissolved in 20 ml of DCM and stirred in the presence of DIPEA (1mL, 5.74 mmol) under N₂ atmosphere at room temperature for 4 hours. After the completion of the reaction as confirmed by TLC and liquid chromatography-mass spectrometry with electrospray ionization source (LCMS-ESI), the reaction mixture was extracted using excess amount of DCM and washed with chilled water. The combined organic layer was treated with brine and dried over Na₂SO₄. The colourless product, which is sparingly soluble in EtOAc, was obtained in 82% isolated yield. The compound was filtered and dried in vacuum and used as such in the next step. ¹H NMR (500 MHz, DMSO-d₆): δ_{H} 11.57 (br, 1H), 8.58 (s, 1H), 7.49-7.47 (m, 2H), 7.41-7.38 (m, 5H), 7.36-7.35 (m, 2H), 5.20 (s, 2H), 5.04 (s, 2H), 5.00 (s, 2H), 3.53-3.52 (q, 2H), 2.61 (s, 2H); ¹³C NMR (100 MHz, DMSO-d₆): δ_{C} 172.96, 171.92, 155.04, 136.81, 135.10, 134.10, 129.35, 128.92, 128.52, 128.46, 128.39, 128.31, 127.98, 127.80, 77.81, 67.59, 66.33, 36.42, 33.09; HRMS (EI): *m/z*: calculated for C₂₀H₂₀N₃O₆ [M-H]⁻ 398.1352, found: 398.1350.

tert-butyl(E)-10-[2-(((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)ethyl)-5-(((benzyloxy) carbonyl)amino]-3,9-dioxo-1-phenyl-2-oxa-4,6,10-triazadodec-4-en-12-oate (**5**). Compound **4** (4.0 mmol), HBTU (4.1 mmol), and NMM (4.1 mmol) were dissolved in anhydrous DMF (5 mL) in a 50 mL RBF, followed by the dropwise addition of compound **2** (4.4 mmol) dissolved in anhydrous DMF (2 mL) under N₂ atmosphere. The reaction mixture was stirred at room temperature for 8 hours and monitored by TLC and liquid chromatography-mass

spectrometry with electrospray ionization source (LCMS-ESI). After the completion of reaction, it was extracted using ethyl acetate, washed with ice cold water, and saturated aqueous NaHCO₃ and KHSO₄ solution. The combined organic layer was treated with brine and dried over Na₂SO₄. The crude was subjected to flash chromatography using EtOAc/Hexane (v/v; 0.3/0.7) to yield the precursor t-butyl ester compound **5** in 52%. The compound exists as two rotamers. The chemical shifts for the peaks of the minor rotamer are given in brackets. ¹H NMR (500 MHz, DMSO-d₆): δH 7.88-7.86 (8.68-8.62) (m, 3H), 7.67-7.63 (m, 2H), 7.41-7.28 (m, 13H), 5.12 (5.18) (s, 2H), 4.99 (5.01) (s, 2H), 4.29 (4.28) (s, 2H), 4.05-4.17 (m, 2H), 3.89 (s, 1H), 3.53-3.51 (m, 2H), 3.15-3.14 (3.10-3.08) (m, 2H), 2.44-2.43 (2.22-2.21) (m, 2H), 2.68-2.62 (m, 2H), 1.37 (1.36) (s, 9H); ¹³C NMR (100 MHz, DMSO-d₆): δC 171.17 (171.55), 168.37 (168.77), 162.99, 154.96, 142.54, 139.38, 137.39 (136.88), 135.12, 128.86, 128.50, 128.43, 128.28, 127.90, 127.88, 127.74, 127.23, 121.31, 119.96, 109.63, 81.50, 80.63, 67.50, 66.27, 50.54, 48.26, 31.24, 31.10, 29.80, 28.95, 27.63, 27.50; HRMS (EI): *m/z*: calculated for C₄₃H₄₈N₅O₉ [M+H]⁺ 778.3452, found: 778.3455.

(E)-10-[2-(((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)ethyl]-5-(((benzyloxy)carbonyl)amino]-3,9-dioxo-1-phenyl-2-oxa-4,6,10-triazadodec-4-en-12-oic acid (**6**). Solution of compound **5** (4.1 mmol) in 50% TFA/DCM (2 ml) was stirred at room temperature for 1 hour. After the completion of the reaction as confirmed by TLC and LCMS-ESI, the solvents were removed by blowing N₂ gas. The contents were precipitated using cold di-ethyl ether and filtered to obtain the product as white solid. The final monomer **6** was obtained in 92 % yield. The compound exists as two rotamers. The chemical shifts for the

peaks of the minor rotamer are given in brackets. ^1H NMR (500 MHz, DMSO- d_6): δ H 8.96-8.89 (m, 1H), 7.94-7.86 (m, 3H), 7.65-7.63 (m, 2H), 7.38-7.30 (m, 12H), 5.13 (5.18) (s, 2H), 5.03 (5.05) (s, 2H), 1.98 (s, 2H), 4.17-4.10 (m, 2H), 3.95 (s, 1H), 3.54 (m, 2H), 3.36-3.33 (m, 2H), 3.16-3.11 (m, 2H), 2.88-2.87 (m, 2H), 2.73-2.67 (m, 2H); ^{13}C NMR (100 MHz, DMSO- d_6): δ C 171.25 (171.69), 170.66 (171.04), 162.29, 158.54, 158.16, 156.21, 154.57, 152.41, 143.82 (143.89), 140.73, 136.62, 135.04, 128.52, 128.48, 128.32, 128.13, 128.00, 127.96, 127.84, 127.57, 127.00, 125.11, 125.00, 120.06, 67.66, 66.55, 65.38, 47.51, 46.75, 35.74, 31.25, 30.75; HRMS (EI): m/z : calculated for $\text{C}_{39}\text{H}_{38}\text{N}_5\text{O}_9$ $[\text{M}-\text{H}]^-$ 720.2670, found: 720.2668.

3.2.3 Solid phase synthesis of PNA oligomers

The unmodified N-(2-aminoethyl)glycine PNA (aegPNA) monomers were acquired from ASM Research Chemicals and used as such for the synthesis of PNA oligomers. The L and Q monomers were synthesized as per the previously reported procedures.^{11, 13} The synthesis of PNA oligomers was done on Methylbenzhydrylamine hydrochloride (MBHA·HCl) polystyrene-based resin as solid support. The initial loading value of 1.5–1.7 mmol/g was brought down to 0.35 mmol/g using acetic anhydride as capping agent. Except for the R monomer, the oligomer synthesis was performed using Boc-strategy in the presence of (Benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyBop) and N,N-Diisopropylethylamine (DIPEA) as the coupling agents. The Fmoc deprotection was done using 20% piperidine in DMF. The final cleavage of the oligomer from the resin was done using the high-low trifluoroacetic acid (TFA)-trifluoromethanesulfonic acid (TFMSA) method once

the coupling has been done for all the unmodified and modified PNA monomers. The crude PNA oligomers were precipitated with ice-cold diethyl ether and dissolved in de-ionized water. The purification of PNA oligomers was done by reverse-phase HPLC (RP-HPLC) using water-CH₃CN-0.1% TFA as the mobile phase. The oligomers were characterized by matrix assisted laser desorption/ionization-time of flight (MALDI-TOF) spectrometry with α -cyano-4-hydroxycinnamic acid (CHCA) as sample crystallization matrix (See **Appendix Table A2, Figure A2.2**). The extinction co-efficient of PNA T, C, Q, L, and R residues are 8.8, 7.3, 7.3, 7.3, and 0 mM⁻¹cm⁻¹, respectively, at 260 nm.^{10, 13}

3.2.4 Nondenaturing polyacrylamide gel electrophoresis

The HPLC-purified RNA and DNA oligonucleotides were acquired from Sigma-Aldrich, Singapore. The nondenaturing polyacrylamide gel electrophoresis (PAGE) (12%) experiments were performed using an incubation buffer containing 200 mM NaCl, 0.5 mM EDTA, 20 mM MES, pH 6.0, or 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5, or 200 mM NaCl, 0.5 mM EDTA, 2 mM MgCl₂, 20 mM HEPES, pH 7.5, or 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 8.0. For the Cy3-labelled rHP1-Cy3 and rHP3-Cy3, the samples were prepared using 50 nM RNA with varied PNA concentrations of 0, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 0.7, 1, 1.5, 2, 5, and 10 μ M in 20 μ L of buffer. For the unlabeled RNA/DNAs, the concentration was 0.25 μ M with varied PNA concentrations of 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 1, 1.6, 2, 4, 10, 16, and 20 μ M in 25 μ L of buffer. Snap cooling was done for RNA/DNA hairpins by directly dipping the sample tubes from 95 °C into ice bath followed by annealing with

PNAs by slow cooling from 65 °C to room temperature. The sample tubes were incubated overnight at 4 °C before the loading. To all the samples, 35% glycerol (20% of the total loaded volume) was added followed by vortexing for a few seconds before the loading into sample wells of the gels. Using a running buffer 1× TBE (Tris–Borate–EDTA), pH 8.3, the gels were run at 4 °C for 6 h with a voltage set at 250 V for all the experiments. The gels were imaged using a Typhoon imager directly for the Cy3-labelled (rHP1-Cy3 and rHP3-Cy3) or after ethidium bromide staining for 30 min for the unlabeled RNA/DNA hairpins.

3.2.5 Fluorescence binding study

The fluorescence emission spectra were recorded using a 1 cm square quartz cuvette on a Varian Cary Eclipse fluorescence spectrophotometer. The required 2-aminopurine-labelled RNA and DNA oligonucleotides purified by RP-HPLC were acquired from Sigma-Aldrich, Singapore. All the fluorescence experiments were carried out using 1 µM RNAs or DNAs and varied concentrations of PNAs at 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10, and 20 µM in 70 µL buffer having 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5. PNA alone samples at 10 µM were used as controls. The samples were prepared by slow cooling of RNAs/DNAs hairpins from 95 °C to room temperature, followed by the addition of PNA oligomers. The solutions were then vortexed for 10 min and left at room temperature for 1 hour and incubated at 4 °C overnight. All the emission spectra were recorded at room temperature over a wavelength range of 330–550 nm, at a fixed excitation wavelength of 303 nm with a slit width of 10 nm.

3.2.6 Thermal melting

The UV-absorbance-detected thermal melting experiments were conducted with the use of an 8-microcell cuvette using the Shimadzu UV-2550 UV-Vis spectrophotometer. The absorbance values at 260 nm with the temperature increasing from 15 to 95 °C followed by the temperature decreasing from 95 to 15 °C were recorded. The optical path length of the 8-microcell cuvette is 1 cm. The temperature ramp rate is 0.5 °C/min. The samples were prepared using 5 µM RNA with and without 5 µM PNA in 130 µL buffer having 200 mM NaCl, 0.5 mM EDTA, 20 mM NaH₂PO₄, pH 7.5. The samples containing the single-stranded RNA and PNA were annealed by slow cooling from 95 °C to room temperature, followed by overnight incubation at 4 °C. Data were normalized at high temperature and the melting temperatures were determined based on the Gaussian fit of the first derivative of the curves.

3.2.7 Cellular uptake and confocal microscopy study

Sf9 (*Spodoptera frugiperda*) insect cells were cultured onto a 4-well chamber slide system (Nunc Lab-Tek) at a cell density of 2×10^5 cells per compartment for 12 h. The cells were washed with PBS and replaced with serum free Sf-900 III medium (Life Technologies Inc., NY, USA) containing 2 µM of R-modified PNA P5-3R-CF or control PNA P1-CF for 12 h at 27 °C. Then, cells were fixed with 4% paraformaldehyde for 15 min at room temperature, rinsed with PBS. The fixed cells were counterstained with DAPI (Invitrogen, Gaithersburg, USA) for 5 min. Fluorescence signals were observed using an inverted confocal microscope (Leica SP8, Leica Microsystems, Mannheim, Germany).

Figure 3.3. RNA, DNA, and PNA molecules studied in this chapter. **(a-h)** Model RNA and DNA hairpins. The underlined U residues in rHP1 and rHP3 are replaced with Cy3 dye to have rHP1-Cy3 and rHP3-Cy3, respectively. **(i-j)** Model ssRNA strands that potentially form parallel and antiparallel duplexes with PNAs. **(k-p)** PNAs studied in this chapter. The letter “R” shown in red represents the PNA R residue. **(q,r)** Stable PNA-RNA₂ triplexes formed by PNA P1-1R4 and P5-3R targeting rHP1. **(s,t)** Unstable duplexes formed by rHP1-ss-p with P5-1R4 and P5-3R, respectively.

3.3 Results and Discussion

3.3.1 Recognition of an internal G-C pair in an RNA duplex (but not DNA duplex) by R-modified PNAs

Our modeling data suggest that an R-modified PNA can recognize an RNA duplex through major-groove PNA·dsRNA triplex formation similar to an L-modified PNA (**Figures 3.1-3.3**). Among the structures with one or more hydrogen bonds formed between R and G or between L and G, we observed two

main populations. It seems that L is more optimal for the simultaneous recognition of the N1 and carbonyl group of G on the Hoogsteen edge, probably because L has a relatively large stacking area and more optimized base-backbone linker. However, the R' residue, with an additional carbon in the base-backbone linker, shows increased flexibility and may result in the loss of hydrogen bonding and stacking interactions (**Figure 3.2g-i**). Thus, we focused on the R monomer (**Scheme 3.1**) in our experimental studies.

Figure 3.4. Nondenaturing PAGE characterization of triplex formation for model RNA hairpins. The incubation buffer is 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5. **(a-c and g)** PAGE characterization and determination of K_d for the binding of dbPNAs to rHP1 having a Watson–Crick G-C pair opposite to R. **(d-f)** PAGE characterization for the binding of dbPNA P1-1R4 to RNA hairpins with a C-G (rHP2), an A-U (rHP3-Cy3), or a U-A (rHP4) base pair opposite to R. rHP1 shows the strongest binding to P1-R4, suggesting a specific recognition of a G-C pair by an R residue. All the Cy3-labelled samples were prepared at 50 nM with 0, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 0.7, 1, 1.5, 2, 5, and 10 μ M of PNA in 20 μ L of buffer. rHP2 and rHP4 were at 0.25 μ M against 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 1, 1.6, 2, 4, 10, 16, and 20 μ M PNA in 25 μ L of buffer.

Figure 3.5. Nondenaturing PAGE characterization of PNAs designed for targeting rHP1-Cy3. **(a-c)** PAGE characterization for the formation of (rHP1-Cy3)·(P1-1R4) triplex. **(d, e, h, i)** PAGE characterization showing no binding of P1-1R4 with rHP2 (C-G) and rHP4 (U-A). **(f,g)** PAGE characterization showing weak binding of P1-1R4 with rHP3-Cy3 (A-U). **(j-l)** PAGE characterization for the formation of (rHP1-Cy3)·(P5-1R4) triplex. **(m-o)** PAGE characterization for the formation of (rHP1-Cy3)·(P5-3R) triplex. **(p-r)** PAGE characterization showing no binding of P5-3R with rHP2 (C-G) and rHP4 (U-A) and a weak binding with rHP3-Cy3 at pH 7.5. **(s-v)** PAGE characterization for the binding of P1 and P5 with rHP1-Cy3 and rHP3-Cy3 at pH 7.5. The samples of rHP1-Cy3

and rHP3-Cy3 were prepared at 50 nM against 0, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 0.7, 1, 1.5, 2, 5, and 10 μM of PNA in 20 μL buffer. Whereas, unlabeled RNAs were at 0.25 μM against PNAs at 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 1, 1.6, 2, 4, 10, 16, and 20 μM (in lanes from left to right) in 25 μL buffer. The buffer contains 200 mM NaCl, 0.5 mM EDTA, 20mM MES, pH 6.0 or 200 mM NaCl, 0.5 mM EDTA, 20mM HEPES pH 7.5 or 8.0. For the buffer containing Mg^{2+} , 2mM MgCl_2 was added at pH 7.5. The running buffer is $1\times$ TBE.

Figure 3.6. Nondenaturing PAGE characterization of P1-1R5 designed for targeting rHP6. (a-d) P1-1R5 binding to rHP6 but not rHP1, rHP5, or rHP7. (e-f) PAGE characterization for the formation of rHP6-P1-1R5 triplex at pH 6.0 and 8.0, respectively. The samples of rHP1-Cy3 were prepared at 50 nM against 0, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 0.7, 1, 1.5, 2, 5 and 10 μM of PNA in 20 μL buffer. Whereas, the unlabeled RNAs were at 0.25 μM against PNAs at 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 1, 1.6, 2, 4, 10, 16, and 20 μM (in lanes from left to right) in 25 μL buffer. The buffer contains 200 mM NaCl, 0.5 mM EDTA, 20mM MES, pH 6.0 or 200 mM NaCl, 0.5 mM EDTA, 20mM HEPES, pH 7.5 or 8.0. The running buffer is $1\times$ TBE.

Figure 3.7. Nondenaturing PAGE characterization of potential binding of PNAs to dHP1. **(a,b)** PAGE characterization showing no binding of P1-1R4 with dHP1. **(c,d)** PAGE characterization showing no binding of P5-1R4 with dHP1. **(e,f)** PAGE characterization showing no binding of P5-3R with dHP1 (**Table 3.1**). DNA hairpin dHP1 was at 0.25 μM against PNAs at 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 1, 1.6, 2, 4, 10, 16, and 20 μM (in lanes from left to right) in 25 μL buffer. The buffer contains 200 mM NaCl, 0.5 mM EDTA, 20mM MES, pH 6.0 or 200 mM NaCl, 0.5 mM EDTA, 20mM HEPES, pH 7.5 or 8.0 The running buffer is 1 $\times$ TBE.

Figure 3.8. K_d determination by nondenaturing PAGE for PNAs binding to RNAs. Unless otherwise specified, the incubation buffer contains 200 mM NaCl, pH 7.5. (a-c) K_d determination for the formation of (rHP1-Cy3)·(P1-IR4) triplex. (d-f) K_d determination for the formation of (rHP3-Cy3)·(P1-IR4) triplex. (g-i) K_d determination for the formation of (rHP1-Cy3)·(P5-IR4) triplex. (j-l) K_d determination for the formation of (rHP1-Cy3)·(P5-3R) triplex. (m) K_d determination for the formation of (rHP3-Cy3)·(P5-3R) triplex. (n-p) K_d determination for the formation of rHP6·(P1-R5) triplex. (q-s) K_d determination

for the formation of (rHP1-Cy3)·P1, (rHP1-Cy3)·P5, and (rHP3-Cy3)·P5 triplexes. **(t-u)** K_d determination for the formation of HIV-HP·(L4-RTTR) triplex. **(v-w)** K_d determination for the formation of (PH-v-Cy3)·(IR-1R) and (PH-v-Cy3)·(IR-3R) triplex. **(x)** K_d determination for the binding of antisense PNA IR-ASP with PH-v hairpin. The fraction of triplex/complex formation as calculated based on equation $Y = I_{\text{triplex}} / (I_{\text{triplex}} + aI_{\text{duplex}})$. The band intensities were normalized according to $a = I_{\text{triplex max}} / I_{\text{duplex max}}$. $I_{\text{duplex max}}$ is the band intensity for hairpin alone without the addition of PNA. $I_{\text{triplex max}}$ is the triplex band intensity with the highest band intensity. In addition, the normalized duplex intensity (see panels **d, e, f, m, p, and q**) was obtained by normalizing the duplex bands: $I_{\text{duplex}} / I_{\text{duplex max}}$. The data were fit to the equation: $Y = Y_0 + (B/2/R_0)(R_0 + X + K_d - ((R_0 + X + K_d)^2 - 4R_0X)^{1/2})$ where R_0 is the RNA hairpin concentration (50 nM or 0.25 μ M). Y_0 and B are the minimum and maximum fractions of triplex formation, respectively. X is the total PNA concentration and K_d is the dissociation constant.

The PNA oligomers were obtained by solid phase peptide synthesis (See **Appendix Table A2, Figure A2.2**). Nondenaturing PAGE experiments (**Figures 3.4-3.8**) were performed to characterize the binding properties of PNAs incorporated with R residue. Replacement of a single C residue at position 4 in PNA P1 with R (P1-1R4, **Figure 3.3k,l**) results in a significant enhancement in binding with the K_d value decreased from 4.9 μ M⁵⁸ to 0.2 μ M at 200 mM NaCl, pH 7.5 (**Figure 3.4, Table 3.1**). Interestingly, P1-R4 has the same binding affinity as that of PNA P5 (**Figure 3.3n**) with three L residues (**Figure 3.4, Table 3.1**). Notably, PNA P1-1R4 shows significantly weakened binding to rHP3 and

no observable binding to rHP2 and rHP4 (**Figure 3.4, Table 3.1**), suggesting that the PNA residue R is highly specific for the recognition of a G-C pair in rHP1 but not C-G, A-U, and U-A pairs in rHP2, rHP3, and rHP4, respectively. Furthermore, the fluorescence titration experiments of P1-1R4 into 2-aminopurine-labelled RNAs (**Figures 3.9-3.10**) reveal a K_d value ($0.4 \pm 0.2 \mu\text{M}$) in binding to dsRNA1-2AP, which has a similar binding site as rHP1. Importantly, no appreciable binding is observed for the duplexes with the base pair opposite to the R residue replaced with a C-G (dsRNA2-2AP), an A-U (dsRNA3-2AP), or a U-A (dsRNA4-2AP) (**Figures 3.9-3.10**). Clearly, incorporating guanidinium into a dbPNA facilitates specific recognition of a G-C pair in a dsRNA.

Table 3.1. K_d values (in μM) for PNAs binding to model RNA and DNA hairpins obtained by nondenaturing PAGE.^a

pH	rHP1-Cy3 (G-C)				rHP2 (C-G)			rHP3-Cy3 (A-U)			rHP4 (U-A)			dHP1 (G-C)	
	6.0	7.5	8.0		6.0	7.5	8.0	6.0	7.5	8.0	6.0	7.5	8.0	6.0	7.5
	No Mg^{2+}		With Mg^{2+}												
P1^b	(0.9±0.2)	(2.1±0.4)	(5.5±0.9)	(27.0±5.3)	-	(NB)	-	-	(NB)	-	-	(NB)	-	-	(NB)
	4.9±1.7								NB						
P5^b	(0.2±0.1)	(0.2±0.1)	(0.2±0.1)	(0.3±0.1)	-	(NB)	-	-	(0.4±0.1)	-	-	(1.9±0.4)	-	-	(>50)
	0.2±0.1								0.4±0.1						
P1-1R4	0.1±0.1	0.2±0.1	0.6±0.1	0.8±0.1	(NB)	(NB)	(NB)	>3	>10	>10	(NB)	(NB)	(NB)	(NB)	(NB)
P5-1R4	0.2±0.1	0.3±0.1	1.3±0.2	0.7±0.1	-	-	-	-	-	-	-	-	-	(NB)	(NB)
P5-3R	0.2±0.1	0.3±0.1	0.4±0.1	0.5±0.1	-	(NB)	-	-	>2	-	-	(NB)	-	(NB)	(NB)
P1-1R5	rHP1-Cy3 (A-U)				rHP5 (C-G)			rHP6 (G-C)			rHP7 (A-U)				
	-	NB	-	-	-	(NB)	-	(0.3±0.1)	(1.1±0.2)	(> 3)	-	(NB)	-	-	-

^aIncubation buffer used: 200 mM NaCl, 0.5 mM EDTA, 20 mM MES at pH 6.0 or 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES at pH 7.5 with or without 2 mM MgCl₂, or 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES at pH 8.0. The PAGE experiments were carried out at 4 °C. “-” indicates that the data were not obtained in this study. “NB” indicates no binding was observed at this condition.

^bThe PNA oligomer sequences were previously studied.^{11, 13, 58} The data shown in parentheses were for the gels done with unlabeled hairpins at 0.25 μM and with the gels post stained by EtBr. The errors shown are standard errors.

Both PAGE and 2-aminopurine fluorescence titration data reveal that PNAs P1-1R4 and P5-1R4 show no binding to dHP1 (**Table 3.1, Figures 3.7, 3.9 and 3.10e**). The dsRNA over dsDNA selectivity of the dbPNAs are consistent with previously reported results.^{11-13, 15, 18, 58} It is likely that a dsDNA has a wider and shallower major groove and a larger twist compared to a dsRNA, resulting in a preferential binding of a PNA to dsRNA over dsDNA.

Figure 3.9. Fluorescence titration of R-modified P1-1R4 into 2-aminopurine (2AP) labelled dsRNAs, a dsDNA, and an ssRNA. (a) Triplex structure of (dsRNA1-2AP)·(P1-1R4). 2AP is shown as “2” in the triplex. (b-c) Fluorescence

titration data. P1-1R4 binding to RNA duplex dsRNA1-2AP results in fluorescence quenching. The G-C pair highlighted in the box of dsRNA1-2AP (panel a) is replaced with a C-G, an A-U, and a U-A pair, respectively, in dsRNA2-2AP, dsRNA3-2AP, dsRNA4-2AP. dsDNA1-2AP is a DNA homolog of dsRNA1-2AP. Titration of P1-1R4 into the dsRNAs with (dsRNA2-2AP, dsRNA3-2AP, and dsRNA4-2AP) and the dsDNA1-2AP results in no significant change in fluorescence intensity. ssRNA1-2AP fluorescence intensity is quenched upon duplex formation with P1-1R4 and indicates a 14-fold weaker binding than dsRNA1-2AP. All the samples were prepared with 1 μ M RNA/DNA in the presence of 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10, and 20 μ M of PNA P1-1R4. For binding to ssRNA1-2AP, two more high-concentration points (35 and 50 μ M) were added. The buffer contains 70 μ L of 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5. The excitation wavelength is 303 nm. The measurements were done at room temperature.

Figure 3.10. Structure-specific binding of P1-1R4 for dsRNA over dsDNA and ssRNA using 2-aminopurine labelled RNAs and DNA at 200 mM NaCl, pH 7.5.

(a) 2-AP labelled RNA and DNA structures used to study the structure specificity of P1-1R4. (b-e) The 2-aminopurine fluorescence in dsRNA1-2AP is quenched by P1-1R4 (see **Figure 3.9**) but not by the other RNA and DNA duplexes tested. (f) The 2-aminopurine fluorescence in ssRNA1-2AP is quenched by P1-1R4 revealing a weakened K_d of $5.6 \pm 1.5 \mu\text{M}$. All the samples were prepared with 1 μM 2-AP labelled RNA/DNA in 70 μL of buffer of 200 mM NaCl, 0.5 mM EDTA, 20mM HEPES, pH 7.5. P1-R4 alone sample was at 10 μM .

Next, we incorporated the R residue at position 5 (P1-1R5) to further characterize the binding affinity and specificity (**Figure 3.3m**). PNA P1-1R5 contains three cytosine residues and may be expected to have a weakened binding with a relatively strong pH dependence due to the requirement of the formation of three $\text{C}^+\cdot\text{G}\cdot\text{C}$ base triples (**Figures 3.1b and 3.3m**). Indeed, the PAGE data show that PNA P1-R5 binds to the target rHP6 (**Figure 3.3f**) with a K_d of $1.1 \pm 0.2 \mu\text{M}$ at 200 mM NaCl, pH 7.5 (**Figures 3.6c and 3.8o**). Again, P1-1R5 shows no observable binding to the hairpins with the corresponding G-C pair in rHP6 replaced with a C-G (rHP5), an A-U (rHP1-Cy3), or a U-A (rHP7) (**Figure 3.6**). Taken together, our data show that dbPNAs having a single R residue bind to dsRNAs highly sequence specifically.

3.3.2 Effects of sequence context, pH, and salt on PNA·RNA₂ triplex formation

PNA P1 contains three C residues and shows a significantly weakened binding ($> 20 \mu\text{M}$, **Table 3.1, Figure 3.4-3.5**) at 200 mM NaCl, pH 8.0. Surprisingly, both P1-1R4 (containing two C residues) and P1-1R5 (containing three C

residues) show a reduced pH dependence compared to P1 in binding to target dsRNAs (**Table 1, Figures 3.4-3.6**). Thus, incorporating R immediately next to C as present in P1-1R5 may be compatible for triplex formation. It is likely that C⁺ and R⁺ (guanidinium) have coupled stacking and hydrogen bonding interactions within a triplex structure, compensating for the expected charge-charge repulsion. The delocalized positive charges in R⁺ and C⁺ may also reduce charge-charge repulsion.

Interestingly, P5-1R4 and P5-3R show similar binding strength compared to P1-1R4 and P5 at various pH's (**Table 3.1, Figures 3.4-3.5 and 3.8**). It is possible that L and R residues at or near the terminal ends of a PNA show a relatively reduced stabilization in binding to a dsRNA.¹¹ Addition of 2 mM MgCl₂ moderately weakens the binding of P5-1R4 and P5-3R, suggesting that R residues stabilize the binding mainly through hydrogen bonding and stacking interactions (**Figures 3.1f,g,k,l and 3.2f**), but not nonspecific electrostatic interactions.

3.3.3 Incorporation of PNA R residues stabilizes and destabilizes PNA·RNA₂ triplex and PNA-RNA duplex, respectively

We further tested the binding of an R-modified PNA to a ssRNA. Our 2-aminopurine fluorescence titration data indicate that P1-1R4 shows binding to ssRNA1-2AP with a K_d of 5.6 ± 1.5 μ M, which is significantly weaker than that for binding to dsRNA1-2AP (0.4 ± 0.2 μ M) (**Figures 3.9 and 3.10a,f**). The dsRNA over ssRNA binding selectivity revealed by fluorescence titration is further corroborated with thermal melting experiments. The UV-absorbance

detected thermal melting data show that P1-1R4 can form parallel and antiparallel duplexes with complementary ssRNAs (rHP1-ss-p and rHP1-ss-ap) with melting temperatures of 39.9 and 34.8 °C, respectively, which are comparatively lower than the duplexes formed by PNA P1 (44.2 and 48.5 °C) (**Figures 3.3 and 3.14**).¹³ Importantly, P5-1R4 shows no melting transitions for both rHP1-ss-p and rHP1-ss-ap as compared to the melting temperature for P5 with rHP1-ss-p (31.5 °C) and rHP1-ss-ap (34.4 °C) (**Figures 3.3 and Figure 3.14**).¹³ The data suggest that, compared to an L-modified PNA, an R-modified PNA is more unfavorable in binding to ssRNAs. Presumably, destabilization of PNA-RNA duplex formation results from an unfavorable Watson–Crick-like R-G pair (**Figure 3.1h**), similar to an unstable Watson–Crick-like L-G pair (**Figure 3.2c**).^{11, 69} Taken together, our data clearly show that incorporating R residues into PNAs facilitates structure specific binding to dsRNAs over ssRNAs.

Figure 3.11. Nondenaturing PAGE characterization of PNAs binding to HIV-1 frameshift-inducing RNA (HIV-HP). The incubation buffer is 0.5 mM EDTA, 20 mM HEPES, pH 7.5, with 200 mM NaCl or 10 mM NaCl. **(a-f)** Sequences and structures of HIV-HP, PNAs, and HIV-HP-L4-RTTR triplex. **(g-j)** PAGE

characterization for the binding of PNAs L4-RR, L4-RTTR, and L4-CTTC to HIV-HP. All the samples were prepared with 0.25 μ M RNA (25 μ L) with varied concentration of PNA (0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 1, 1.6, 2, 4, 10, 16, and 20 μ M, in lanes from left to right). Reducing the NaCl concentration from 200 mM to 10 mM results in a slight enhancement in binding for L4-RTTR (panels **h** and **i**), suggesting that the two R residues recognize G-C pairs with specific hydrogen bonding and stacking but not nonspecific ionic interactions.

Figure 3.12. Nondenaturing PAGE characterization of triplex formation for HIV-1 frameshift-inducing RNA (HIV-HP) with dbPNAs L4-RR, and L4-CTTC. The incubation buffer is 10 mM NaCl, 0.5 mM EDTA, 20mM HEPES, pH 7.5. Both samples in panels **a** and **b** were prepared with 0.25 μ M RNA (25 μ L) with varied concentration of PNA (0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 1, 1.6, 2, 4, 10, 16, and 20 μ M, in lanes from left to right).

3.3.4 Binding to HIV-1 minus-one ribosomal frameshift inducing RNA hairpin and influenza viral panhandle structure by dbPNAs containing multiple R residues

We tested the application of R- and L-modified dbPNAs for targeting HIV-1 frameshift site RNA hairpin⁷⁰⁻⁷³ and influenza viral panhandle structure (PH-v).^{55, 74-79} We made a series of R-containing PNAs targeting HIV-1 frameshift stimulatory RNA hairpin (HIV-HP, **Figure 3.11**). Our previous work shows that

a 6-mer PNA L4 binds to HIV-HP with a K_d of $0.3 \pm 0.1 \mu\text{M}$ (at 200 mM NaCl, pH 7.5).⁵⁸ Incorporating two non-consecutive R residues (L4-RTTR) into L4 does not affect the binding significantly ($K_d = 0.6 \pm 0.1 \mu\text{M}$). Reducing the NaCl concentration from 200 mM to 10 mM results in a K_d of $0.4 \mu\text{M}$ for L4-RTTR (**Figure 3.11**), further suggesting that an R-modified PNA binds to a dsRNA through hydrogen bonding and stacking but not nonspecific ionic interactions. However, replacing the two R residues with C residues (L4-CTTC) results in no binding at both 200 mM and 10 mM NaCl (**Figures 3.11-3.12**), suggesting the stabilizing effect of R residues. Interestingly, incorporating two consecutive R residues into L4 (L4-RR) causes no binding to HIV-HP at both 200 mM and 10 mM NaCl (**Figures 3.11-3.12**), which suggests that consecutive R residues are not favorable for PNA·dsRNA triplex formation, presumably due to the unfavorable stacking of consecutive R residues. We speculate that the stacking pattern restricted by the helical structure of a PNA·dsRNA triplex does not allow the favorable self-stacking geometry of two guanidinium as observed in other systems.²³ Thus, R residues may be incorporated into PNAs immediately adjacent to T, C, and L but not another R for enhanced recognition of G-C pairs in dsRNAs.

Influenza A viruses contain eight viral RNA (vRNA) segments with highly conserved panhandle duplex structure formed between the first 13 nucleotides at the 5' end and the last 12 nucleotides at the 3' end.^{55, 74-79} A 10-mer dbPNA IR-1 binds to PH-v with a K_d of $0.3 \mu\text{M}$ at 200 mM NaCl, pH 7.5.^{55, 69} We replaced one and three L residues in IR-1 with R to obtain IR-1R and IR-3R, respectively (**Figure 3.13**). The PAGE data show that IR-1R and IR-3R bind to a Cy3-labelled

and **i**) A stable PH-v-IR-3R triplex and an unstable (PH-v-ss)-(IR-3R) duplex. PH-v but not the ssRNA segment of PH-v (PH-v-ss) is targeted by dbPNAs IR-1R and IR-3R. **(h)** A duplex invasion complex formed between PH-v and IR-ASP. **(j-l)** nondenaturing PAGE characterization. The dbPNAs IR-1R and IR-3R with L residues replaced with R at one or more positions show a similar binding affinity to PH-v-Cy3.^{55,69} The 10-mer dbPNAs IR-1R and IR-3R show a stronger binding affinity than the unmodified 11-mer asPNA IR-ASP. The Cy3-labelled RNA is at 50 nM (20 μ L) with varied concentration of PNA (0, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 0.7, 1, 1.5, 2, 5, and 10 μ M of PNA in lanes from left to right).

Figure 3.14. Thermal melting results for potential PNA-RNA duplexes. The heating curves are shown. The melting temperatures are shown for the curves with melting transitions. The ssRNAs and PNAs were at 5 μ M (130 μ L). The buffer is 200 mM NaCl, 0.5 mM EDTA, 20 mM NaH₂PO₄, pH 7.5. **(a)** Thermal melting results for control PNAs P1 and P5 with rHP1-ss-p and rHP1-ss-ap. Compared to unmodified P1, P5 contains L residues and shows reduced binding to ssRNAs, consistent with our previous work.¹¹ **(b)** Thermal melting results for R-modified PNAs and an asPNA IR-ASP. Incorporation of PNA R residues causes destabilization of PNA-RNA duplexes, due to unfavorable Watson-Crick-

like base pairing, which is similar to PNA L and Q residues (see **Figures 3.1h and 3.2c,d**).

3.3.5 Cellular uptake of an R-modified PNA

It is known that incorporation of the guanidinium moiety facilitates enhanced cellular uptake of various cargos including PNAs.^{13, 23, 38-49} We attached carboxyfluorescein to PNAs P1 (H₂N-Lys-TCTCTTTC-CONH₂) and P5-3R (H₂N-Lys-TRTRTTTR-CONH₂) to obtain P1-CF and P5-3R-CF, respectively (**Appendix Table A2**), to assess how incorporation of R residues may affect cellular permeability. At first, we tested the cellular uptake properties of carboxyfluorescein-labelled P5-3R-CF and P1-CF on HeLa and Madin-Darby Canine Kidney (MDCK) cell lines. Unfortunately, we observed poor fluorescence signals in both of the cell lines. Interestingly, incubation of Sf9 cells with carboxyfluorescein-labelled P5-3R-CF and P1-CF for 12 hours shows fluorescence signal for P5-3R-CF mainly in cytoplasm (**Figure 3.15a,b,e,f**), although it is known that the fixation of the cells may perturb the cellular localization of the PNA. However, no significant fluorescence signal was observed for the cells treated with P1-CF (**Figure 3.15c,d,g,h**), consistent with the fact that unmodified PNAs are not taken up by cells. P5-3R-CF is a derivative of P1-CF with the C residues replaced with R. Thus, incorporating R residues improves cell permeability.

Figure 3.15. Cell uptake study of an R-modified PNA. Confocal microscopy images shown are for Sf9 cells treated with 2 μ M (**a, b, e, f**) R-modified PNA P5-3R-CF (CF-H₂N-Lys-TCTCTTTC-CONH₂) and (**c, d, g, h**) control PNA P1-CF (CF-H₂N-Lys-TRTRTTTR-CONH₂). CF: caryboxyfluorescein. (**a, c**) Carboxyfluorescein imaging. Ex, 492nm; Em, 517nm. (**b, d**) Bright field imaging. (**e, g**) DAPI imaging. Ex, 358nm; Em, 461nm. (**f, h**) Merged images. Scale bar: 20 μ m.

Our previous work shows that incorporating multiple Q but not L residues may not improve cellular permeability,^{13, 55} although both Q and L have a guanidine moiety (**Figure 3.1**). It is possible that, compared to PNA residue L, which has the guanidine moiety within an aromatic base, Q and R have a positive charge and more optimal flexibility.^{13, 39, 41, 47, 49} Both positive charges and optimal flexibility are critical for facilitating tight binding with the cell surface counter anions (such as phosphate and sulfate groups present on cell membrane) for initiating cellular uptake.^{43, 44, 49} Taken together, our results suggest that one may

incorporate multiple R and/or Q residues into dbPNAs for an improved cellular permeability.

3.4 Conclusions

In summary, we have developed a simple guanidinium-based PNA residue R for the specific recognition of Watson-Crick G-C base pairs in RNA duplexes at physiological pH conditions. The synthetic scheme for the PNA R monomer involves a few facile and simple steps, which is advantageous compared to our previously developed PNA L monomer.¹¹ Our study exemplifies the potential broad scope of the newly developed R-modified dbPNAs in the sequence and structure specific targeting of dsRNAs over ssRNAs and dsDNAs. To the best of our knowledge, it is the first report on incorporating a simple guanidinium into a PNA scaffold as a pairing and stacking partner for the recognition of RNA G-C pairs. In addition, our dbPNAs are uniquely positioned on the major groove for molecular recognition. Our work shows that incorporating nonconsecutive R residues into our dbPNA platform containing other modified L and Q residues facilitates the pairing of R with G on the Hoogsteen edge (**Figure 3.1f,g,k,l**) and avoids other potential pairing geometries with any other natural bases including C on the Watson–Crick face with two hydrogen bond acceptors.^{27, 28, 80} The presence of guanidinium group in single-stranded PNAs also serves to improve the cellular uptake (**Figure 3.15**), with potentially improved solubility and bioactivity. It is possible that one may combine multiple R and/or Q¹³ residues together with other cell penetrating moiety such as neamine⁵⁵ for further improved cellular uptake and activities of dbPNAs.

References

1. Cech, T. R.; Steitz, J. A. *Cell* **2014**, 157, 77-94.
2. Disney, M. D.; Yildirim, I.; Childs-Disney, J. L. *Org. Biomol. Chem.* **2014**, 12, 1029-1039.
3. Kole, R.; Krainer, A. R.; Altman, S. *Nat. rev. Drug discovery* **2012**, 11, 125-140.
4. Kaczmarek, J. C.; Kowalski, P. S.; Anderson, D. G. *Genome Med.* **2017**, 9, 60.
5. Han, H.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, 90, 3806-3810.
6. Roberts, R. W.; Crothers, D. M. *Science* **1992**, 258, 1463-1466.
7. Escude, C.; Francois, J. C.; Sun, J. S.; Ott, G.; Sprinzl, M.; Garestier, T.; Helene, C. *Nucleic Acids Res.* **1993**, 21, 5547-5553.
8. Zhou, Y.; Kierzek, E.; Loo, Z. P.; Antonio, M.; Yau, Y. H.; Chuah, Y. W.; Geifman-Shochat, S.; Kierzek, R.; Chen, G. *Nucleic Acids Res.* **2013**, 41, 6664-6673.
9. Patil, K. M.; Chen, G. *Modified Nucleic Acids in Biology and Medicine*, Erdmann, V. A.; Jurga, S.; Barciszewski, J., Eds. Springer: 2016; pp 299-317.
10. Toh, D.-F. K.; Patil, K. M.; Chen, G. *J Visualized Exp.* **2017**, 127, e56221.
11. Devi, G.; Yuan, Z.; Lu, Y.; Zhao, Y.; Chen, G. *Nucleic Acids Res.* **2014**, 42, 4008-4018.
12. Devi, G.; Zhou, Y.; Zhong, Z.; Toh, D.-F. K.; Chen, G. *Wiley Interdiscip Rev. RNA* **2015**, 6, 111-128.
13. Toh, D.-F. K.; Devi, G.; Patil, K. M.; Qu, Q.; Maraswami, M.; Xiao, Y.; Loh, T.-P.; Zhao, Y.; Chen, G. *Nucleic Acids Res.* **2016**, 44, 9071-9082.
14. Puah, R. Y.; Jia, H.; Maraswami, M.; Toh, D.-F. K.; Ero, R.; Yang, L.; Patil, K. M.; Ong, A. A. L.; Krishna, M. S.; Sun, R.; Tong, C.; Huang, M.; Chen, X.; Loh, T. P.; Gao, Y. G.; Liu, D. X.; Chen, G. *Biochemistry* **2018**, 57, 149-159.
15. Zengeya, T.; Gupta, P.; Rozners, E. *Angew. Chem. Int. Ed.* **2012**, 51, 12593-12596.
16. Hnedzko, D.; McGee, D. W.; Karamitas, Y. A.; Rozners, E. *RNA* **2017**, 23, 58-69.
17. Muse, O.; Zengeya, T.; Mwaura, J.; Hnedzko, D.; McGee, D. W.; Grewer, C. T.; Rozners, E. *ACS Chem. Biol.* **2013**, 8, 1683-1686.
18. Gupta, P.; Zengeya, T.; Rozners, E. *Chem. Commun.* **2011**, 47, 11125-11127.
19. Wojciechowski, F.; Hudson, R. H. *Curr. Top. Med. Chem.* **2007**, 7, 667-679.
20. Li, M.; Zengeya, T.; Rozners, E. *J. Am. Chem. Soc.* **2010**, 132, 8676-8681.
21. Kotikam, V.; Kennedy, S. D.; MacKay, J. A.; Rozners, E. *Chem. Eur. J.* **2019**, 25, 4367-4372.
22. Ono, A.; Tso, P. O. P.; Kan, L. S. *J. Am. Chem. Soc.* **1991**, 113, 4032-4033.
23. Vazdar, M.; Heyda, J.; Mason, P. E.; Tesei, G.; Allolio, C.; Lund, M.; Jungwirth, P. *Acc. Chem. Res.* **2018**, 51, 1455-1464.
24. Gund, P. *J. Chem. Edu.* **1972**, 49, (2), 100-103.
25. Puglisi, J. D.; Tan, R.; Calnan, B. J.; Frankel, A. D.; Williamson, J. R. *Science* **1992**, 257, 76-80.
26. Seeman, N. C.; Rosenberg, J. M.; Rich, A. *Proc. Natl. Acad. Sci. U. S. A.* **1976**, 73, 804-808.
27. Cheng, A. C.; Chen, W. W.; Fuhrmann, C. N.; Frankel, A. D. *J. Mol. Biol.* **2003**, 327, 781-796.
28. Kondo, J.; Westhof, E. *Nucleic Acids Res.* **2011**, 39, 8628-8637.
29. Kim, J. H.; Resende, R.; Wennekes, T.; Chen, H. M.; Bance, N.; Buchini, S.; Watts, A. G.; Pilling, P.; Streltsov, V. A.; Petric, M.; Liggins, R.; Barrett, S.; McKimm-Breschkin, J. L.; Niikura, M.; Withers, S. G. *Science* **2013**, 340, 71-75.
30. Bottcher, T.; Kolodkin-Gal, I.; Kolter, R.; Losick, R.; Clardy, J. *J. Am. Chem. Soc.* **2013**, 135, 2927-2930.

31. McKeever, C.; Kaiser, M.; Rozas, I. *J. Med. Chem.* **2013**, 56, 700-711.
32. Takayama, K.; Mori, K.; Taketa, K.; Taguchi, A.; Yakushiji, F.; Minamino, N.; Miyazato, M.; Kangawa, K.; Hayashi, Y. *J. Med. Chem.* **2014**, 57, 6583-6593.
33. Wu, H.; Esteve, E.; Tremaroli, V.; Khan, M. T.; Caesar, R.; Manneras-Holm, L.; Stahlman, M.; Olsson, L. M.; Serino, M.; Planas-Felix, M.; Xifra, G.; Mercader, J. M.; Torrents, D.; Burcelin, R.; Ricart, W.; Perkins, R.; Fernandez-Real, J. M.; Backhed, F. *Nat. Med.* **2017**, 23, 850-858.
34. Heinzl, G. A.; Huang, W.; Yu, W.; Giardina, B. J.; Zhou, Y.; MacKerell, A. D., Jr.; Wilks, A.; Xue, F. *J. Med. Chem.* **2016**, 59, 6929-6942.
35. Saczewski, F.; Balewski, L. *Expert Opin. Ther. Pat.* **2009**, 19, 1417-1448.
36. Holmes, T. C.; May, A. E.; Zaleta-Rivera, K.; Ruby, J. G.; Skewes-Cox, P.; Fischbach, M. A.; DeRisi, J. L.; Iwatsuki, M.; Ōmura, S.; Khosla, C. *J. Am. Chem. Soc.* **2012**, 134, 17797-17806.
37. von Itzstein, M.; Wu, W. Y.; Kok, G. B.; Pegg, M. S.; Dyason, J. C.; Jin, B.; Phan, T. V.; Smythe, M. L.; White, H. F.; Oliver, S. W.; Colman, P. M.; Varghese, J. N.; Michael Ryan, D.; Woods, J. M.; Bethell, R. C.; Hotham, V. J.; Cameron, J. M.; Penn, C. R. *Nature* **1993**, 363, 418-423.
38. Gupta, P.; Muse, O.; Rozners, E. *Biochemistry* **2012**, 51, 63-73.
39. Thomas, S. M.; Sahu, B.; Rapireddy, S.; Bahal, R.; Wheeler, S. E.; Procopio, E. M.; Kim, J.; Joyce, S. C.; Contrucci, S.; Wang, Y.; Chiosea, S. I.; Lathrop, K. L.; Watkins, S.; Grandis, J. R.; Armitage, B. A.; Ly, D. H. *ACS Chem. Biol.* **2013**, 8, 345-352.
40. Shiraishi, T.; Nielsen, P. E. *Methods Mol. Biol.* **2011**, 751, 209-221.
41. Lu, X. W.; Zeng, Y.; Liu, C. F. *Org Lett* **2009**, 11, 2329-2332.
42. Ohmichi, T.; Kuwahara, M.; Sasaki, N.; Hasegawa, M.; Nishikata, T.; Sawai, H.; Sugimoto, N. *Angew. Chem. Int. Ed.* **2005**, 44, 6682-6685.
43. Patil, K. M.; Naik, R. J.; Rajpal; Fernandes, M.; Ganguli, M.; Kumar, V. A. *J. Am. Chem. Soc.* **2012**, 134, 7196-7199.
44. Lattig-Tunnemann, G.; Prinz, M.; Hoffmann, D.; Behlke, J.; Palm-Apergi, C.; Morano, I.; Herce, H. D.; Cardoso, M. C. *Nat. Commun.* **2011**, 2, 453.
45. Wu, C. H.; Weng, M. H.; Chang, H. C.; Li, J. H.; Cheng, R. P. *Bioorg. Med. Chem.* **2014**, 22, 3016-3020.
46. Du, S.; Liew, S. S.; Li, L.; Yao, S. Q. *J. Am. Chem. Soc.* **2018**, 140, (47), 15986-15996.
47. Zhou, P.; Wang, M.; Du, L.; Fisher, G. W.; Waggoner, A.; Ly, D. H. *J. Am. Chem. Soc.* **2003**, 125, 6878-6879.
48. Gait, M. J.; Arzumanov, A. A.; McClorey, G.; Godfrey, C.; Betts, C.; Hammond, S.; Wood, M. J. A. *Nucleic Acid Ther.* **2019**, 29, 1-12.
49. Wexselblatt, E.; Esko, J. D.; Tor, Y. *J. Org. Chem.* **2014**, 79, 6766-6774.
50. Huang, L.; Wang, J.; Wilson, T. J.; Lilley, D. M. J. *Cell Chem. Biol.* **2017**, 24, 1407-1415 e2.
51. Sherlock, M. E.; Malkowski, S. N.; Breaker, R. R. *Biochemistry* **2017**, 56, 352-358.
52. Wilds, C. J.; Maier, M. A.; Manoharan, M.; Egli, M. *Helv. Chim. Acta* **2003**, 86, 966-978.
53. Fox, K. R.; Brown, T. *Biochem. Soc. trans.* **2011**, 39, 629-634.
54. Ohkubo, A.; Yamada, K.; Ito, Y.; Yoshimura, K.; Miyauchi, K.; Kanamori, T.; Masaki, Y.; Seio, K.; Yuasa, H.; Sekine, M. *Nucleic Acids Res.* **2015**, 43, 5675-5686.

55. Kesy, J.; Patil, K. M.; Kumar, S. M.; Shu, Z.; Yee, Y. H.; Zimmermann, L.; Ong, A. A. L.; Toh, D.-F. K.; Krishna, M. S.; Yang, L.; Decout, J. L.; Luo, D.; Prabakaran, M.; Chen, G.; Kierzek, E. *Bioconjugate Chem.* **2019**, *30*, 931-943.
56. Semenyuk, A.; Darian, E.; Liu, J.; Majumdar, A.; Cuenoud, B.; Miller, P. S.; Mackerell, A. D., Jr.; Seidman, M. M. *Biochemistry* **2010**, *49*, 7867-7878.
57. Ong, A. A. L.; Toh, D. F. K.; Patil, K. M.; Meng, Z.; Yuan, Z.; Krishna, M. S.; Devi, G.; Haruehanroengra, P.; Lu, Y.; Xia, K.; Okamura, K.; Sheng, J.; Chen, G. *Biochemistry* **2019**, *58*, 1319-1331.
58. Patil, K. M.; Toh, D.-F. K.; Yuan, Z.; Meng, Z.; Shu, Z.; Zhang, H.; Ong, A. A. L.; Krishna, M. S.; Lu, L.; Lu, Y.; Chen, G. *Nucleic Acids Res.* **2018**, *46*, 7506-7521.
59. Zgarbova, M.; Otyepka, M.; Sponer, J.; Mladek, A.; Banas, P.; Cheatham, T. E., III; Jurecka, P. *J. Chem. Theory Comput.* **2011**, *7*, 2886-2902.
60. Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E., III; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817-3829.
61. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157-1174.
62. Jakalian, A.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2002**, *23*, 1623-1641.
63. Mark, P.; Nilsson, L. *J. Phys. Chem. A* **2001**, *105*, 9954-9960.
64. Kresse, G.; Hafner, J. *Phys. Rev. B Condens. Matter Mater. Phys.* **1994**, *49*, 14251-14269.
65. Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L. P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. *PLoS Comput. Biol.* **2017**, *13*, e1005659.
66. McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernandez, C. X.; Schwantes, C. R.; Wang, L. P.; Lane, T. J.; Pande, V. S. *Biophys. J.* **2015**, *109*, 1528-1532.
67. Thomson, S. A.; Josey, J. A.; Cadilla, R.; Gaul, M. D.; Fred Hassman, C.; Luzzio, M. J.; Pipe, A. J.; Reed, K. L.; Ricca, D. J.; Wiethe, R. W.; Noble, S. A. *Tetrahedron* **1995**, *51*, 6179-6194.
68. Bernatowicz, M. S.; Wu, Y.; Matsueda, G. R. *Tetrahedron Lett.* **1993**, *34*, 3389-3392.
69. Krishna, M. S.; Toh, D. F. K.; Meng, Z.; Ong, A. A. L.; Wang, Z.; Lu, Y.; Xia, K.; Prabakaran, M.; Chen, G. *Anal. Chem.* **2019**, *91*, 5331-5338.
70. Baril, M.; Dulude, D.; Gendron, K.; Lemay, G.; Brakier-Gingras, L. *RNA* **2003**, *9*, 1246-1253.
71. Garcia-Miranda, P.; Becker, J. T.; Benner, B. E.; Blume, A.; Sherer, N. M.; Butcher, S. E. *J. Virol.* **2016**, *90*, 6906-6917.
72. Watts, J. M.; Dang, K. K.; Gorelick, R. J.; Leonard, C. W.; Bess, J. W., Jr.; Swanstrom, R.; Burch, C. L.; Weeks, K. M. *Nature* **2009**, *460*, 711-716.
73. Gareiss, P. C.; Miller, B. L. *Curr. Opin. Investig. Drugs* **2009**, *10*, 21-28.
74. Lenartowicz, E.; Kesy, J.; Ruszkowska, A.; Soszynska-Jozwiak, M.; Michalak, P.; Moss, W. N.; Turner, D. H.; Kierzek, R.; Kierzek, E. *PLoS One* **2016**, *11*, e0148281.
75. Hsu, M. T.; Parvin, J. D.; Gupta, S.; Krystal, M.; Palese, P. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 8140-8144.
76. Bae, S. H.; Cheong, H. K.; Lee, J. H.; Cheong, C.; Kainosho, M.; Choi, B. S. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 10602-10607.
77. Gultyaev, A. P.; Fouchier, R. A. M.; Olsthoorn, R. C. L. *Int. Rev. Immunol.* **2010**, *29*, 533-556.
78. Medina, R. A.; Rojas, M.; Tuin, A.; Huff, S.; Ferres, M.; Martinez-Valdebenito, C.; Godoy, P.; Garcia-Sastre, A.; Fofanov, Y.; SantaLucia, J., Jr. *J. Clin. Microbiol.* **2011**, *49*, 335-344.

79. Te Velthuis, A. J. W.; Long, J. C.; Bauer, D. L. V.; Fan, R. L. Y.; Yen, H. L.; Sharps, J.; Siegers, J. Y.; Killip, M. J.; French, H.; Oliva-Martin, M. J.; Randall, R. E.; de Wit, E.; van Riel, D.; Poon, L. L. M.; Fodor, E. *Nat. Microbiol.* **2018**, 3, 1234-1242.
80. Chen, Y.; Yang, F.; Zubovic, L.; Pavelitz, T.; Yang, W.; Godin, K.; Walker, M.; Zheng, S.; Macchi, P.; Varani, G. *Nat. Chem. Biol.* **2016**, 12, 717-723.

Chapter 4

Conclusion and Future Work

RNA has diverse structures and is a vital molecule found in all cells.¹ The diverse secondary and tertiary structures determine the various biological functions of RNA.² To understand and target RNAs, scientists have developed many modified nucleic acid mimics.³ The focus has been mainly on regulating gene expression and establishing structural insights into the functions of DNA and RNA.^{4, 5} Although, the last few decades witnessed an enormous research interest in modified nucleic acid chemistry, there have been only a single-digit number of therapeutically useful FDA-approved nucleic acid oligomers.^{6, 7} There are several problems posed in the development of therapeutic oligomers from the design and development in a research lab till the clinical trials. The key issue is the delivery to the target site, as many of the oligomers with traditional sugar phosphate backbone are readily cleaved by the action of nucleases and proteases. This feature of naturally occurring nucleic acids and analogues not only limits their application in therapy but also in diagnosis. Phosphorodiamidate morpholino oligomers (PMOs) and peptide nucleic acids (PNAs), both with neutral backbones, can withstand the enzymatic degradations and avoid non-specific charge-charge interactions with proteins.^{8,9}

Currently, a PMO has been successfully developed as an antisense drug in the treatment of duchenne muscular dystrophy,⁷ a genetic disorder involving

progressive muscle degeneration and weakness. Apart from sharing a poor cellular delivery for both PMOs and PNAs, a major setback posed by PMOs is sufficiently long oligomer length required to observe any promising outcome.¹⁰ Furthermore, PMOs act through the antisense mechanism involved in Watson–Crick duplex formation and typically limit the applications to the single-stranded, unstructured or partially structured regions of RNA. The structured regions in the RNA targets may significantly weaken the binding strength and slow down the binding kinetics of oligonucleotides.¹¹ Therefore, triplex formation provide an alternative strategy which involves targeting not only the sequences but also the secondary structures without disturbing the original target structure . Therefore, triplex formation presumably facilitates fast binding kinetics with the target regions and achieves strong binding using relatively short (6-10 mer) oligomer sequence.⁹

In this thesis, we have established a method for the direct and selective detection of duplex RNA in a programmable fashion over single-stranded RNA and duplex DNA through the fluorescence enhancement upon triplex formation under physiological pH conditions.¹² The triplex fluorescence enhancement is less explored and relatively challenging to achieve due to its weak Hoogsteen hydrogen bonding compared to Watson–Crick pairing under physiological pH conditions. The designed second-generation ^{bt}U monomer displays an increased fluorescence upon triplex formation with the complementary duplex RNA mainly due to the unique structural environments experienced by ^{bt}U within a PNA•dsRNA triplex. In contrast to the unmodified PNA C monomer, which requires lower pH condition for the duplex binding, our previously reported L

monomer facilitate strong binding with Watson–Crick G-C base pair at pH 7.5.¹³ Further, L monomer is selective for duplex over single-stranded RNA by stabilizing Hoogsteen L·G-C Paring and destabilizing Watson–Crick L-G pairing. A similar structural selectivity was observed for Q monomer, which binds specifically to inverted Watson–Crick C-G base pair in an RNA duplex. Combining L and Q residues with ^{bt}U facilitates the construction of a fluorescent turn-on probe for the direct and selective detection of duplex RNA over single-stranded RNA and duplex DNA. The method also extends its application for the direct detection of influenza viral panhandle structure and HIV-1 -1 ribosomal frameshift inducing RNA hairpin by designing complementary PNA strands incorporated with ^{bt}U monomer. However, in the current study the sensitivity and fluorescent enhancement of ^{bt}U monomer is limited with only 2 to 6-fold and requires excitation at 319 nm, which make it hard to employ under in vivo conditions. Our future research goal will be focused on designing and developing novel PNA monomers displaying near infra-red emission, with fluorescence enhancement on the order of several hundred-fold. The possible ways of achieving this goal include extending π -conjugation on uracil monomer as shown in **figure 4.1a-c** and incorporating new green fluorescent protein (GFP) based PNA monomers (**figure 4.1d**). Due to its high sensitivity and fluorescence efficiency, the application of GFP has been extensively studied in the live cell imaging.¹⁴ As reported earlier for small molecules binding to RNA aptamers, it is possible to program the fluorescent spectrum of GFP like chromophore for different wavelength by simple substituent modification at R, R' and R'' (**figure 4.1d**).¹⁵ Similar to thiazole orange (TO) and benzo[c,d]indole-quinoline (BIQ) (**figure 1.6a,b**), the proposed GFP like PNA monomer (**figure 4.1d**) expected to

serve as universal base surrogate with remarkable photophysical properties.¹⁴⁻¹⁷ Whereas, the modified π -extended uracil monomers (**figure 4.1a-c**) is expected to show fluorescent turn-on properties upon binding to complimentary Watson-Crick A-U base pair. Therefore, fluorescent dsRNA binding PNAs will facilitate understanding and mapping of complex secondary structures of RNA in biological environments in a programmable fashion.

Figure 4.1. Proposed near infra-red fluorescent turn-on PNA monomers derived from uracil (**a-c**) and GFP like chromophore (**d**).

Despite the several advantages of PNAs over natural nucleic acids and analogues with traditional sugar phosphate backbone, a major demerit of PNA is its cellular uptake and therefore its activity. In order to facilitate the improved solubility and cellular delivery, PNAs have been conjugated with amino acids like lysine and arginine having positive charge,^{18, 19} cell-penetrating peptides,²⁰ and/or carbohydrates such as neamine.²¹⁻²⁴

Meanwhile, cytosine requires lower pH condition for the duplex binding with Watson-Crick G-C base pairs through the Hoogsteen pairing.²⁵ Although, PNA L monomer is selective for duplex RNA over single-stranded RNA and duplex DNA, it suffers from complex synthetic protocol and is unable to improve the

cellular uptake. In order to surpass this, we have designed and developed a PNA R monomer having guanidine moiety attached to the PNA backbone through a two-carbon linker.²⁶ The chemical synthesis of R residue is simple, facile and involves only few well reported steps.²⁶ Similar to L monomer, R monomer shows a strong and selective binding to Watson–Crick G-C base pair in RNA duplex over DNA duplex and single-stranded RNA under physiological pH conditions as confirmed by PAGE, 2-aminopurine fluorescence titration and thermal melting experiments.¹⁸ In addition, compared to L, R monomer is more specific in binding to Watson–Crick G-C pair over an A-U pair. The cellular uptake of a PNA oligomer incorporated with R residues were tested by attaching carboxyfluorescein at N-terminal end over Sf9 cells. The confocal imaging reveals for the presence of PNA sequence inside cytoplasm of Sf9 cells. To further boost the delivery of PNAs both inside the cytoplasm and in the nucleus and to improve PNAs biological activity, one may need to incorporate multiple G-C pair recognizing R residues coupled with another cell-penetrating moieties such as neamine.^{19, 23}

The presence of single stranded loop in between duplex RNA structure in naturally occurring RNAs sometime limits our duplex binding strategy to achieve remarkable binding at sub-nanomolar level and thenceforth its biological activity. Currently, we are exploring novel PNA oligomers which can recognize both structured and single stranded regions by combining chemically modified dbPNAs (T, L or R, Q and E or S, **Figure 1.5**) and traditional unmodified antisense PNAs (A, G, C and T). Meanwhile, covalent binding of dbPNAs with

the complementary RNA duplex regions provides an alternative strategy to achieve a stronger binding affinity with moderate length.^{27, 28}

In summary, in this thesis we have reported a method in employing PNAs as probes to (i) selectively detect duplex RNAs at physiological pH conditions by designing PNA residue ^btU monomer, and (ii) achieve pH independent duplex binding of Watson–Crick G-C base pairs by incorporating into PNAs with R residues, which may show an increased solubility and cellular uptake. Both ^btU and R may be further modified to improve the binding properties, fluorescence light-up effect, and cellular uptake efficiency. The work provides the foundation for further developing RNA duplex structure-specific PNA-based probes and therapeutics.

References

1. Serganov, A.; Patel, D. J. *Nat. Rev. Genet.* **2007**, 8, 776-790.
2. Butcher, S. E.; Pyle, A. M. *Acc. Chem. Res.* **2011**, 44, 1302-1311.
3. Kaur, H.; Babu, B. R.; Maiti, S. *Chem. Rev.* **2007**, 107, 4672-4697.
4. Meister, G.; Tuschl, T. *Nature* **2004**, 431, 343-349.
5. Plasterk, R. H. *Cell* **2006**, 124, 877-881.
6. Li, N.; Li, Q.; Tian, X.-Q.; Qian, H.-Y.; Yang, Y.-J. *Am. J. Cardiovascular Drugs* **2014**, 14, 367-376.
7. Lim, K. R.; Maruyama, R.; Yokota, T. *Drug. Des. Devel. Ther.* **2017**, 11, 533-545.
8. Nan, Y.; Zhang, Y. J. *Front. Microbiol.* **2018**, 9, 750.
9. Devi, G.; Zhou, Y.; Zhong, Z.; Toh, D. F.; Chen, G. *Wiley Interdiscip Rev. RNA* **2015**, 6, 111-128.
10. Deere, J.; Iversen, P.; Geller, B. L. *Antimicrob. Agents Chemother.* **2005**, 49, 249-255.
11. Armitage, B. A. *Drug Discov Today* **2003**, 8, 222-228.
12. Krishna, M. S.; Toh, D. K.; Meng, Z.; Ong, A. A. L.; Wang, Z.; Lu, Y.; Xia, K.; Prabakaran, M.; Chen, G. *Anal. Chem.* **2019**, 91, 5331-5338.
13. Devi, G.; Yuan, Z.; Lu, Y.; Zhao, Y.; Chen, G. *Nucleic Acids Res.* **2014**, 42, 4008-4018.
14. Newman, R. H.; Fosbrink, M. D.; Zhang, J. *Chem. Rev.* **2011**, 111, 3614-3666.
15. Paige, J. S.; Wu, K. Y.; Jaffrey, S. R. *Science* **2011**, 333, 642-646.
16. Sato, T.; Sato, Y.; Nishizawa, S. *J. Am. Chem. Soc.* **2016**, 138, 9397-9400.
17. Yoshino, Y.; Sato, Y.; Nishizawa, S. *Anal. Chem.* **2019**, 91, 14254-14260.
18. Nielsen, P. E.; Haaime, G.; Lohse, A.; Buchardt, O. *Angew. Chem. Int. Ed.* **1996**, 35, 1939-1942.
19. Gupta, P.; Muse, O.; Rozners, E. *Biochemistry* **2012**, 51, 63-73.
20. Kilk, K.; Langel, U. *Methods Mol. Biol.* **2005**, 298, 131-41.
21. Mishra, R.; Su, W.; Pohmann, R.; Pfeuffer, J.; Sauer, M. G.; Ugurbil, K.; Engelmann, J. *Bioconjugate Chem.* **2009**, 20, 1860-1868.
22. Riguete, E.; Tripathi, S.; Chaubey, B.; Desire, J.; Pandey, V. N.; Decout, J. L. *J Med Chem* **2004**, 47, 4806-4809.
23. Kesu, J.; Patil, K. M.; Kumar, S. R.; Shu, Z.; Yong, H. Y.; Zimmermann, L.; Ong, A. A. L.; Toh, D. K.; Krishna, M. S.; Yang, L.; Decout, J. L.; Luo, D.; Prabakaran, M.; Chen, G.; Kierzek, E. *Bioconjugate Chem.* **2019**, 30, 931-943.
24. Ong, A. A. L.; Tan, J.; Bhadra, M.; Dezanet, C.; Patil, K. M.; Chong, M. S.; Kierzek, R.; Decout, J. L.; Roca, X.; Chen, G. *Molecules* **2019**, 24, (16).
25. Li, M.; Zengeya, T.; Rozners, E. *J. Am. Chem. Soc.* **2010**, 132, 8676-8681.
26. Krishna, M. S.; Wang, Z.; Zheng, L.; Bowry, J.; Ong, A. A. L.; Mu, Y.; Prabakaran, M.; Chen, G. *Biochemistry* **2019**, 58, 3777-3788.
27. Takasugi, M.; Guendouz, A.; Chassignol, M.; Decout, J. L.; Lhomme, J.; Thuong, N. T.; Helene, C. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, 88, 5602-5606.
28. Decout, J. L.; Lekchiri, Y.; Lhomme, J. J. *Photochem. Photobio.* **1992**, 55, 639-645.

Appendix

Appendix 1.1. ^1H , ^{13}C NMR Spectra for the synthesized compounds used in chapter 2

5IU mon est; CDCl₃, AV500

5IU mon est; CDCl₃, BBFO2 400

UBth Monomer ester-DMSOd6

UBth Monomer Ester 13C DMSOd6

Krishna_UBth monomer

UBth Monomer 13C 2nd, DMSO

Appendix A1.2. MALDI-TOF analysis of PNA oligomers studied in chapter 2

Table A1. MALDI-TOF data of PNA sequences studied in chapter 2.

PNA	Oligomer Sequence	Chemical Formula	Calculated MW	Obtained MW
P5-XL	H ₂ N-Lys- <u>X</u> LTLTTL-CONH ₂	C ₉₈ H ₁₂₆ N ₃₈ O ₂₇ S ₄	2394.85	2397.78
P5-LXL	H ₂ N-Lys-TL <u>X</u> LTTL-CONH ₂	C ₉₈ H ₁₂₆ N ₃₈ O ₂₇ S ₄	2394.85	2398.45
P5-LXT	H ₂ N-Lys-TLTL <u>X</u> TTL-CONH ₂	C ₉₈ H ₁₂₆ N ₃₈ O ₂₇ S ₄	2394.85	2398.45
P5-TXT	H ₂ N-Lys-TLTLTX <u>T</u> L-CONH ₂	C ₉₈ H ₁₂₆ N ₃₈ O ₂₇ S ₄	2394.85	2398.30
P4-CXT	H ₂ N-Lys-TLTC <u>X</u> TTL-CONH ₂	C ₉₈ H ₁₂₆ N ₃₈ O ₂₈ S ₃	2380.50	2380.22
P5-LXLXT	H ₂ N-Lys-TL <u>X</u> L <u>X</u> TTL-CONH ₂	C ₁₀₅ H ₁₂₈ N ₃₈ O ₂₇ S ₅	2512.84	2516.90
P5-LXXT	H ₂ N-Lys-TLTL <u>X</u> XTL-CONH ₂	C ₁₀₅ H ₁₂₈ N ₃₈ O ₂₇ S ₅	2512.84	2516.53
P5t-TXTTT	H ₂ N-Lys-TLT <u>X</u> TTL-CONH ₂	C ₉₉ H ₁₂₇ N ₃₇ O ₂₉ S ₃	2395.52	2394.04
P5t-TTXXT	H ₂ N-Lys-TLTT <u>X</u> TTL-CONH ₂	C ₉₉ H ₁₂₇ N ₃₇ O ₂₉ S ₃	2395.52	2395.46
IR-1X	H ₂ N-Lys-TLT <u>X</u> TQTLLL- CONH ₂	C ₁₂₂ H ₁₆₃ N ₅₁ O ₃₂ S ₅	3016.27	3016.67
L4-TXL	H ₂ N-Lys-LLT <u>X</u> LL-CONH ₂	C ₇₅ H ₉₇ N ₃₁ O ₁₇ S ₅	1865.10	1865.14
L4-LXT	H ₂ N-Lys-LL <u>X</u> TLL-CONH ₂	C ₇₅ H ₉₇ N ₃₁ O ₁₇ S ₅	1865.10	1864.90

The oligomers P4, P5, L4, and IR-1 were previously studied. The letter “X” represents a ^{bt}U residue in the sequences.

Appendix 2.1. ^1H , ^{13}C NMR Spectra for the synthesized compounds used in chapter 3

DiCbz_ethyl_aminioacid_dmsod6

DiCbz_aminopropionic_acid_dmsd_13C

R_ester2_dmsod6

R_ester3_dmsod6_13C

R_acid_dmsod6

R_acid2_dmsod6_13C

Appendix A2.2. MALDI-TOF analysis of PNA oligomers studied in chapter 3

Table A2. MALDI-TOF data of PNA sequences studied in chapter 3.

PNA	Oligomer Sequence	Chemical Formula	Calculated MW	Obtained MW
P1	H ₂ N-Lys-TCTCTTTC-CONH ₂	C ₉₁ H ₁₂₄ N ₃₈ O ₃₀	-	-
P1-1R4	H ₂ N-Lys-TCT <u>R</u> TTTC-CONH ₂	C ₈₉ H ₁₂₅ N ₃₈ O ₂₉	2189.95	2191.10
P1-1R5	H ₂ N-Lys-TCTC <u>R</u> TTTC-CONH ₂	C ₈₈ H ₁₂₄ N ₃₉ O ₂₈	2174.95	2176.44
P5	H ₂ N-Lys-TLTLTTTL-CONH ₂	C ₉₁ H ₁₂₄ N ₃₈ O ₂₇ S ₃	-	-
P5-1R4	H ₂ N-Lys-TL <u>T</u> RTTTL-CONH ₂	C ₈₉ H ₁₂₅ N ₃₈ O ₂₇ S ₂	2221.90	2245.05 [M + Na ⁺]
P5-3R	H ₂ N-Lys-TR <u>T</u> RTT <u>T</u> R-CONH ₂	C ₈₅ H ₁₂₇ N ₃₈ O ₂₇	2113.19	2152.39 [M + K ⁺]
IR-1	H ₂ N-Lys-TLTTTQTL <u>L</u> L-CONH ₂	C ₁₁₅ H ₁₆₁ N ₅₁ O ₃₂ S ₄	-	-
IR-1R	H ₂ N-Lys-TLTTTQTR <u>L</u> L-CONH ₂	C ₁₁₃ H ₁₆₃ N ₅₁ O ₃₂ S ₃	2842.19	2842.22
IR-3R	H ₂ N-Lys-TRTTTQTR <u>L</u> R-CONH ₂	C ₁₀₉ H ₁₆₇ N ₅₁ O ₃₂ S	2735.92	2753.32 [M + Na ⁺]
IR-ASP	H ₂ N-Lys-CCCTTGTTTCT-CONH ₂	C ₁₂₅ H ₁₆₉ N ₅₃ O ₄₀	3054.06	3056.26
L4	H ₂ N-Lys-LLTTLL-CONH ₂	C ₆₈ H ₉₅ N ₃₁ O ₁₇ S ₄	-	-
L4-RR	H ₂ N-Lys-LLT <u>T</u> RR-CONH ₂	C ₆₄ H ₉₇ N ₃₁ O ₁₇ S ₂	1636.80	1675.10 [M + K ⁺]
L4-RTTR	H ₂ N-Lys-L <u>R</u> TT <u>R</u> L-CONH ₂	C ₆₄ H ₉₇ N ₃₁ O ₁₇ S ₂	1636.80	1675.31 [M + K ⁺]
L4-CTTC	H ₂ N-Lys-LCTTCL-CONH ₂	C ₆₈ H ₉₅ N ₃₁ O ₁₉ S ₂	1714.83	1718.01
P1-CF	CF-H ₂ N-Lys-TCTCTTTC-CONH ₂	C ₁₁₂ H ₁₃₄ N ₃₈ O ₃₆	2588.53	2598.21
P5-3R-CF	CF-H ₂ N-Lys-TR <u>T</u> RTT <u>T</u> R-CONH ₂	C ₁₀₆ H ₁₄₀ N ₃₈ O ₃₃	2474.52	2516.53 [M + K ⁺]

^aThe oligomers P1, P5, L4 and IR-1 were previously characterized (1,2,6,7).

“CF” represents a carboxyfluorescein attachment.

Hello.!!

I am Krishna M. S., a graduate student at Nanyang Technological University, Singapore. It is my privilege to briefly introduce myself to you.

In 2008, I obtained my bachelor's degree in chemistry from Bangalore University and was placed among top 15 students out of several hundred. Later in 2009, I was enrolled the master's program in chemistry in the Department of Chemistry, Central college campus, Bangalore University, Bangalore. Here, I got strong theoretical and practical training in different aspects of chemistry, which include broad area of inorganic, organic, physical and analytical chemistry in the first year. I specialized in organic chemistry during the second year of my MSc program. In 2011, I cleared Council of Scientific and Industrial Research-Junior Research Fellowship (CSIR-JRF), securing all India ranking 69 in chemical science.

With this background, I moved to Department of Chemistry, Indian Institute of Technology, Kanpur, and spent my time as a junior and senior research fellow from 2012-2015. Here, my research interest was focused on the rational design and development of novel metal-organic framework and lattice inclusion compounds. I acquired technical skills, including the synthesis and characterization of various organic compounds and inorganic complexes having great importance and computer-based skills such as solving the crystal structures of the compounds based on X-ray crystallography. During this period, I attended an international summer school on "Topological Methods for Expert Systems in Materials Science" at Samara State University, Samara, Russia.

In 2016, I joined Nanyang Technological University, Singapore to pursue my doctoral degree under the supervision of Asst. Prof. Chen Gang. My graduate research interest was focused on (i) the rational design and synthesis of duplex binding Peptide Nucleic Acid (dbPNA) monomers displaying sequence and structure specific binding to RNA duplex through the Hoogsteen hydrogen bonding and (ii) fluorescent PNA monomers, which can be used as a diagnosing

tool to understand the nucleic acid microenvironment and structural polymorphism that prevail in the nature.

Overall, the eight years of my research journey helped me to develop many skills in different fields of science from synthesis, characterization, to applications. I had an opportunity to mentor more than twenty students for their masters/undergraduate dissertations, summer research, ureca projects, Nanyang Research Program (NRP) and Science Mentorship Programs (SMP). I am very keen to acquire the knowledge and build my career in science.

Thanking you,

Sincerely,

Krishna.

CURRICULUM VITAE

Shivakumar M. Krishna

Doctoral Student, Chemistry and Biological Chemistry (CBC),
School of Physical and Mathematical Science (SPMS), Nanyang Technological
University (NTU), Singapore, 639798

Email: krishnam003@e.ntu.edu.sg; krishnams.orgchem@gmail.com,

Mob: +65 9891 4336

Academic Qualifications

2016-2019	PhD in Chemistry, Nanyang Technological University (NTU), SINGAPORE
2009-2011	MSc in Chemistry, Bangalore University, INDIA
2005-2008	BSc in Chemistry, Bangalore University, INDIA

Summary of Working Experience

2016-Present	Graduate student, CBC, SPMS, NTU, SINGAPORE
2015-2016	Guest faculty, Tumkur University, INDIA
2014-2015	Senior Research Fellow, IIT Kanpur, INDIA
2012-2014	Junior Research Fellow, IIT Kanpur, INDIA
2011-2012	Research Assistant, Bangalore University, INDIA

Awards and Achievements

2011	Recipient of scholarship from Vision Group of Science and Technology, Government of Karnataka, INDIA
2011	Qualified Joint CSIR-UGC NET (All in India Rank 69), INDIA
2012	CSIR-Junior Research Fellowship, INDIA
2014	CSIR-Senior Research Fellowship, INDIA
2015	Qualified GATE in Chemical Science, INDIA
2016	NTU Research Scholarship, SINGAPORE
2016	Receipt of Fellowship to Pursue PhD at University of Heidelberg, GERMANY (declined)

Technical Skills

- A strong research background to design and synthesize organic compounds and inorganic complexes having academic and industrial importance.
- Ability to handle the instruments like NMR, MALDI-TOF, LCMS, GCMS, UV Spectrophotometer, Fluorescence spectrophotometer, FT-IR, Powder XRD, HPLC, CD spectropolarimeter, PCR, Quantachrome etc.,
- Ability to analyze, interpret and characterize the compounds using different spectroscopic techniques.
- Purifying compounds using techniques like column chromatography, sublimation, distillation and crystallization methods.
- A strong research background in solid phase peptide synthesis and non-denatured polyacrylamide gel electrophoresis (PAGE).
- Good oral and written communication skills. Use of English at a suitable standard in the preparation of manuscripts and presentations for scientific journals, meetings, conferences and classes.

Computer Skills

- Adequate knowledge on the use of various computer programs like MS Office, ChemOffice (Chemdraw and Chem3D), Origin and Chemical search engines (Scifinder & Reaxys).
- Proficient knowledge to solve and analyze the crystal structure using the software's like *winGX*, *OlexSys*, *Mercury*, *Diamond* and *ToposPro*.
- A good knowledge of solving organic compounds using NMR software *TopSpin*.

Teaching Experiences

2014-2015	Teaching Assistant for CHM101A practical's, Department of Chemistry, IIT Kanpur, INDIA.
2015-2016	Guest faculty to teach undergraduate Chemistry both theory and practical's, Department of Chemistry, Tumkur University, INDIA.
2018-2019	Teaching Assistant for CM1031, CM2061 and CM3061 synthetic organic chemistry practical's, CBC, SPMS, NTU, SINGAPORE.

Student Mentorships

Students from different groups worked for their master thesis, final year projects (FYP), Ureca, summer research, Nanyang research program (NRP) and science mentorship program (SMP).

Number of Students Trained – 20

• Master Thesis		
Jogesh Bowry	July 2018-Dec 2018	Uni. of Southampton, UK
• Undergraduate Thesis		
Lina Nur Lyana Binte Abdul Rahim	Aug 2019-Dec 2019	NTU, Singapore
Yeo Li Sin	Aug 2019-Dec 2019	NTU, Singapore
Zheng Anning	Jan 2018-May 2018	Uni. of Bristol, UK
Ang Min Hui	Jan 2018-May 2018	NTU, Singapore
Goh Celestine	Jan 2017-May 2017	NTU, Singapore
• Ureca projects		
Dashreen Kaur D/O Gurmit Singh	Aug 2018-Jun 2019	NTU, Singapore
Tee Shiau Hang	Aug 2017-Jun 2018	NTU, Singapore
Hou Hay Sy	Aug 2017-Jun 2018	NTU, Singapore
Ang Min Hui	Aug 2016-Jun 2017	NTU, Singapore
• Summer Projects		
Kenneth Yeo	May 2019-Jul 2019	NTU, Singapore
Teng Hon		
Lam Min Hui	May 2019-Jul 2019	NTU, Singapore
Kelvin Poh Guo En	May 2019-Jul 2019	NTU, Singapore
Lee Qian Ping	May 2018-Jul 2018	NTU, Singapore
Toh Yun Ting, Jacqueline	May 2018-Jul 2018	NTU, Singapore
Wesley Lim Jun Jie	May 2017-Jul 2017	NTU, Singapore
• Nanyang Research Program (NRP) (secured Gold Award)		

Li Yixin	Apr 2017-Apr 2018	River Valley High School
Tan Qi Lun	Apr 2017-Apr 2018	Innova Junior College
Stephanie		
<ul style="list-style-type: none"> Science Mentorship Program (SMP) (3M Best Project Award, MOE Singapore) 		
Nallapuraju Atchyuta	Feb 2018-Sep 2018	Raffles Institution
Srinivasa Varma		
Vishal Vijay Kumar	Feb 2018-Sep 2018	Raffles Institution
Daniel Loo Kang Le	Feb 2018-Sep 2018	Raffles Institution

Special Acquirements

- Completed certified course, HWG702 University Teaching for Teaching Assistants, Aug-Nov 2017.
- Completed certified course, Diploma in Information Technology, 2008.

International Workshops

Aug 2014 International Summer School on "*Topological Methods for Expert Systems in Materials Science*" at Samara State University, Samara, RUSSIA

Oral Presentation

Jul 2018 Developing Fluorescent Light-Up Probes for Detecting RNA Structures. Singapore RNA Biology Alliance Group Meeting, NUS, SINGAPORE

Poster Presentations

Mar 2019 Developing Fluorescent Light-Up Probes for Detecting RNA Structures. Singapore-Paris Chemical Biology Workshop, NTU, SINGAPORE

Sep 2018 Developing Fluorescent Light-Up Probes for Detecting RNA Structures. 4th RNA Biology Symposium, NUS, SINGAPORE

Mar 2015 A Fluorescent Metal-Organic Framework Based on Pyrene-Tetra Carboxylic Acid: Solvatochromism and Sensing of Nitroaromatic Compounds. National Symposium on Radiation and Photochemistry (NSRP-2015), IIT Kanpur, INDIA

Citation Summary

Number of publications co-authored: 10

Google Scholar Citations: 60 h-index: 4 i10-index: 2

<https://scholar.google.co.in/citations?user=scZtVYkAAAAJ&hl=en>

List of First Author Publications

1. **Krishna, M. S.**; Wang, Z.; Zheng, L.; Bowry, J.; Ong, A. A. L.; Mu, Y.; Prabakaran, M.; Chen, G. *Biochemistry* **2019**, 58, 3777-3788.
2. **Krishna, M. S.**; Toh, D. -F. K.; Meng, Z.; Ong, A. A. L.; Wang, Z.; Lu, Y.; Xia, K.; Prabakaran, M.; Chen, G. *Anal. Chem.* **2019**, 91, 5331-5338.

Referees

- Asst. Prof. **Chen Gang** (PhD advisor)
Chemistry and Biological Chemistry, School of Physical and Mathematical Science, Nanyang Technological University (NTU), Singapore,
Email: rnachen@ntu.edu.sg
- Prof. **Xing Bengang** (Course Instructor)
Chemistry and Biological Chemistry, School of Physical and Mathematical Science, Nanyang Technological University (NTU), Singapore,
Email: bengang@ntu.edu.sg