Part I Targeting RNA duplexes with modified triplex-forming oligonucleotides Part II Allostery within a Watson-Crick RNA duplex and its application in studying RNA recognition by RIG-I

Zhou, Yuan

2014

Zhou, Y. (2014). Part I Targeting RNA duplexes with modified triplex-forming oligonucleotides Part II Allostery within a Watson-Crick RNA duplex and its application in studying RNA recognition by RIG-I. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/56303>

<https://doi.org/10.32657/10356/56303>

**PART I: TARGETING RNA DUPLEXES WITH MODIFIED
TRIPLEX-FORMING OLIGONUCLEOTIDES**

**PART II: ALLOSTERY WITHIN A WATSON-CRICK RNA
DUPLEX AND ITS APPLICATION IN STUDYING RNA
RECOGNITION BY RIG-I**

ZHOU YUAN

SCHOOL OF PHYSICAL AND MATHEMATICAL SCIENCES

2014

**PART I: TARGETING RNA DUPLEXES WITH MODIFIED
TRIPLEX-FORMING OLIGONUCLEOTIDES**

**PART II: ALLOSTERY WITHIN A WATSON-CRICK RNA
DUPLEX AND ITS APPLICATION IN STUDYING RNA
RECOGNITION BY RIG-I**

ZHOU YUAN

School of Physical and Mathematical Sciences

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

2014

ACKNOWLEDGEMENT

Foremost, I would like to express my deepest gratitude to my supervisors, Associate Professor LI Tianhu and Assistant Professor CHEN Gang, for their strong support, constant guidance, conscientious suggestions during my graduate study in School of Physical and Mathematical Sciences in Nanyang Technological University.

Here I would like to acknowledge Professors KIERZEK Ryszard, Susana GEIFMAN-SHOCHAT, PHAN Anh Tuan, XING Bengang, SHAO Fangwei and ZHANG Dawei for providing research materials and access to their instruments.

I am grateful to Dr. DEVI Gitali, Dr. NARAYANAN Ranjani, Ms TONG Cailing, Ms YUAN Zhen, Mr ZHONG Zhensheng, and Mr LIU Yiran in our research group, for their great help in scientific discussions and advice.

I also want to express my thanks for all administrative and technical staff in our division.

Last but not most, I want to show my gratitude to my parents from the bottom of my heart, who bring me up. Here I dedicate this thesis to them.

TABLE OF CONTENTS

ACKNOWLEDGEMENT	I
TABLE OF CONTENTS	II
ABSTRACT	III
LIST OF ABBREVIATION	VII
Part I: <i>Targeting RNA Duplexes with Modified Triplex-Forming Oligonucleotides</i>	
- INTRODUCTION	1
- MATERIALS AND METHODS	6
- RESULTS AND DISCUSSION	11
- CONCLUSION	39
- REFERENCES	40
Part II: <i>Allostery within a Watson-Crick RNA Duplex and Its Application in Studying RNA Recognition by RIG-I</i>	
- INTRODUCTION	45
- MATERIALS AND METHODS	51
- RESULTS AND DISCUSSION	58
- CONCLUSION	88
- REFERENCES	89
LIST OF PUBLICATIONS	94

ABSTRACT

Part I: Targeting RNA Duplexes with Modified Triplex-Forming Oligonucleotides

Triplex is emerging as an important RNA tertiary structure motif, in which consecutive base triples form between a duplex and a third strand. RNA duplex region is also often functionally important site for protein binding. Thus, Triplex-Forming Oligonucleotides (TFOs) may be developed to regulate various biological functions involving RNA such as viral ribosomal frame-shifting and reverse transcription. How chemical modification in TFOs affects RNA triplex stability, however, is not well understood. Here, we incorporated Locked Nucleic Acid (LNA), 2-thio U, and 2'-O methyl modified residues in a series of all pyrimidine RNA TFOs, and studied the binding to two RNA hairpin structures. 12-Base-triple major-groove pyrimidine-purine-pyrimidine triplex structures form between the RNA/DNA hairpins and the complementary RNA TFOs. UV-absorbance-detected thermal melting studies reveal that the LNA and 2-thio U modifications in TFOs strongly enhance triplex formation with both parental RNA and DNA duplex regions. In addition, we found that incorporation of 2'-O methyl modified residues in a TFO destabilizes and stabilizes triplex formation with RNA and DNA duplex regions, respectively. The (de)stabilization of RNA triplex formation may be facilitated through modulation of van der Waals contact, base

stacking, hydrogen bonding, backbone preorganization, geometric compatibility, and/or dehydration energy. Better understanding of the molecular determinants of RNA triplex structure stability lays the foundation for designing and discovering novel sequence-specific duplex-binding ligands as diagnostic and therapeutic agents targeting RNA.

Part II: Allostery within a Watson-Crick RNA Duplex and Its Application in Studying RNA Recognition by RIG-I

RNA has rich local and global conformational dynamics, which may often result from allosteric effects. However, allostery within a Watson-Crick RNA duplex is not well known. Here, we report the direct observation that the local structural dynamics of a Watson-Crick RNA duplex is modulated allosterically by the specific binding of human Retinoic acid-inducible gene I (RIG-I) C-terminal domain (CTD). To detect the allosteric effect, we incorporated into Watson-Crick RNA duplexes a fluorescent analogue of adenine, 2-aminopurine (2AP), which is a sensitive probe of RNA local structure and dynamics. 2AP fluorescence increases upon ligand binding, suggesting ligand-induced changes of the local dynamics of the (2AP)-U pair. The enhancement of 2AP fluorescence is (1) distance dependent with the (2AP)-U pair positioned within about 13 base pairs away from the ligand binding site, and (2) structure dependent with no fluorescence changes observed upon the incorporation of a non-Watson-Crick pair in between the (2AP)-U pair and the ligand binding site. Binding of full-length RIG-I and a triplex-forming oligonucleotide (TFO) also induces 2AP fluorescence enhancement. Allosteric 2AP fluorescence changes facilitate quantifications of binding affinities. Thus, Watson-Crick RNA duplexes may not be considered as merely rigid structural scaffolds. Ligand binding- and tertiary interaction-induced allostery within Watson-Crick DNA/RNA duplexes may facilitate the regulation of functions

involving nucleic acid structure dynamics and nucleic acid-protein assembly, such as catalysis, metabolite sensing, helicase/chaperone activity, and the detection of single mismatches and damaged/modified residues in nucleic acid duplexes.

LIST OF ABBREVIATIONS

TFOs	Triplex-Forming Oligonucleotides
2'-OMe	2'-O methyl
LNA	Locked Nucleic Acid
2',4'-BNA	2',4'-Bridged Nucleic Acid
U ^L	LNA U
s ² U	2-thio U
s ² U ^M	2'-OMe 2-thio U
U ^M	2'-OMe U
C ^M	2'-OMe C
SPR	Surface Plasmon Resonance
rHP	RNA hairpin
dHP	DNA hairpin
<i>T</i> _m	melting temperature
EDTA	Ethylenediaminetetraacetic acid
TAE	Tris-Acetate, EDTA
MES	2-(N-morpholino)ethanesulfonic acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
CM5	carboxymethylated dextran coated sensor chip
RIG-I	Retinoic acid-inducible gene I
CARDs	caspase activation and recruitment domains

MAVS	mitochondrial antiviral-signaling protein
CTD	C-terminal domain
ssRNA	single-strand RNA
dsRNA	double-strand RNA
5'-ppp	5'-triphosphate
2AP	2-aminopurine
LB	Luria broth
OD	optical density
IPTG	isopropylthiogalactoside
DTT	dithiothreitol
FPLC	Fast Performance Liquid Chromatography
AMA	ammoniamethylamine
DMSO	dimethyl sulfoxide
TEA	triethylamine
TBDMS	tert-butyldimethylsilyl
RP-HPLC	reverse phase high performance liquid chromatography
TEAA	triethylammonium acetate
ACN	acetonitrile
DMT	dimethoxytrityl
APS	ammonium persulfate
K_D	equilibrium dissociation constant
TCSPC	time-correlated single photon counting

τ	lifetime
A	amplitudes
E	glutamic acid
K	lysine
PAGE	Polyacrylamide gel electrophoresis

Part I: Targeting RNA Duplexes with Modified Triplex-Forming Oligonucleotides

1. INTRODUCTION

Gene regulation, catalysis, and immunomodulation are some of the biological functions of RNAs. The diverse biological functions of RNA are determined by the complex structures of RNA stabilized by both secondary and tertiary interactions. Better understanding of the structures and functions of RNAs provides the foundation for designing nucleic acid-based diagnostics, therapeutics, and nanobiotechnologies (1-7).

Triplex is a biologically important RNA tertiary structure, that is found in pseudoknots and many large RNAs in which consecutive base triples form between RNA duplex and a third strand (8-18) in the major or minor groove of a Watson-Crick base paired stem. Non-Watson-Crick pairs in a double-stranded stem may also stabilize tertiary triplex structure formation, although the secondary structure is destabilized. RNA duplex region is also often functionally important for protein binding (19-25). Thus, it is of great therapeutic potential to design and discover chemically modified Triplex-Forming Oligonucleotides (TFOs) and other therapeutic ligands targeting RNA duplex regions to reprogram RNA-RNA and RNA-protein interactions in cells and viruses.

Base triple is a structure motif important for shaping RNAs into complex three-dimensional architectures for diverse biological functions (26). For example, base triples in tRNAs help maintain the L-shaped three-dimensional structure (27-30). A major-groove U·A-U base triple (Figures 1A) in viral TAR (trans-activation response element) RNAs facilitates specific recognition by small molecules and Tat proteins (31-35). Highly selective recognition of purine metabolites by purine-binding riboswitches results from the base triple formation with shape and chemical complementarity at the binding pocket (36). RNA triplex is a complex structure stabilized by multiple base triples formed between an RNA duplex and a third strand (of a natural RNA or an artificial nucleic acid). An RNA duplex may accommodate a third strand to form a major-groove and a minor-groove triplex, respectively.

A major-groove RNA triplex structure is stable in isolation by forming consecutive pyrimidine motif major-groove base triples such as U·A-U, C⁺·G-C, and U·G-C. Minor-groove RNA triplexes, e.g., A-minor motif triplex, are often found in pseudoknots and almost all large RNAs. A-minor motif minor-groove triplexes show intrinsic structural plasticity; the adenosine residues in the third strand may have various local structures in the minor groove and may even be replaced with other residues.

Major-groove pyrimidine·purine·pyrimidine triplex formation with DNA duplexes

(Figures 1 and 2) has been utilized for sequence-specific DNA cleavage (37,38) and inhibition of transcription (39). Many studies have been reported to use chemically modified TFOs to improve the binding affinity with DNA duplexes (40,41). For example, TFO modification with 2'-O methyl (2'-OMe) (Figure 1B) facilitates triplex formation (42,43) with parental DNA duplexes. Incorporation of Locked Nucleic Acid (LNA) (also known as 2',4'-Bridged Nucleic Acid (2',4'-BNA)) monomers (Figure 1B) significantly favors triplex formation with DNA duplexes (44-46). 2-thio U (Figure 1B) is a base modification, which has been incorporated in TFOs to effectively stabilize triplexes with parental DNA duplexes (47).

Surprisingly, there are very few studies about targeting RNA duplex with triplex formation. How incorporation of LNA and 2-thio U modified residues into TFOs affects RNA triplex stability is not well understood. In the thesis, LNA U (U^L), 2-thio U (s^2U), 2'-OMe 2-thio U (s^2U^M), 2'-OMe U (U^M), and 2'-OMe C (C^M) residues were incorporated into 12-nucleotide (nt) RNA TFOs with all pyrimidine residues, and the formation of major-groove triplex with the modified TFOs binding to RNA and DNA hairpins (Figure 2) were studied. The melting temperatures of RNA duplexes and triplexes were measured by UV-absorbance-detected thermal melting at various NaCl concentrations and pH's. Gel electrophoresis and Surface Plasmon Resonance (SPR) experiments were also used to further confirm the triplex formation.

Figure 1. (A) Base triples studied in this thesis. The N3 atom of cytosine base in TFOs needs to be protonated (shown in bold) to form C⁺-G-C base triple. s²U-A-U base triple is more stable than unmodified U-A-U base triple, partly because of enhanced van der Waals interaction between the thio group in 2-thio U base and H8 hydrogen (shown in bold) in adenine base. (B) Chemical structures of unmodified and modified RNA nucleosides: RNA U (U), 2'-OMe U (U^M), LNA U (U^L), 2-thio U (s²U), and 2'-OMe 2-thio U (s²U^M).

Figure 2. Hairpin and triplex structures studied in this thesis. TFO strands are shown in gray with the modified residues shown in black. (A) A 14-bp 32-nt RNA hairpin (rHP1). (B) A triplex made by TFO 1-R12 binding with rHP1. (C) A mutated (U-A inversion, shown in box) 14-bp 32-nt RNA hairpin (rHP2). (D) A hypothetical triplex formed between non-complementary TFO 1-R12 and rHP2. A mismatched base triple (U·U-A) is shown. (E) A 14-bp 32-nt DNA hairpin (dHP1), which is homologous to the RNA hairpin rHP1. (F) A triplex formed between TFO 1-R12 and dHP1. (G) A 12-bp 28-nt RNA hairpin (rHP3). (H) A triplex formed between TFO 3-R12 and rHP3.

2. MATERIALS AND METHODS

2.1 Oligonucleotides

Chemically synthesized and HPLC purified DNA hairpin 1 (dHP1), RNA hairpin 1 (rHP1), mutated RNA hairpin 2 (rHP2), RNA hairpin 3 (rHP3), and unmodified RNA TFOs (Figure 2) were purchased from Sigma-Aldrich in Singapore. Modified RNA TFOs were chemically synthesized, deprotected, and purified as reported (48-50). All oligonucleotides were characterized by ESI-MS. The sequences of TFOs are listed in Table 1.

Table 1. TFO sequences studied in this paper.

TFO	Sequence ^a											
1-D12 ^b	T	C	T	C	T	C	T	C	T	T	T	C
1-R12	U	C	U	C	U	C	U	C	U	U	U	C
Control	A	U	C	U	G	U	U	C	C	A	C	U
1-RU ^L 1	U	C	U	C	U	C	<u>U^L</u>	C	U	U	U	C
1-RU ^L 2	U	C	U	C	<u>U^L</u>	C	<u>U^L</u>	C	U	U	U	C
1-RU ^L 3	U	C	U	C	<u>U^L</u>	C	<u>U^L</u>	C	<u>U^L</u>	U	U	C
1-Rs ² U1	U	C	U	C	U	C	<u>s²U</u>	C	U	U	U	C
1-Rs ² U2	U	C	U	C	<u>s²U</u>	C	<u>s²U</u>	C	U	U	U	C
1-Rs ² U3	U	C	U	C	<u>s²U</u>	C	<u>s²U</u>	C	<u>s²U</u>	U	U	C
1-RU ^M 3	U	C	U	C	U ^M	C	U ^M	C	U ^M	U	U	C
1-M12	U ^M	C ^M	U ^M	C ^M	U ^M	C ^M	U ^M	C ^M	U ^M	U ^M	U ^M	C ^M
1-MU ^L 1	U ^M	C ^M	U ^M	C ^M	U ^M	C ^M	<u>U^L</u>	C ^M	U ^M	U ^M	U ^M	C ^M
1-MU ^L 2	U ^M	C ^M	U ^M	C ^M	<u>U^L</u>	C ^M	<u>U^L</u>	C ^M	U ^M	U ^M	U ^M	C ^M
1-MU ^L 3	U ^M	C ^M	U ^M	C ^M	<u>U^L</u>	C ^M	<u>U^L</u>	C ^M	<u>U^L</u>	U ^M	U ^M	C ^M
1-Ms ² U ^M 1	U ^M	C ^M	U ^M	C ^M	U ^M	C ^M	<u>s²U^M</u>	C ^M	U ^M	U ^M	U ^M	C ^M
1-Ms ² U ^M 2	U ^M	C ^M	U ^M	C ^M	<u>s²U^M</u>	C ^M	<u>s²U^M</u>	C ^M	U ^M	U ^M	U ^M	C ^M
1-Ms ² U ^M 3	U ^M	C ^M	U ^M	C ^M	<u>s²U^M</u>	C ^M	<u>s²U^M</u>	C ^M	<u>s²U^M</u>	U ^M	U ^M	C ^M
3-R12	C	C	U	C	U	C	C	U	C	C	C	U
3-Rs ² U2	C	C	U	C	<u>s²U</u>	C	C	<u>s²U</u>	C	C	C	U
3-RU ^L 2	C	C	U	C	<u>U^L</u>	C	C	<u>U^L</u>	C	C	C	U

^a All listed sequences have RNA residues unless otherwise noted. The modified nucleotides studied are LNA U (U^L), 2'-thio U (s²U), 2'-OMe U (U^M), 2'-OMe C (C^M), and 2'-OMe 2-thio U (s²U^M). In the nomenclature for TFOs,

the first number (1 or 3) indicates the TFO's complementary hairpin sequence (rHP1, dHP1, or rHP3), and the last number (1, 2, 3, or 12) indicates the number of modified residues. The first letters (D, R, and M) indicate the main backbone compositions of DNA, RNA and 2'-OMe RNA, respectively.

^b In TFO 1-D12, all 12 residues are unmodified deoxyribonucleotides.

Triplexes were annealed as follows:

1. 1.5 mL microtube containing RNA hairpin was placed inside a heating block for 5 minutes at 95 °C.
2. The tube was then transferred inside an ice block for 5 minutes in order to cool the tube to around 0 °C quickly.
3. The hairpin solution was then transferred to a tube with dried triplex-forming oligonucleotide, and the final concentrations of the TFO, hairpin, and triplex are all 5 µM.
4. The 1.5 mL microtube with 5 µM triplex was placed inside a heating block for 5 minutes at 65 °C.
5. After heating, heat block was powered off and the tube was left in the heat block for about 2 hours, and then the tube was left inside 4 °C refrigerator for at least three hours in order to form stable triplex structure.
6. Triplex was transferred into the Hellma UV quartz cuvette, and then silicon oil was added on the top of the RNA solution. The cuvettes were centrifuged at 4000 rpm for 10 minutes to evenly spread the silicon oil. The oil was added to prevent water evaporation.

2.2 UV-absorbance-detected thermal melting

Beckman DU 800 spectrophotometer with high performance transport and multiple-cell holder was used to measure UV absorbance versus temperature thermal melting. The spectrophotometer was connected to a computer for data collection and analysis. Using a Peltier temperature controller, the sample was heated up from 15 to 95 °C and then cooled down to 15 °C at a ramp rate of 0.2 °C/min (or 1 °C/min for measurement of duplexes formed between TFO strands and a 12-nt all purine strand, 5'-AGAGAGAGAAAG-3'), and the absorbance at 260 nm or 265 nm (only for 1-Rs²U3) was recorded every 0.5 °C. For UV-absorbance-detected thermal melting studies, the samples contained 5 µM hairpin and 5 µM TFO in 100-1000 mM NaCl, 20 mM MES (pH 5.0, 5.5, and 6.0) or 20 mM HEPES (pH 6.5, 6.8, and 7.0) and 0.5 mM EDTA. MES and HEPES buffers were used in this research because their pK_a values are relatively temperature independent (51). The first derivative curves of UV thermal melting were fit to Gaussian functions, and the temperatures with maximum first derivatives (at Gaussian peaks) in the melting (heating) curves were taken as the melting temperatures (T_{m1} for triplex to hairpin transition, and T_{m2} for hairpin to single strand transition). Hysteresis between heating and cooling curves was observed for triplex to hairpin transitions. For hairpin to single strand transition, no hysteresis was observed. Due to the hysteresis, equilibrium thermodynamic parameters for the triplex to hairpin transitions were not obtained. Equilibrium thermodynamic parameters for the duplex and hairpin formation were calculated

by fitting to a two-state model with the MeltWin program (52).

2.3 Surface Plasmon Resonance (SPR) assay

All SPR experiments were done on a Biacore 3000 (BIAcore AB, GE Healthcare) with a carboxymethylated dextran coated sensor chip (CM5) with a fixed flow rate of 5 μ L/min running buffer (200 mM NaCl, 0.5 mM EDTA, 20 mM MES at pH 6.0) at 25 $^{\circ}$ C. Biotinylated RNA hairpin (5'-biotin-TTTTUAGAGAGAGAGAAAGUUUCGACUUUCUCUCUCUA-3' and 5'-biotin-TTTTGGAGAGGAGGGAUUUUUCCCUCCUCUCC-3', with four DNA thymine (T) residues were incorporated as a linker between biotin and RNA hairpins, rHP1 and rHP3) and serially diluted TFOs were prepared in the same running buffer. Biotinylated RNA hairpin was immobilized on sensor chip surfaces. A serial injection of unmodified and modified RNA TFOs, or an RNA control (20.6 pM, 61.7 pM, 185.2 pM, 555.6 pM, 1.67 nM, and 5.0 nM) provided the response of triplexes simultaneously. After a dissociation period with 1800 seconds, 30 μ L (3 min) of 10 mM HCl (53) was injected in order to completely remove the bound TFOs by denaturing the triplexes. After that, 150 μ L (15 min) of running buffer was injected to refold rHP1 and stabilize the baseline (54) to restore the TFO binding ability of the sensor chip for further experiments. All the sensorgrams were then corrected by subtraction of the control. Data were processed by globally fitting to a simple 1:1 interaction model using numeric integration to yield the individual affinities.

2.4 Gel electrophoresis

A 16.5 × 22 cm gel containing 12% acrylamide (Acrylamide/Bis-acrylamide = 19:1) was used in native polyacrylamide gel electrophoresis experiments. RNA hairpin and TFOs were both 1 μM, and they were incubated in 40 μL of buffer (100 mM NaCl, 20 mM MES and 0.5 mM EDTA at pH 5.5), and left for 3 hours before loading. Then 8 μL of 35% glycerol was added into the 40 μL loading sample. The running buffer contains 10 mM NaCl, 2 mM MgCl₂, and 1× TAE (40 mM Tris-Acetate, 1 mM EDTA) at pH 6.0. The gel electrophoresis experiments were run at 160 V for 12-16 hours at 4 °C, and the gel was stained by ethidium bromide (1 μg/mL) for 30 minutes. After that the gel was imaged by a Typhoon scanner (GE Healthcare).

3. RESULTS AND DISCUSSION

Table 2. Melting temperatures of triplexes (T_{m1}) and hairpins (T_{m2})^a

TFO sequence	Duplex target	pH	[NaCl] (mM)	T_{m1} (°C)	T_{m2} (°C)
No TFO	rHP1	6.5	200	N/A	84.1
	rHP1	6.0	200	N/A	84.0
	rHP1	5.5	1000	N/A	91.3
	rHP1	5.5	500	N/A	87.5
	rHP1	5.5	300	N/A	86.0
	rHP1	5.5	200	N/A	83.5
	rHP1	5.5	100	N/A	79.3
	dHP1	6.5	200	N/A	73.7
	dHP1	6.0	200	N/A	73.5
	dHP1	5.5	200	N/A	74.4
	dHP1	5.5	100	N/A	66.3
	rHP2	6.0	200	N/A	83.8
	rHP2	5.5	200	N/A	83.8
	rHP3	7.0	200	N/A	91.9
	rHP3	6.0	200	N/A	92.9
	rHP3	5.5	100	N/A	92.2
1-R12	rHP1	6.5	200	<20.0	84.3
	rHP1	6.0	200	28.1	84.3
	rHP1	5.5	1000	41.0	89.4
	rHP1	5.5	500	41.5	88.0
	rHP1	5.5	300	42.5	86.2
	rHP1	5.5	200	43.6	83.5
	rHP1	5.5	100	46.4	79.3
	dHP1	6.5	200	29.4	74.8
	dHP1	6.0	200	45.3	73.4
	dHP1	5.5	100	(67.5)	(67.5)
1-RU ^L 1	rHP2	6.0	200	<20.0	84.7
	rHP1	6.5	200	<20.0	84.8
	rHP1	6.0	200	36.5	84.6
	rHP1	5.5	200	57.5	82.5
	rHP1	5.0	200	(83.5)	(83.5)
	dHP1	6.0	200	52.3	73.6
	dHP1	5.5	200	(73.5)	(73.5)
	dHP1	5.0	200	(73.5)	(73.5)
	rHP2	6.0	200	<20.0	85.4
	rHP2	5.5	200	26.7	83.8
	rHP2	5.0	200	46.4	83.1

1-RU ^L 2	rHP1	7.0	200	<20.0	84.6
	rHP1	6.5	200	31.8	85.1
	rHP1	6.0	200	45.4	84.6
	dHP1	6.0	200	57.0	73.6
	rHP2	6.0	200	<20.0	84.7
1-RU ^L 3	rHP1	7.0	200	<20.0	84.5
	rHP1	6.5	200	36.6	85.3
	rHP1	6.0	200	52.9	84.3
	dHP1	6.0	200	(75.4)	(75.4)
	rHP2	6.0	200	<20.0	84.5
1-Rs ² U1	rHP1	6.5	200	26.0	84.6
	rHP1	6.0	200	40.1	84.5
	rHP1	5.5	1000	54.4	90.9
	rHP1	5.5	500	53.9	87.0
	rHP1	5.5	300	55.8	86.7
	rHP1	5.5	200	56.9	83.5
	rHP1	5.5	100	58.3	80.0
	dHP1	6.0	200	56.2	74.0
	rHP2	6.0	200	<20.0	84.2
	rHP2	5.5	200	44.9	83.6
1-Rs ² U2	rHP1	7.0	200	<20.0	84.6
	rHP1	6.5	200	38.3	84.0
	rHP1	6.0	200	51.9	84.0
	rHP1	5.5	1000	64.4	90.9
	rHP1	5.5	500	63.9	87.0
	rHP1	5.5	300	64.6	85.9
	rHP1	5.5	200	66.2	83.5
	rHP1	5.5	100	68.2	79.6
	dHP1	6.0	200	60.0	75.0
	rHP2	6.0	200	<20.0	84.7
1-Rs ² U3	rHP1	7.0	200	23.7	84.9
	rHP1	6.5	200	41.9	84.1
	rHP1	6.0	200	54.4	83.1
	rHP1	5.5	1000	(90.9)	(90.9)
	rHP1	5.5	500	(87.0)	(87.0)
	rHP1	5.5	300	(86.0)	(86.0)
	rHP1	5.5	200	(84.0)	(84.0)
	rHP1	5.5	100	(80.2)	(80.2)
	dHP1	6.0	200	(74.1)	(74.1)
	rHP2	6.0	200	<20.0	83.9
1-RU ^M 3	rHP1	6.5	200	<20.0	83.1
	rHP1	6.0	200	<20.0	83.1
	rHP1	5.5	100	25.7	77.9
	rHP1	5.5	100	25.7	77.9

1-M12	rHP1	5.0	100	41.5	79.1
	dHP1	6.5	200	34.6	75.1
	dHP1	5.5	100	(68.7)	(68.7)
	rHP1	5.5	100	<20.0	78.2
	dHP1	6.5	200	37.1	75.4
1-MU ^L 1	dHP1	5.5	100	(69.5)	(69.5)
	rHP1	6.5	100	<20.0	79.2
	rHP1	6.0	200	<20.0	83.8
	rHP1	5.5	200	<20.0	82.5
	rHP1	5.5	100	<20.0	78.2
1-MU ^L 2	dHP1	7.0	1000	<20.0	80.6
	dHP1	6.8	1000	<20.0	82.1
	dHP1	6.5	200	41.4	73.7
	dHP1	6.0	1000	53.4	81.2
	dHP1	6.0	200	(73.5)	(73.5)
	rHP1	6.0	200	<20.0	84.0
	rHP1	5.5	100	<20.0	77.7
	dHP1	7.0	1000	<20.0	80.7
	dHP1	6.8	1000	39.6	82.1
	dHP1	6.5	200	(74.0)	(74.0)
1-MU ^L 3	dHP1	6.0	1000	(81.5)	(81.5)
	dHP1	6.0	200	(73.5)	(73.5)
	rHP1	6.0	200	<20.0	84.1
	rHP1	5.5	100	<20.0	78.2
	dHP1	7.0	1000	<20.0	80.7
	dHP1	6.8	1000	50.5	81.7
	dHP1	6.5	200	(74.0)	(74.0)
	dHP1	6.0	1000	(81.2)	(81.2)
	dHP1	6.0	200	(73.5)	(73.5)
	rHP1	5.5	100	<20.0	78.6
1-Ms ² U ^M 1	dHP1	7.0	1000	<20.0	80.6
	dHP1	6.8	1000	36.9	82.1
	dHP1	6.5	200	42.9	74.0
	dHP1	6.0	1000	(81.2)	(81.2)
	dHP1	6.0	200	(73.5)	(73.5)
1-Ms ² U ^M 2	rHP1	5.5	100	<20.0	79.1
	dHP1	7.0	1000	<20.0	80.7
	dHP1	6.8	1000	42.0	82.1
	dHP1	6.5	200	(74.0)	(74.0)
	dHP1	6.0	1000	(81.5)	(81.5)
1-Ms ² U ^M 3	dHP1	6.0	200	(73.5)	(73.5)
	rHP1	5.5	100	<20.0	78.5
	dHP1	7.0	1000	<20.0	80.7
	dHP1	6.8	1000	50.0	81.7

3-R12	dHP1	6.5	200	(74.0)	(74.0)
	dHP1	6.0	1000	(81.2)	(81.2)
	dHP1	6.0	200	(73.5)	(73.5)
	rHP3	7.0	200	<20.0	91.9
	rHP3	6.5	200	<20.0	91.2
	rHP3	6.0	1000	<20.0	>95.0
	rHP3	6.0	200	25.9	93.2
	rHP3	5.5	1000	25.2	93.5
	rHP3	5.5	200	43.6	93.5
3-Rs ² U2	rHP3	5.5	100	53.0	93.7
	rHP3	7.0	200	<20.0	91.9
	rHP3	6.0	200	43.6	93.0
	rHP3	5.5	100	76.1	93.9
3-RU ^L 2	rHP3	7.0	200	<20.0	92.1
	rHP3	6.5	200	24.7	93.4
	rHP3	6.0	200	43.8	93.4
	rHP3	5.5	200	68.2	93.0
	rHP3	5.5	100	74.2	93.0

^aFour hairpin structures (Figure 2) were studied: two RNA duplex targets (rHP1 and rHP3), a mutated RNA duplex target (rHP2), and a DNA duplex target (dHP1). The values shown in parentheses are for sequences with T_{m1} merging with T_{m2} . The absorbance at 265 nm was recorded for 1-Rs²U3 binding to rHP1, rHP2, and dHP1. Experimental errors of T_m values are estimated to be 1 °C.

3.1 Effect of pH on RNA triplex formation

In order to form hydrogen bonds with N7 atoms of guanine in the purine strand of a duplex (Figures 1 and 2), N3 atoms of cytosine in the TFO need to be protonated. Therefore, the effect of pH on triplex formation was investigated. Thermal melting studies for RNA triplexes formed between TFOs and RNA hairpins rHP1 and rHP3 (Figure 2A,G) indicate that the melting temperatures of the RNA hairpins (T_{m2}) do not change obviously with pH (Figure 3A,B, Table 2), which is consistent with preceding results for RNA duplex structures (55). The melting temperatures of the RNA triplexes (T_{m1}) decrease more than 20 °C with increasing pH from pH 5.5 to 6.5 (Figure 3A-C, Table 2), which is consistent with the result that the pK_a of the

N3 atoms of cytosine in TFOs is around 7.0 for the major-groove RNA triplexes

(9).

Figure 3. Effects of pH and NaCl concentration on RNA triplex thermal stability.

(A) UV-absorbance-detected thermal melting curves of TFO 1-R12 binding to the RNA hairpin (rHP1, Figure 2A) at pH 5.5, 6.0 and 6.5 in 200 mM NaCl. (B) First derivatives of thermal melting curves in panel (A) reveal T_{m1} values of 43.6, 28.1, and < 20 °C, respectively, at pH 5.5, 6.0, and pH 6.5. (C) Melting temperatures of triplexes (T_{m1}) at various pH's in 200 mM NaCl. (D) Thermal melting curves of TFO 1-R12 binding to rHP1 at pH 5.5 in varying NaCl concentrations. (E) First derivatives of thermal melting curves in panel (D). With increasing NaCl concentration, triplex melting temperature (T_{m1}) decreases, while hairpin (rHP1) melting temperature (T_{m2}) increases. (F) Melting temperatures of triplexes (T_{m1}) at pH 5.5 in different NaCl concentrations.

3.2 Effect of salt concentration on RNA triplex formation

When the concentration of NaCl is increased from 100 mM to 1 M, the melting temperatures of the RNA hairpins (T_{m2}) increase as we expected; conversely the melting temperatures of the RNA triplexes (T_{m1}) decrease modestly (Figure 3D-F, Table 2). The favorable charge-charge attraction between C^+ and phosphate backbone decreases with increasing salt concentration, resulting in reduced triplex thermal stability. In addition, when the concentration of NaCl is increased, the pK_a of N3 of cytosines in the TFOs decreases. The pK_a decrease may partly explain why triplexes are destabilized with increasing the concentration of NaCl. These results are consistent with the previous results that Na^+ and H^+ are competitive in binding to unmodified DNA triplexes containing both $C^+ \cdot G-C$ and $T \cdot A-T$ base triples (53,56,57).

We observe a stronger salt dependent triplex thermal stability for rHP3 (8 G-C pairs, Figure 2B) than for rHP1 (5 G-C pairs, Figure 2H) (Figure 3F), which could be explained that the formation of $C^+ \cdot G-C$ base triples results in the release of Na^+ . Based on the pH and salt dependence measurements, the chemical modification results shown below can be more easily interpreted

3.3 Incorporation of LNA U and 2-thio U in a TFO enhances its binding to an RNA duplex region

Figure 4. Effects of base and sugar modifications of TFO on RNA triplex thermal stability in 200 mM NaCl at pH 6.0 unless otherwise noted. (A) UV-absorbance-detected thermal melting curves of LNA modified TFOs binding to rHP1 (Figure 2A). First derivatives of thermal melting curves reveal T_{m1} values of 28.1, 36.5, 45.4, and 52.9 °C, respectively, for 1-R12, 1-RU^L1, 1-RU^L2, and 1-RU^L3. (B) 2-Thio U modified TFOs binding to rHP1. The absorption wavelength for 1-Rs²U3 binding to rHP1 was taken at 265 nm. T_{m1} values are 28.1, 40.1, 51.9, and 54.4 °C, respectively, for 1-R12, 1-Rs²U1, 1-Rs²U2, and 1-Rs²U3. (C) 2'-OMe modified TFOs binding to rHP1 in 100 mM NaCl at pH 5.5. T_{m1} values are 46.4 °C, 25.7 °C, and < 20 °C, respectively, for 1-R12, 1-RU^M3, and 1-M12. The results suggest that incorporation of 2'-OMe residues in TFOs destabilizes (RNA)₂-(TFO) triplex formation. (D) An LNA modified TFO binding to rHP3. T_{m1} values are 25.9 °C and 43.8 °C, respectively, for 3-R12 and 3-RU^L2. (E) A 2-thio U modified

TFO binding to rHP3. T_{m1} values are 25.9 °C and 43.6 °C, respectively, for 3-R12 and 3-Rs²U2. (F) 2'-OMe modified TFOs binding to dHP1 in 200 mM NaCl at pH 6.5. T_{m1} values are 29.4, 34.6, and 37.1 °C, respectively, for 1-R12, 1-RU^M3, and 1-M12. The results suggest that incorporation of 2'-OMe residues in TFOs stabilize (DNA)₂-(TFO) triplex formation.

LNA modification (Figure 1B) favors the formation of duplexes with complementary single strand DNA or RNA, and the formation of triplexes with parental DNA duplexes (44-46,49,58,59), because LNA preorganizes oligonucleotides in an A-form-like structure by locking the sugar pucker in C3'-endo conformation. Although C3'-endo sugar pucker in TFOs is favored in both RNA and DNA triplex formation (14,15,60-62), how LNA incorporated in TFOs affects binding to an RNA duplex is not well known. Here we demonstrate by thermal melting experiment that LNA modification in RNA TFOs improves triplex stability with parental RNA duplex structure segment within a hairpin (rHP1, Figure 2A) (Figure 4A, Table 2). Triplexes with more uridines modified are more stable than those with less uridines modified (Figure 4A, Table 2). For the triplex with unmodified RNA TFO (1-R12) in 200 mM NaCl at pH 6.0, the T_{m1} is 28.1 °C. In the same buffer condition, the T_{m1} 's for the triplexes with modified TFOs with one, two and three LNA U residues (1-RU^L1, 1-RU^L2 and 1-RU^L3) are 36.5, 45.4, and 52.9 °C, respectively. Similar stabilizing effect is observed for the complementary TFO (3-RU^L2) binding to rHP3 (Figure 4D). rHP3 is thermally stable and is only partially melted at around 90 °C. This is consistent with

previously reported study (63).

T_{m1} is increased by around 9 °C for each LNA modification (Figure 4A,D). Therefore, it can be concluded that stabilization effect of LNA incorporation is relatively positional and sequence-environmental independent. As we know, LNA modification may disrupt a potential hydrogen bond between 2'-OH in the TFO and a non-bridging oxygen in the purine strand of a duplex (Figure 5) (14,15,60,64,65). The significance of this hydrogen bond is verified by the fact that a DNA TFO (1-D12, Table 1) does not bind to rHP1 (Figure 10A), which is consistent with previous results (63,66). Thus, it can be assumed the energy penalty due to the loss of the hydrogen bonds is compensated by the preorganization of TFO backbone. Thus, our results suggest that LNA may be incorporated into TFOs to enhance their binding to RNA duplex regions.

Previous studies have shown that, 2-thio U modification (Figure 1) in an oligonucleotide stabilizes RNA duplex formation with complementary single strand DNA or RNA, and triplex formation with a parental DNA duplex, respectively (47-49,61). As shown by our thermal melting results, incorporation of 2-thio U in a TFO also significantly stabilizes a triplex with a parental RNA duplex structure segment within a hairpin (rHP1, Figure 2A). At 200 mM NaCl at pH 6.0, the T_{m1} 's for the triplexes with one, two and three 2-thio U residues (1-Rs²U1, 1-Rs²U2 and 1-Rs²U3) in the TFOs are 40.1, 51.9, and 54.4 °C, respectively

(Figure 4B, Table 2). Similar stabilizing effect is also observed for the complementary TFO (3-RS²U2) binding to rHP3 (Figure 4E).

2-Thio U modification in a TFO stabilizes triplex formation, and it is partly caused by the steric repulsion between the 2'-hydroxyl group of the ribose and the bulky 2-thio group of 2-thio U, which favors C3'-endo conformation of the ribose (67), and the C3'-endo conformation facilitates stable triplex formation (60). Base-base hydrogen bonding interaction (between the imino proton H3 in 2-thio U base and N7 in adenine base) is also enhanced with decreased pK_a of N3 (from 9.3 to 8.8) upon thiolation of U (61,68). Moreover, 2-thio U modification enhances TFO binding by reduced thermodynamic cost of dehydration and improved van der Waals contact between sulfur atom in 2-thio U base and H8 hydrogen in adenine base (Figures 1A and 5) (48,69).

Figure 5. Two RNA U·A·U base triples (taken from PDB 3P22) are stabilized by hydrogen bonds between 2'-hydroxyl groups in the third strand and non-bridging oxygen in the purine strands of the RNA duplex (green dashed line). 2'-OMe U modification in the third strand (TFO strand) disrupts this hydrogen bond and the methyl group may cause steric clash with the backbone of the purine strand. LNA U modification in the TFO strand may not cause steric clash due to the fact that the methylene group is highly constrained. The U·A·U base triple is also stabilized by base-base hydrogen bonds (black dashed lines) and van der Waals contacts (yellow dashed lines). Therefore, 2-thio U modification in the TFO strand enhances base-base interactions.

Furthermore, the strong stacking effect of the 2-thio group with adjacent bases facilitates the enhancement of the triplex stability (Figure 5) (47). Base-base

stacking interactions are probably sequence and environmental dependent (55,56,70,71), which may explain one extra 2-thio U modification in 1-Rs²U3, which is flanked by C and U (Table 1), increases T_{m1} by only ~3 °C relative to 1-Rs²U2 (Figure 4B). The 2-thio U modifications present in 1-Rs²U1, 1-Rs²U2, and 3-Rs²U2 are all flanked by two C's, with each modification increasing T_{m1} by ~10 °C (Figure 4B,E). 2-thio U is a naturally occurring modification in RNA (72), which means nature may have evolved to utilize the conservative atomic mutation (2-thio U) to modulate RNA secondary and tertiary structure stability and function.

Figure 6. UV-absorbance-detected thermal melting curves (heating and cooling) of various TFOs binding to the RNA hairpin (rHP1) in 200 mM NaCl at pH 6.0. It reveals that T_{m1} in the heating curve is higher than that in the cooling curve,

indicating a hysteresis for triplex to hairpin transition. T_{m2} is almost the same for heating and cooling curves, indicating no hysteresis for rHP1 to single strand transition.

Due to the hysteresis observed in triplex to hairpin transitions (see Figure 6 for example), we did not obtain equilibrium thermodynamic parameters. Our preliminary Isothermal Titration Calorimetry (ITC) results (data not shown) suggest that the binding dissociation constants are on the order of μM . In addition, we carried out SPR experiments to monitor the triplex formation in real time (Figure 7). As suggested by the SPR results, the designed triplex structures (Figure 2B,H) indeed form. We note that immobilization of hairpins on negatively charged surface may affect binding kinetics and thermodynamics (6,54).

Native polyacrylamide gel electrophoresis experiments (Figure 8) were performed to confirm triplex formation for all TFOs binding to unlabeled hairpins. We post stained the gel by ethidium bromide which intercalates into hairpin and triplex structures. In the lane 2 shown in Figure 8A, rHP1 was mixed with an RNA TFO control sequence (5'-AUCUGUCCACU-3', Table 1), which is not sequence-complementary to the hairpin to form a triplex. As expected, RNA triplex did not form between rHP1 and RNA TFO control. Two bands were observed in lane 6 (shown in Figure 8A) with a molar ratio of rHP1 to TFO (1-RU^L1) at 1:0.5. It suggests that a mixture of ca. 50% triplex and ca. 50% RNA hairpin are present

in lane 6 (Figure 8A). Direct imaging of all pyrimidine TFO bands by staining with Syber Green II was tested, but we observed only hairpin and triplex bands. Consistently, no apparent thermal melting transitions were observed for TFOs alone (data not shown). We did observe a band for an all purine 12-nt single strand (5'-AGAGAGAGAAAG-3', Figure 11C,D) by both ethidium bromide and Syber Green II staining (data not shown). Thus, our thermal melting and gel results suggest that the designed RNA triplex structures (Figure 2B,H) form for unmodified RNA TFO and modified RNA TFOs incorporated with LNA U and 2-thio U residues.

Gel electrophoresis was used to measure the dissociation constant (K_D) for some of the TFOs. TFOs with varying concentrations and 5 nM Cy3 labelled rHP1 were annealed. The RNA triplexes were then incubated overnight at 4 °C before loading into the gels. The results are shown in Figure 9. Quantification of the gels shows that the K_D of rHP1 binding to RNA TFO with LNA modification (1-RU^L1, 257.35 ± 46.94 nM) or 2-thio U modification (1-Rs²U1, 289.94 ± 5.23 nM) is lower than the unmodified RNA TFO (1-R12, 1.27 ± 0.27 μ M), which means both LNA and 2-thio U modifications in TFO enhance triplex formation, consistent with our thermal melting results.

TFO	k_a ($M^{-1}s^{-1}$)	k_d ($\times 10^{-5} s^{-1}$)	K_D ($\times 10^{-9} M$)
3-R12 (n=6)	270 ± 23	2.5 ± 0.3	93 ± 13
3-Rs ² U2	330 ± 2	19.0 ± 5.3	570 ± 162
3-RU ¹ 2 (n=4)	1360 ± 333	0.95 ± 0.01	7.0 ± 6.9

Figure 7. SPR measurements of the interactions between various TFOs and RNA hairpin, rHP3. 82 nM, 247 nM, 741 nM, 2.2 μ M, 6.7 μ M, and 20 μ M of TFOs were serially injected, and the highest response is 20 μ M. The binding dissociation constants, K_D , were calculated based on $K_D = k_d / k_a$. We note that both dissociation and association rates are slow, which makes it difficult to measure kinetics accurately.

Figure 8. Native polyacrylamide gel for the RNA hairpin (rHP1) and triplexes. (A) rHP1 binds to modified RNA TFOs incorporated with LNA U residues. In lane 2, the control TFO (see Table 1) does not bind to rHP1. In lane 6, rHP1 to 1-RU^L1 molar ratio is 1:0.5. The fact that only two bands (rHP1 and triplexes) were observed suggests that the designed RNA triplexes form without alternative structures. (B) rHP1 binds to unmodified and modified RNA TFOs incorporated with 2-thio U residues. (C) RNA triplex bands were not observed for TFOs incorporated with three or more 2'-OMe residues. The results suggest that 2'-OMe modifications in TFOs destabilize RNA triplex formation. (D) rHP3 binds to unmodified and modified TFOs incorporated with 2-thio U and LNA U residues.

Figure 9. Gel electrophoresis for the binding between 5 nM Cy3 labelled rHP1 and RNA TFOs in 200 mM NaCl at pH 5.5. (A) Unmodified RNA TFO 1-R12 (concentrations from lanes 1-13 are 10, 20, 50, 100, 200, 500, 1000, 2000, 5000, 10000, 20000, 50000, and 0 nM, respectively) binding with Cy3 labelled rHP1. (B) LNA modified RNA TFO 1-RU^L1 (concentrations for lanes 1-13 are 1, 2, 5, 10, 20, 50, 100, 200, 500, 1000, 2000, 5000, and 0 nM, respectively) binding with Cy3 labelled rHP1. (C) 2-Thio U modified RNA TFO 1-Rs²U1 (concentrations for lanes 1-13 are 1, 2, 5, 10, 20, 50, 100, 200, 500, 1000, 2000, 5000, and 0 nM, respectively) binding with Cy3 labelled rHP1.

Figure 10. Native polyacrylamide gel for RNA hairpin 1 (rHP1, Fig 2A), mutated RNA hairpin 2 (rHP2, Fig 2C), DNA hairpin 1 (dHP1, Fig 2E), and triplexes. (A)

TFOs 1-D12, 1-MU^L1, 1-MU^L2, and 1-MU^L3 do not form triplexes with rHP1. **(B)** TFOs 1-Ms²U^M1, 1-Ms²U^M2, and 1-Ms²U^M3 do not form triplexes with rHP1. **(C)** TFOs 1-R12, 1-RU^L1, 1-RU^L2, and 1-RU^L3 form triplexes with dHP1. **(D)** TFOs 1-MU^L1, 1-MU^L2, 1-MU^L3, 1-Ms²U^M1, 1-Ms²U^M2, and 1-Ms²U^M3 form triplexes with dHP1. **(E)** TFOs 1-Rs²U1, 1-Rs²U2, 1-Rs²U3, 1-RU^M3 and 1-M12 form triplexes with dHP1. **(F)** TFOs 1-R12, 1-RU^L1, 1-RU^L2, and 1-RU^L3 do not form triplexes with rHP2. **(G)** TFOs 1-RU^M3, 1-M12, 1-Rs²U1, 1-Rs²U2, and 1-Rs²U3 do not form triplexes with rHP2. **(H)** TFOs 1-MU^L1, 1-MU^L2, 1-MU^L3, 1-Ms²U^M1, 1-Ms²U^M2, and 1-Ms²U^M3 do not form triplexes with rHP2.

3.4 Binding specificity assay

We further tested the RNA TFOs binding to an RNA hairpin (rHP2, Figure 2C,D) with one A-U base pair inverted compared to rHP1. As suggested by our thermal melting results in 200 mM NaCl at varying pH, the triplex melting temperatures (T_{m1}) are below 20 °C at pH 6.0 (Table 2). We tested TFOs 1-RU^L1 and 1-Rs²U3 binding to rHP2 at low pH (5.5 and 5.0), and T_{m1} values were found to be more than 30 °C lower than those of sequence-complementary triplexes when T_{m1} and T_{m2} are not overlapped (Tables 2). The native gel results further confirm that triplexes do not form between rHP2 and TFOs with one mismatch (Figure 10F,G,H). The results indicate that the TFOs have good sequence specificity in binding to an RNA duplex region.

3.5 Incorporation of 2'-OMe residues in a TFO destabilizes its binding to an RNA duplex region

We then investigated how 2'-OMe modification (Figure 1B) in a TFO affects the stability of a triplex with a parental RNA duplex region (Figure 2B). The T_{m1} of triplex with TFO 1-RU^M3 (25.7 °C), which has three 2'-OMe residues, is lower than that with RNA TFO (1-R12) (46.4 °C) in 100 mM NaCl at pH 5.5 (Figure 4C and Table 2). All the complementary modified TFOs with more than three 2'-OMe modified residues do not bind to rHP1 (Figure 4C and Table 2). No obvious gel electrophoresis mobility difference was seen between rHP1 alone and rHP1 mixed with TFOs with three or more 2'-OMe residues (Figures 9C and 10A,B), further suggesting that 2'-OMe TFOs do not bind to rHP1. The results indicate that 2'-OMe modified TFOs destabilize triplex formation with RNA duplex region.

It is probable that substitution of 2'-OH with 2'-OMe disrupts the hydrogen bond between 2'-OH in the TFO and a non-bridging oxygen in the purine strand of a duplex (Figure 5) (14,15,60,64,65) and thus disfavors RNA triplex formation (43). In addition, unlike the highly restrained methylene group in LNA, the exposed methyl group in 2'-OMe RNA (Figure 1B) in the TFO may cause steric clash with the narrow major groove of the RNA duplex region.

3.6 TFOs binding to a homologous DNA duplex region

The formation of a thermodynamically stable triplex structure in the major groove without chemical modification is limited to a pyrimidine TFO strand binding to a duplex region with purines on one strand and pyrimidines on the other (62,73-75). Previous thermodynamic studies reveal that to form stable triplexes, RNA is preferred on both pyrimidine strands, whereas DNA is preferred on the purine strand (63). If the purine strand of a parental duplex region is RNA, only RNA TFO strand binds tightly to the parental duplex region (66).

To study the effect of target strand composition on triplex formation, we measured the binding of the complementary TFOs to a DNA hairpin target (dHP1, Figure 2E,F) with the same sequence as the RNA hairpin (rHP1, Figure 2A,B). Our thermal melting results reveal that the triplex melting temperatures (T_{m1}) are always higher with a DNA duplex region than with a homologous RNA duplex region (ΔT_{m1} ranging from 6 to 21 °C in 200 mM NaCl at pH 6.0) (Table 2). The DNA hairpin, dHP1, has a melting temperature (T_{m2}) of about 74.0 °C in 200 mM NaCl, which is about 10 °C lower than the RNA hairpin, rHP1 (Figure 4, Table 2), consistent with the prediction from nearest-neighbor models (70,76,77).

The relatively wider major groove of a DNA duplex (compared to an RNA duplex) may provide relatively easier access for TFO binding. Consistently, incorporation of 2'-OMe residues in TFOs was found to enhance binding to the DNA hairpin,

dHP1 (Figure 4F, Table 2), which is in contrast to the significant destabilization effect observed in TFOs binding to the RNA hairpin, rHP1 (see above, Figure 4C, Table 2). It is likely that the major groove of a DNA duplex is more geometrically compatible for accommodating a 2'-OMe modified TFO. The results suggest that rules for enhancing DNA triplex formation may not be applicable for RNA triplex formation. One may simply incorporate 2'-OMe or other relatively bulky residues in a TFO to selectively target a DNA duplex over a homologous RNA duplex region.

3.7 The hairpin rHP1 forms without alternative structures

The strand concentration independent T_{m2} of rHP1 (Table 3) suggests that the designed hairpin structure forms without appreciable bimolecular duplex (with an internal loop).

Table 3. Melting temperatures and thermodynamic parameters of RNA hairpin (rHP1, Fig. 2A) in 200 mM NaCl at pH 6.0 in varying concentrations^a

Concentration (μ M)	T_m ($^{\circ}$ C)	$-\Delta H^0$ (kcal/mol)	$-\Delta S^0$ (cal/mol $\cdot$ K)	$-\Delta G^0_{37}$ (kcal/mol)
54.3	83.6	116.2	325.7	15.16
31.8	83.8	118.4	331.7	15.54
18.2	83.2	117.2	328.8	15.19
11.1	83.6	117.1	328.1	15.29
7.1	83.9	126.1	353.3	16.55
4.6	84.1	126.4	353.8	16.67
2.3	83.6	116.9	327.7	15.26
1.5	83.1	121.0	339.7	15.65
1.0	82.9	122.5	344.2	15.78

^a T_m values and thermodynamic parameters were obtained by fitting to a two-state model with the MeltWin program (<http://www.meltwin3.com>) (41). Experimental errors of T_m values are estimated as 1 $^{\circ}$ C. Experimental errors of ΔH^0 , ΔS^0 , ΔG^0_{37} (free energy at 37 $^{\circ}$ C) are estimated to as 12%, 13.5%, and 4%, respectively (59).

It is likely that chemically modified TFOs may displace the pyrimidine segment of the parental RNA hairpin to form a parallel or antiparallel duplex structure as shown in Figure 11A,B. To test the possibility of the formation of such alternative structures due to strand invasion, we carried out thermal melting studies for the duplexes formed between a 12-nt all purine strand (5'-AGAGAGAGAAAG-3', Figure 11C,D) and various TFOs. The results (Table 4) reveal that most of the duplex formation is relatively pH independent (ΔT_m is within 3 $^{\circ}$ C from pH 7.0 to

5.5) and have lower T_m than that of the parental RNA hairpin, rHP1 (Figure 12 and Tables 2 and 4). The pH independence of melting temperatures suggests that 12-bp parallel RNA duplex structures (formed between TFO sequences and an all purine strand, Figure 11D) do not form, because the stability of parallel DNA duplex is pH dependent (78). Therefore, 9-bp antiparallel duplex structures form between the 12-nt all purine strand and TFOs (see Figure 11C for example).

Table 4. Melting temperatures and thermodynamic parameters of duplexes formed between various TFO strands (5 μ M) and a purine strand (5'-AGAGAGAGAAAG-3') (5 μ M) in 200 mM NaCl at varying pH^a

TFO sequence	pH	T_m (°C)	$-\Delta H^0$ (kcal/mol)	$-\Delta S^0$ (cal/mol·K)	$-\Delta G^0_{37}$ (kcal/mol)
1-R12	7.0	54.0 (54.3)	89.6	247.8	12.69
	6.0	54.3 (55.0)	93.7	259.6	13.13
	5.5	51.9 (53.0)	87.3	241.8	12.30
1-RU ^L 1	7.0	59.3 (59.4)	86.4	234.0	13.83
	6.0	59.9 (60.3)	87.3	236.2	14.04
	5.5	57.6 (58.3)	84.9	230.3	13.47
1-RU ^L 2	7.0	65.7 (66.4)	93.2	248.9	16.01
	6.0	67.9 (68.0)	96.5	258.0	16.51
	5.5	63.2 (65.1)	96.2	258.4	16.02
1-RU ^L 3	7.0	68.3 (70.1)	108.1	289.0	18.37
	6.0	72.0 (72.7)	104.7	278.1	18.48
	5.5	67.2 (69.3)	111.5	299.6	18.52
1-Rs ² U1	7.0	57.6 (58.6)	82.7	223.9	13.27
	6.0	57.4 (58.2)	79.1	213.1	13.05
	5.5	56.4 (57.0)	80.4	217.7	12.88
1-Rs ² U2	7.0	62.2 (63.5)	76.1	200.2	14.01
	6.0	61.8 (63.7)	78.4	206.9	14.21
	5.5	60.7 (62.2)	79.3	210.6	14.00
1-Ms ² U ^M 1	6.0	62.1 (63.4)	88.3	236.6	14.92
1-Ms ² U ^M 2	6.0	65.3 (67.1)	118.2	321.3	18.55
1-Ms ² U ^M 3	6.0	67.9 (68.9)	133.5	364.0	20.55
1-MU ^L 1	6.0	67.8 (68.4)	111.9	320.7	18.03
1-MU ^L 2	6.0	73.1 (73.5)	125.2	335.7	21.08
1-MU ^L 3	6.0	75.5 (76.5)	135.2	360.1	23.17
rHP1	6.5	84.1 (84.7)	115.3	322.2	15.35
	6.0	84.0 (84.5)	114.6	320.3	15.22
	5.5	83.5 (84.1)	110.4	309.1	14.55
dHP1	6.5	73.7 (73.9)	89.5	257.8	9.51
	6.0	73.5 (73.7)	92.2	265.8	9.77
	5.5	74.4 (74.6)	88.7	255.1	9.59
rHP2	6.0	83.8 (84.5)	114.1	319.1	15.15
	5.5	83.8 (84.7)	117.1	327.1	15.61

^a T_m values shown in parentheses and thermodynamic parameters for duplex formation were obtained by fitting to a two-state model with the MeltWin program (<http://www.meltwin3.com>) (41). Thermal melting results for RS3 are not listed here due to the fact that the thermal melting transitions are non-two-state (see Fig. S6). The bottom three sequences are for RNA, DNA, and mutated RNA hairpins (see Fig. 2). Experimental errors of T_m values are estimated as 1 °C. Experimental errors of ΔH^0 ,

ΔS^0 , ΔG^0_{37} (free energy at 37 °C) are estimated to as 12%, 13.5%, and 4%, respectively (59).

Figure 11. Hypothetical structures formed involving TFO strands. A representative TFO strand is shown in gray with the residues to be modified shown in black. Note that our results suggest that structures shown in panels (A, B, and D) do not form. **(A)** Parallel duplex structure formed by a modified TFO binding to the RNA hairpin sequence, rHP1 (Fig. 2A). **(B)** Antiparallel duplex structure formed by a TFO binding to rHP1. **(C)** 9-bp antiparallel duplex structure formed between a 12-nt purine strand and a TFO. The 12-nt purine strand has the same sequence as the purine segment of rHP1. **(D)** 12-bp parallel duplex structure formed between the 12-nt purine strand and a TFO.

The fact that most of the T_m values of the 9-bp antiparallel duplexes (Figure 11C) are in between T_{m1} and T_{m2} of triplexes (Figure 12, Tables 2 and 4) suggests that neither parallel nor antiparallel duplex structure due to strand invasion (Figure

11A,B) forms. In addition, there are only two bands observed in native gels (Figures 8-10) further suggesting no appreciable alternative structures form (74). Thus, our designed 12-base-triple triplex structures (Figure 2B) form between rHP1 and complementary unmodified and modified RNA TFOs incorporated with LNA U and 2-thio U residues.

Figure 12. Comparison of melting temperatures of triplexes (Figure 2B, T_{m1}), RNA hairpin rHP1 (Figure 2A, T_{m2}), and 9-bp duplexes (Figure 11C, T_m) in 200 mM NaCl at varying pH.

4. CONCLUSION

In this part, both base and sugar modifications in TFOs have been found to regulate the triplex formation with RNA duplex regions. Both sugar methylation and base thiolation are naturally occurring RNA modifications. The effects of sugar modifications are probably sequence environment independent. 2'-OMe sugar modification may be selected by nature to weaken RNA base triple interaction due to the loss of a hydrogen bond and steric clash. LNA sugar modification in TFOs stabilizes RNA triplex structures through backbone preorganization, despite the loss of a hydrogen bond. 2-Thio U base modification in TFOs stabilizes RNA triplex formation in a sequence environment dependent manner, through the combined effects of enhanced van der Waals contact, hydrogen bonding, base stacking, backbone preorganization, and reduced free energy penalty of dehydration. A thio group may be selected by nature to enhance van der Waals interaction with hydrogen atoms in a duplex major groove (H8 in adenine and guanine, H5 and H6 in uracil and cytosine) and other interactions, and therefore favor RNA base triple and other tertiary structures. Our results provide useful insights into rational design of more potent and selective triplex forming ligands targeting biologically important RNA duplex regions.

5. REFERENCES

1. Kole R, Krainer AR, Altman S (2012) RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat. Rev. Drug Discov.* 11: 125-140.
2. Wesolowski D, Tae HS, Gandotra N, Llopis P, Shen N, Altman S (2011) Basic peptide-morpholino oligomer conjugate that is very effective in killing bacteria by gene-specific and nonspecific modes. *Proc. Natl. Acad. Sci. U.S.A.* 108: 16582-16587.
3. Guan LR, Disney MD (2012) Recent advances in developing small molecules targeting RNA. *ACS Chem. Biol.* 7: 73-86.
4. Akinc A, Zumbuehl A, Goldberg M, Leshchiner ES, Busini V, Hossain N, Bacallado SA, Nguyen DN, Fuller J, Alvarez R, Borodovsky A, Borland T, Constien R, de Fougerolles A, Dorkin JR, Jayaprakash KN, Jayaraman M, John M, Kotliansky V, Manoharan M, Nechev L, Qin J, Racie T, Raitcheva D, Rajeev KG, Sah DWY, Soutschek J, Toudjarska I, Vornlocher HP, Zimmermann TS, Langer R, Anderson DG (2008) A combinatorial library of lipid-like materials for delivery of RNAi therapeutics. *Nat. Biotechnol.* 26: 561-569.
5. Barczak AK, Gomez JE, Kaufmann BB, Hinson ER, Cosimi L, Borowsky ML, Onderdonk AB, Stanley SA, Kaur D, Bryant KF, Knipe DM, Sloutsky A, Hung DT (2012) RNA signatures allow rapid identification of pathogens and antibiotic susceptibilities. *Proc. Natl. Acad. Sci. U.S.A.* 109: 6217-6222.
6. Zhang X, Servos MR, Liu J (2012) Instantaneous and quantitative functionalization of gold nanoparticles with thiolated DNA using a pH-assisted and surfactant-free route. *J. Am. Chem. Soc.* 134: 7266-7269.
7. Jaeger L, Chworos A (2006) The architectonics of programmable RNA and DNA nanostructures. *Curr. Opin. Struct. Biol.* 16: 531-543.
8. Chen G, Chang KY, Chou MY, Bustamante C, Tinoco I, Jr. (2009) Triplex structures in an RNA pseudoknot enhance mechanical stability and increase efficiency of -1 ribosomal frameshifting. *Proc. Natl. Acad. Sci. U.S.A.* 106: 12706-12711.
9. Theimer CA, Blois CA, Feigon J (2005) Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Mol. Cell* 17: 671-682.
10. Kim NK, Zhang Q, Zhou J, Theimer CA, Peterson RD, Feigon J (2008) Solution structure and dynamics of the wild-type pseudoknot of human telomerase RNA. *J. Mol. Biol.* 384: 1249-1261.
11. Gilbert SD, Rambo RP, Van Tyne D, Batey RT (2008) Structure of the SAM-II riboswitch bound to S-adenosylmethionine. *Nat. Struct. Mol. Biol.* 15: 177-182.
12. Qiao F, Cech TR (2008) Triple-helix structure in telomerase RNA contributes to catalysis. *Nat. Struct. Mol. Biol.* 15: 634-640.
13. Liu F, Theimer CA (2012) Telomerase Activity Is Sensitive to Subtle Perturbations of the TLC1 Pseudoknot 3' Stem and Tertiary Structure. *J. Mol. Biol.* 423: 719-735.
14. Mitton-Fry RM, DeGregorio SJ, Wang J, Steitz TA, Steitz JA (2010) Poly(A) tail recognition by a viral RNA element through assembly of a triple helix. *Science* 330: 1244-1247.

15. Ulyanov NB, James TL (2010) RNA structural motifs that entail hydrogen bonds involving sugar-phosphate backbone atoms of RNA. *New J. Chem.* 34: 910-917.
16. Leontis NB, Westhof E (2003) Analysis of RNA motifs. *Curr. Opin. Struc. Biol.* 13: 300-308.
17. Ogle JM, Brodersen DE, Clemons WM, Jr., Tarry MJ, Carter AP, Ramakrishnan V (2001) Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science* 292: 897-902.
18. Demeshkina N, Jenner L, Westhof E, Yusupov M, Yusupova G (2012) A new understanding of the decoding principle on the ribosome. *Nature* 484: 256-259.
19. Puglisi EV, Puglisi JD (2011) Secondary structure of the HIV reverse transcription initiation complex by NMR. *J. Mol. Biol.* 410: 863-874.
20. Levengood JD, Rollins C, Mishler CH, Johnson CA, Miner G, Rajan P, Znosko BM, Tolbert BS (2012) Solution structure of the HIV-1 exon splicing silencer 3. *J. Mol. Biol.* 415: 680-698.
21. Runyon ST, Puglisi JD (2003) Design of a cyclic peptide that targets a viral RNA. *J. Am. Chem. Soc.* 125: 15704-15705.
22. Battiste JL, Mao HY, Rao NS, Tan RY, Muhandiram DR, Kay LE, Frankel AD, Williamson JR (1996) a helix-RNA major groove recognition in an HIV-1 Rev peptide RRE RNA complex. *Science* 273: 1547-1551.
23. Wiedenheft B, Sternberg SH, Doudna JA (2012) RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482: 331-338.
24. Nallagatla SR, Toroney R, Bevilacqua PC (2011) Regulation of innate immunity through RNA structure and the protein kinase PKR. *Curr. Opin. Struc. Biol.* 21: 119-127.
25. Davidson A, Leeper TC, Athanassiou Z, Patora-Komisarska K, Karn J, Robinson JA, Varani G (2009) Simultaneous recognition of HIV-1 TAR RNA bulge and loop sequences by cyclic peptide mimics of Tat protein. *Proc. Natl. Acad. Sci. U.S.A.* 106: 11931-11936.
26. Abu Almakarem AS, Petrov AI, Stombaugh J, Zirbel CL, Leontis NB (2012) Comprehensive survey and geometric classification of base triples in RNA structures. *Nucleic Acids Res.* 40: 1407-1423.
27. Kim SH, Suddath FL, Quigley GJ, McPherson A, Sussman JL, Wang AH, Seeman NC, Rich A (1974) Three-dimensional tertiary structure of yeast phenylalanine transfer RNA. *Science* 185: 435-440.
28. Robertus JD, Ladner JE, Finch JT, Rhodes D, Brown RS, Clark BF, Klug A (1974) Structure of yeast phenylalanine tRNA at 3 Å resolution. *Nature* 250: 546-551.
29. Quigley GJ, Rich A (1976) Structural domains of transfer RNA molecules. *Science* 194: 796-806.
30. Stout CD, Mizuno H, Rubin J, Brennan T, Rao ST, Sundaralingam M (1976) Atomic coordinates and molecular conformation of yeast phenylalanyl tRNA. An independent investigation. *Nucleic Acids Res.* 3: 1111-1123.
31. Puglisi JD, Tan R, Calnan BJ, Frankel AD, Williamson JR (1992) Conformation of the TAR RNA-arginine complex by NMR spectroscopy. *Science* 257: 76-80.
32. Puglisi JD, Chen L, Blanchard S, Frankel AD (1995) Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex. *Science* 270: 1200-1203.
33. Ye X, Kumar RA, Patel DJ (1995) Molecular recognition in the bovine immunodeficiency virus Tat peptide-TAR RNA complex. *Chem. Biol.* 2: 827-840.

34. Calabro V, Daugherty MD, Frankel AD (2005) A single intermolecular contact mediates intramolecular stabilization of both RNA and protein. *Proc. Natl. Acad. Sci. U.S.A* 102: 6849-6854.
35. Davidson A, Leeper TC, Athanassiou Z, Patora-Komisarska K, Karn J, Robinson JA, Varani G (2009) Simultaneous recognition of HIV-1 TAR RNA bulge and loop sequences by cyclic peptide mimics of Tat protein. *Proc. Natl. Acad. Sci. U.S.A* 106: 11931-11936.
36. Batey RT (2012) Structure and mechanism of purine-binding riboswitches. *Q. Rev. Biophys.* 45: 345-381.
37. Moser HE, Dervan PB (1987) Sequence-specific cleavage of double helical DNA by triple helix formation. *Science* 238: 645-650.
38. Le Doan T, Perrouault L, Praseuth D, Habhouh N, Decout JL, Thuong NT, Lhomme J, Helene C (1987) Sequence-specific recognition, photocrosslinking and cleavage of the DNA double helix by an oligo-[a]-thymidylate covalently linked to an azidoproflavine derivative. *Nucleic Acids Res.* 15: 7749-7760.
39. Morgan AR, Wells RD (1968) Specificity of the three-stranded complex formation between double-stranded DNA and single-stranded RNA containing repeating nucleotide sequences. *J. Mol. Biol.* 37: 63-80.
40. Duca M, Vekhoff P, Oussedik K, Halby L, Arimondo PB (2008) The triple helix: 50 years later, the outcome. *Nucleic Acids Res.* 36: 5123-5138.
41. Fox KR, Brown T (2011) Formation of stable DNA triplexes. *Biochem. Soc. Trans.* 39: 629-634.
42. Shimizu M, Konishi A, Shimada Y, Inoue H, Ohtsuka E (1992) Oligo(2'-O-methyl)ribonucleotides: effective probes for duplex DNA. *FEBS Lett.* 302: 155-158.
43. Wang S, Kool ET (1995) Relative stabilities of triple helices composed of combinations of DNA, RNA and 2'-O-methyl-RNA backbones: chimeric circular oligonucleotides as probes. *Nucleic Acids Res.* 23: 1157-1164.
44. Torigoe H, Hari Y, Sekiguchi M, Obika S, Imanishi T (2001) 2'-O,4'-C-methylene bridged nucleic acid modification promotes pyrimidine motif triplex DNA formation at physiological pH: Thermodynamic and kinetic studies. *J. Biol. Chem.* 276: 2354-2360.
45. Brunet E, Alberti P, Perrouault L, Babu R, Wengel J, Giovannangeli C (2005) Exploring cellular activity of locked nucleic acid-modified triplex-forming oligonucleotides and defining its molecular basis. *J. Biol. Chem.* 280: 20076-20085.
46. Vester B, Wengel J (2004) LNA (Locked nucleic acid): High-affinity targeting of complementary RNA and DNA. *Biochemistry* 43: 13233-13241.
47. Okamoto I, Seio K, Sekine M (2006) Triplex forming ability of oligonucleotides containing 2'-O-methyl-2-thiouridine or 2-thiothymidine. *Bioorg. Med. Chem. Lett.* 16: 3334-3336.
48. Testa SM, Disney MD, Turner DH, Kierzek R (1999) Thermodynamics of RNA-RNA duplexes with 2-or 4-thiouridines: Implications for antisense design and targeting a group I intron. *Biochemistry* 38: 16655-16662.
49. Carlucci M, Kierzek E, Olejnik A, Turner DH, Kierzek R (2009) Chemical synthesis of LNA-2-thiouridine and its influence on stability and selectivity of oligonucleotide binding to RNA. *Biochemistry* 48: 10882-10893.

50. Kierzek E, Mathews DH, Ciesielska A, Turner DH, Kierzek R (2006) Nearest neighbor parameters for Watson-Crick complementary heteroduplexes formed between 2'-O-methyl RNA and RNA oligonucleotides. *Nucleic Acids Res.* 34: 3609-3614.
51. Moody EM, Lecomte JTJ, Bevilacqua PC (2005) Linkage between proton binding and folding in RNA: A thermodynamic framework and its experimental application for investigating pK_a shifting. *RNA* 11: 157-172.
52. McDowell JA, Turner DH (1996) Investigation of the structural basis for thermodynamic stabilities of tandem GU mismatches: solution structure of (rGAGGUCUC)₂ by two-dimensional NMR and simulated annealing. *Biochemistry* 35: 14077-14089.
53. Sugimoto N, Wu P, Hara H, Kawamoto Y (2001) pH and cation effects on the properties of parallel pyrimidine motif DNA triplexes. *Biochemistry* 40: 9396-9405.
54. Alberti P, Arimondo PB, Mergny JL, Garestier T, Helene C, Sun JS (2002) A directional nucleation-zipping mechanism for triple helix formation. *Nucleic Acids Res.* 30: 5407-5415.
55. Chen G, Kennedy SD, Turner DH (2009) A CA⁺ pair adjacent to a sheared GA or AA pair stabilizes size-symmetric RNA internal loops. *Biochemistry* 48: 5738-5752.
56. Soto AM, Loo J, Marky LA (2002) Energetic contributions for the formation of TAT/TAT, TAT/CGC⁺, and CGC⁺/CGC⁺ base triplet stacks. *J. Am. Chem. Soc.* 124: 14355-14363.
57. Lee HT, Arciniegas S, Marky LA (2008) Unfolding thermodynamics of DNA pyrimidine triplexes with different molecularities. *J. Phys. Chem. B* 112: 4833-4840.
58. Kierzek E, Ciesielska A, Pasternak K, Mathews DH, Turner DH, Kierzek R (2005) The influence of locked nucleic acid residues on the thermodynamic properties of 2'-O-methyl RNA/RNA heteroduplexes. *Nucleic Acids Res.* 33: 5082-5093.
59. Kierzek E, Pasternak A, Pasternak K, Gdaniec Z, Yildirim I, Turner DH, Kierzek R (2009) Contributions of stacking, preorganization, and hydrogen bonding to the thermodynamic stability of duplexes between RNA and 2'-O-methyl RNA with locked nucleic acids. *Biochemistry* 48: 4377-4387.
60. Holland JA, Hoffman DW (1996) Structural features and stability of an RNA triple helix in solution. *Nucleic Acids Res.* 24: 2841-2848.
61. Kumar RK, Davis DR (1997) Synthesis and studies on the effect of 2-thiouridine and 4-thiouridine on sugar conformation and RNA duplex stability. *Nucleic Acids Res.* 25: 1272-1280.
62. Gotfredsen CH, Schultze P, Feigon J (1998) Solution structure of an intramolecular pyrimidine-purine-pyrimidine triplex containing an RNA third strand. *J. Am. Chem. Soc.* 120: 4281-4289.
63. Roberts RW, Crothers DM (1992) Stability and properties of double and triple helices: dramatic effects of RNA or DNA backbone composition. *Science* 258: 1463-1466.
64. Escude C, Francois JC, Sun JS, Ott G, Sprinzl M, Garestier T, Helene C (1993) Stability of triple helices containing RNA and DNA strands: experimental and molecular modeling studies. *Nucleic Acids Res.* 21: 5547-5553.
65. Wilds CJ, Damha MJ (1999) Duplex recognition by oligonucleotides containing 2'-deoxy-2'-fluoro-D-arabinose and 2'-deoxy-2'-fluoro-D-ribose. Intermolecular 2'-OH-phosphate contacts versus sugar puckering in the stabilization of triple-helical complexes. *Bioconjugate Chem.* 10: 299-305.

66. Han H, Dervan PB (1993) Sequence-specific recognition of double helical RNA and RNA:DNA by triple helix formation. *Proc. Natl. Acad. Sci. U.S.A.* 90: 3806-3810.
67. Yamamoto Y, Yokoyama S, Miyazawa T, Watanabe K, Higuchi S (1983) NMR analyses on the molecular mechanism of the conformational rigidity of 2-thioribothymidine, a modified nucleoside in extreme thermophile transfer-RNAs. *FEBS Lett.* 157: 95-99.
68. Gilli P, Pretto L, Bertolasi V, Gilli G (2009) Predicting hydrogen-bond strengths from acid-base molecular properties. The pK_a slide rule: toward the solution of a long-lasting problem. *Acc. Chem. Res.* 42: 33-44.
69. Siegfried NA, Kierzek R, Bevilacqua PC (2010) Role of unsatisfied hydrogen bond acceptors in RNA energetics and specificity. *J. Am. Chem. Soc.* 132: 5342-5344.
70. Xia T, SantaLucia J, Jr., Burkard ME, Kierzek R, Schroeder SJ, Jiao X, Cox C, Turner DH (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry* 37: 14719-14735.
71. Chen G, Kierzek R, Yildirim I, Krugh TR, Turner DH, Kennedy SD (2007) Stacking effects on local structure in RNA: changes in the structure of tandem GA pairs when flanking GC pairs are replaced by isoG-isoC pairs. *J. Phys. Chem. B* 111: 6718-6727.
72. Cantara WA, Crain PF, Rozenski J, McCloskey JA, Harris KA, Zhang X, Vendeix FA, Fabris D, Agris PF (2011) The RNA Modification Database, RNAMDB: 2011 update. *Nucleic Acids Res.* 39: D195-201.
73. Plum GE, Pilch DS, Singleton SF, Breslauer KJ (1995) Nucleic acid hybridization: triplex stability and energetics. *Annu. Rev. Biophys. Biomol. Struct.* 24: 319-350.
74. Blose JM, Silverman SK, Bevilacqua PC (2007) A simple molecular model for thermophilic adaptation of functional nucleic acids. *Biochemistry* 46: 4232-4240.
75. Radhakrishnan I, Patel DJ (1994) DNA Triplexes: Solution Structures, Hydration Sites, Energetics, Interactions, and Function. *Biochemistry* 33: 11405-11416.
76. Mathews DH, Disney MD, Childs JL, Schroeder SJ, Zuker M, Turner DH (2004) Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc. Natl. Acad. Sci. U.S.A.* 101: 7287-7292.
77. SantaLucia J, Jr., Hicks D (2004) The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.* 33: 415-440.
78. Ohmichi T, Kawamoto Y, Wu P, Miyoshi D, Karimata H, Sugimoto N (2005) DNA-based biosensor for monitoring pH in vitro and in living cells. *Biochemistry* 44: 7125-7130.

Part II: Allostery within a Watson-Crick RNA Duplex and Its Application in Studying RNA Recognition by RIG-I

1. INTRODUCTION

RNA has rich local and global conformational dynamics due to the relatively flexible non-Watson-Crick regions and relatively weak tertiary contacts compared to Watson-Crick duplexes (1-14). Watson-Crick RNA duplexes are usually considered as rigid scaffolds in large complex RNAs. RNA-binding proteins and other ligands often induce changes in the local structures of non-Watson-Crick regions of the binding site. The local structure changes can be transmitted through relatively rigid Watson-Crick RNA duplexes facilitating modulation of RNA global structures and dynamics (15-18). Inversely, ligand binding and/or formation of long-range tertiary structures may cause allosteric effect resulting in structural rearrangement in the catalytic site of a ribozyme or the metabolite binding site of a riboswitch (19-27).

Allosteric effect within a Watson-Crick RNA duplex is not well known. Here, we report the direct observation that the local structural dynamics of a Watson-Crick RNA duplex is modulated allosterically by the specific binding of a ligand. The RNA duplex-binding ligands studied in this paper are human Retinoic acid-inducible gene I (RIG-I), which is a newly discovered cytosolic pattern-recognition receptor that senses viral RNAs released from replicating

viruses in infected cells (28). Its structure (Figure 1) can be divided into three parts. At the N-terminus, there are two caspase activation and recruitment domains (CARDs). Upon binding to an RNA ligand, CARDs are exposed and interact with mitochondrial associated antiviral signaling molecule (MAVS) (29), which in turn activates the downstream signaling processes. The central DExD/H helicase region assists in the detection of viral RNAs. The C-terminal domain (CTD) recognizes and interacts with various RNAs (30,31).

Figure 1. Illustration of the RIG-I domain structure.

The crystal structures of RIG-I CTD in complex with a short dsRNA reveal that RIG-I CTD binds to the blunt terminal end of 5'-ppp dsRNA (see Figure 2B for example) (32,33). RIG-I CTD is in contact with the backbone in both major and minor grooves of the first two base pairs, but reaches into the major groove starting from the third base pair (32-35). Close inspection of the crystal structures reveals that multiple lysine residues surround the 5'-ppp group. The only nucleobase specific interaction is the stacking between a phenylalanine residue of RIG-I and the terminal base pair of RNA. Compared to 5'-phosphorylated RNA, non-phosphorylated dsRNA binds to RIG-I CTD slightly differently. The different amino acid residues, and molecular interactions (charge-charge ionic interactions, hydrogen bonds, and aromatic stacking) involved are shown in Figure 3.

Recent crystal structures of the complexes of multidomain RIG-I and dsRNA (36-38) suggest that RIG-I helicase and CTD domains bind synergistically to a dsRNA with CTD partially capping one terminal end of the dsRNA and helicase domain surrounding the adjacent duplex region (Figure 2A,C). However, it is not clear how RIG-I activity is modulated by RNA ligands with non-Watson-Crick structures and chemical modifications (33,39-41).

Figure 2. (A) Cartoons of dsRNA in complex with RIG-I CTD, RIG-I helicase-CTD, and full length RIG-I, respectively. Drawing is not to scale and is only for illustrative purpose. (B) Crystal structure (PDB: 3NCU) of the RIG-I

C-terminal domain (CTD) (shown in blue) in complex with a 12-bp dsRNA with 5'-pp. RIG-I CTDs are bound to both ends of the dsRNA. The terminal G-C base pairs (shown in cyan lines) of the dsRNA stack on phenylalanine 853 (shown in orange lines) of RIG-I. (C) Crystal structure (PDB: 2YKG) of the two-domain RIG-I helicase-CTD in complex with a 10-bp dsRNA with 5'-OH. RIG-I CTD is bound to one terminal end of the dsRNA, with the helicase domain wrapping around the dsRNA. A terminal G-C base pair (shown in cyan lines) of the dsRNA stacks on phenylalanine 853 (shown in orange lines) of RIG-I.

Figure 3. (A) Binding interactions between RIG-I CTD and dsRNA without 5'-ppp (34). (B) Binding interactions between RIG-I CTD and 5'-ppp dsRNA (33).

2-Aminopurine (2AP), which forms a Watson-Crick-like base pair with uracil (U)

(Figure 4), is a fluorescent isomorphous nucleobase analogue of adenine (A, 6-aminopurine) with absorption and emission wavelengths at 303 and 370 nm, respectively (42-45). When 2AP is incorporated into DNA, its fluorescence will be quenched (46-48), and the quantum yield will be reduced from that of the free 2AP (0.65). This reduction is caused by nucleobases stacking interaction. The 2AP fluorescence decay is single exponential with a lifetime of ~10 ns; but when the 2AP is incorporated into a DNA duplex, the lifetime has four decay components, from 50 ps to 8 ns (46). The short decay time of 2AP is attributed to the fully stacked state (49). The stacked states can be affected by temperature, flanking bases, and protein binding. Thus, 2AP is a very useful probe for the detection of local conformational and environmental changes due to complex interactions.

There is a great potential in tuning anti-viral innate immune response by novel short nucleic acid-based agonists/antagonists targeting RIG-I. In this thesis, we aim to develop a 2AP fluorescence based method to study RNA binding with RIG-I. We have substituted Watson-Crick pair (A-U or G-C) with (2AP)-U pair (Figure 5) in various RNA duplexes or hairpins to probe their binding to RIG-I CTD and full length RIG-I using steady state and time-resolved fluorescence assays. Binding was confirmed by gel electrophoresis. Cy3 was incorporated in the loop of an RNA hairpin to quantify the binding affinity with RIG-I CTD by gel electrophoresis.

Figure 4. Chemical structures of Watson-Crick G-C, Watson-Crick A-U, and Watson-Crick-like (2AP)-U pairs. Fluorescent 2AP (2-aminopurine) is an isomer of A (6-aminopurine).

A	B	C	D	E	F	G	H	I	J	K	L
										U C	
									U C	U G	
									U G	C-G	
									C-G	C-G	A A
									U-2	U-2	A G
									13G-C	13G-C	25C-G 20
									A-U	A-U	C-G
									A-U	A-U	2-U
									A-U	A-U	C-G
									G-C	G-C	G-C
									A-U	A-U	3'A-U 15
									G-C	G-C	5'G-C
									A-U	A-U	C-G
									A-U	A-U	G-C
									A-U	A-U	5G-C 10
									A-U	A-U	G A
									A-U	A-U	A A
5' 3'	5' 3'	5' 3'	5' 3'	5' 3'	5' 3'	5' 3'	5' 3'	5' 3'	5' 3'	5' 3'	5' 3'
12-bp RNA1	12-bp RNA2	rHP1-Cy3	rHP1-5	rHP2-9	rHP2-9-AA	rHP3-11	rHP4-12	rHP5-13	rHP6-13	rHP7-14	dumbbell RNA

Figure 5. RNA hairpins and duplexes labelled with 2AP or Cy3. The number “2” represents 2AP. (A,B) self-complementary 12-bp RNA1 (with 2AP at 5'-termini) and 12-bp RNA2 (with 2AP at internal positions). (C,D) 8-bp 20-nt RNA hairpins labelled with Cy3 (rHP1-Cy3) and 2AP (rHP1-5). (E,F) 12-bp 28-nt RNA hairpin

in the absence (rHP2-9) and presence (rHP2-9-AA) of an A-A mismatch. **(G)** A 14-bp 32-nt RNA hairpin (rHP3-11). **(H)** A 15-bp 34-nt RNA hairpin (rHP4-12). **(I)** A 14-bp 32-nt RNA hairpin (rHP5-13). **(J)** A 16-bp 36-nt RNA hairpin (rHP6-13). **(K)** A 17-bp 38-nt RNA hairpin (rHP7-14). **(L)** An 11-bp 30-nt dumbbell RNA structure (dumbbell RNA).

2. MATERIALS AND METHODS

2.1 Protein expression and purification

400 mL Luria broth (LB) for *Escherichia coli* bacteria culture medium was prepared in a 1 L conical flask by adding 4 g peptone, 4 g NaCl and 2 g yeast extract and 400 mL of ultrapure water. The flask was autoclaved at 121 °C for 20 minutes before the addition of 200 µL of 34 mg/ mL chloramphenicol and 800 µL of 50 mg/ mL ampicillin.

10 mL of autoclaved LB solution was added into six 15 mL plastic centrifuge tubes. Human RIG-I CTD is on vector pET22b. Rosetta/pET22b/RIG-I CTD bacteria were spread on a solid agar gel plate with 17 µg/mL chloramphenicol and 100 µg/mL ampicillin, and we placed the gel plate at 37 °C incubator inversely for 12 hours. A single colony of *Escherichia coli* bacteria was picked from the agar plates and transferred to the 15 mL tubes. The tubes were then incubated at 37 °C with shaking at 200 rpm for 6 hours until the optical density (OD) at 600 nm is

approximately 0.6-0.8.

The culture with 10 mL LB solution was transferred into the 1 L conical flask with around 400 mL LB solution. Then, it was incubated at 37 °C with shaking at 200 rpm for 4 hours until the OD at 600 nm is approximately 0.8. 160 µL of 1 M of IPTG (isopropylthiogalactoside) was added into the flask, and it was incubated at 16 °C with shaking at 200 rpm for 12 hours in order to induce the expression of the RIG-I CTD.

200 mL bacterial culture was evenly divided into four 50 mL tubes and centrifuged at 8000 rpm at 25 °C for 5 minutes. We removed the supernatant and collected the sediment. 25 mL lysis buffer with 20 mM NaH₂PO₄, 0.5 M NaCl, 40 mM imidazole, 5% glycerol, and 5 mM of dithiothreitol (DTT) at pH 8.0 was added to the tubes in order to suspend the sediment. Then the sediment was sonicated for 15 minutes and centrifuged at 8000 rpm for 1 hour at 4 °C. After that, the supernatant was collected. Purification was performed on HisTrap column via Fast Performance Liquid Chromatography (FPLC). The RIG-I CTD are tagged with six consecutive histidine residues (His-tag), which interacts with the nickel ions in the HisTrap column. The column was equilibrated with washing buffer containing 50 mM sodium phosphate, 300 mM NaCl, 10% glycerol, 40 mM imidazole, and 1 mM DTT at pH 8.0. The supernatant was then injected into the FPLC, and the proteins without His-tag were eluted out first. Then, the elution buffer, consisting

of the same components as washing buffer but 500 mM imidazole, was used to elute out the RIG-I CTD with His-tag. The RIG-I CTD was purified by Superdex 75 size exclusion column in FPLC, using a running buffer with 20 mM Tris-HCl, 150 mM NaCl, and 1 mM DTT at pH 7.4.

Expression and purification of full-length RIG-I were carried out following the protocol provided by Dr Cheng Kao (Indiana University). Full-length human RIG-I is on pET28a. Rosseta/pET28a/RIG-I bacteria were spread on a solid agar gel plate with 17 µg/mL chloramphenicol and 50 µg/mL kanamycin. Bacteria resistant to both kanamycin and chloramphenicol were grown on an agar plate at 37 °C overnight and a single colony was picked to prepare the starter culture in 20 mL of LB in a 50 mL tube. The culture was incubated for about 5 hours at 37 °C with constant shaking of 200 rpm, before the culture was transferred equally to the six 1 L conical flasks. These flasks were then left for 8 hours in the incubator at 37°C at 200 rpm.

The incubator temperature was lowered to 4 °C. 1.5 hours later, 10 mL ethanol (final concentration of ethanol, v/v=2.5%) was added, and the flask was incubated for another 1.5 hours. Then, 0.2 mM of isopropyl-β-D-1-thiogalactopyranoside (IPTG) was added to induce the expression of RIG-I full length protein overnight for 16 hours at 16 °C at 200 rpm.

The LB medium was then centrifuged for 10 minutes at 10000 rpm. The

supernatant was removed and the sediments were collected. 25 mL of lysis buffer, which was prepared with 50 mM sodium phosphate, 10 mM imidazole, 300 mM NaCl, 10% glycerol, 5 mM beta-mercaptoethanol at pH 8.0, was added together with 800 μ L of protease inhibitor cocktail and 0.1 mg/mL of lysozyme. The cells were then sonicated for 15 minutes, with 5 seconds “on” and 5 seconds “off” in an ice water bath. The resulting mixture was then centrifuged at 8000 rpm for 60 minutes. The supernatant was collected and filtered.

Purification of RIG-I full length was performed on HisTrap column and Superdex 200 by Fast Performance Liquid Chromatography (FPLC).

2.2 RNA synthesis and purification

The RNA oligonucleotides were synthesized using Applied Biosystems 394 DNA/RNA synthesizer (ABI). The synthetic scale is 1 μ mole. Acetonitrile was used to dilute various phosphoramidites to 0.1 M. The synthesis was done from the 3' to 5' direction.

The synthesis column was dried thoroughly using nitrogen gas. Then, the solid support with RNA oligonucleotide was poured into a 15 mL Greiner tube. For cleaving the oligonucleotide from the support and removing the cyanoethyl protecting group, 2 mL of ammoniamethylamine (AMA) was added to the tube, then the tube was heated at 65 °C for 10 minutes. The supernatant solution was

transferred into another 1.5 mL centrifuge tube. We dried the supernatant using a speed-vac machine. After that 115 μ L of dimethyl sulfoxide (DMSO) was added to dissolve the oligonucleotides. Then, 60 μ L triethylamine (TEA) and 75 μ L TEA $\cdot$ 3HF were added and heated at 65 $^{\circ}$ C for 3 hours to remove tert-butyldimethylsilyl (TBDMS) protecting group.

The synthesized RNA oligonucleotides were purified using reverse phase high performance liquid chromatography (RP-HPLC). One buffer contains 0.1 M triethylammonium acetate (TEAA) at pH 7.0 and another one is 100% acetonitrile (ACN). Two peaks appeared during HPLC. The first peak showed oligonucleotides without dimethoxytrityl (DMT) protecting group and the other one was the oligonucleotides with the DMT protecting group. We only collected the RNA with DMT protecting group. Then 1 mL 80 % acetic acid was added and incubated at room temperature for one hour in order to remove the DMT protecting group. HPLC was carried out again to check the purity of the desired RNA.

2.3 RNA binding studies by Native Polyacrylamide Gel Electrophoresis (PAGE)

The glass plates are washed with ultrapure water and ethanol, and then they are cleaned with kimwipes. The spacers are positioned at the edges of glass plates and the plates are held together. The glass plates were sealed with water-proof tapes. 50 mL 12% (w/v) native polyacrylamide gel solution was prepared. 5.7 g acrylamide

and 0.3 g bisacrylamide were added into a 50 mL sterile centrifuge tube. 50 mL of 1x TBE buffer at pH 7.4 was added to dissolve the solid. The tube was then heated in a water bath at 70 °C for 20 minutes to remove air bubbles. After that, it was centrifuged at 25 °C, 4000 rpm for 20 minutes. 250 µL 10% ammonium persulfate (APS) and 50 µL TEMED were added, soon after that we poured the solution into the assembled glass plates carefully and make sure no air bubbles formed. The bottom of the plates was clipped to prevent leakage. The ten or sixteen wells comb was inserted when the solution went up to the top of the plates and the plates were left at room temperature for 1 hour for the acrylamide to polymerize.

1x TBE electrophoresis buffer at pH 7.4 was used to rinse all the wells. Pre-run was carried out at 160 V for 1 hour with 10 µL loading dye to make sure that the gel is prepared well and to remove excess APS. The samples were then loaded and the gel was run at 160 V at 4 °C for 6 hours. Except for Cy3 labelled RNA, the gels were stained by ethidium bromide. 50 µL ethidium bromide was diluted in 500 mL ultrapure water and used to stain the gel. Typhoon scanner was used for imaging.

2.4 Steady-state fluorescence measurement and determination of dissociation constants

RIG-I CTD (with a stock concentration of 500 or 1000 µM and pre-mixed with 1 µM 2AP-labelled RNA) was titrated into the cuvette with 70 µL 2AP-labelled RNA at 1 µM. The concentrations of RIG-I CTD were increased from 0 to ~82 µM

during the titration. Full-length RIG-I and TFO were pre-mixed with 1 μM 2AP-labelled RNA, and incubated overnight before fluorescence measurement to allow the system to reach equilibrium (50). All the steady-state fluorescence data were measured at 20 $^{\circ}\text{C}$ unless otherwise noted. The buffers were shown in the Results and Discussion section. Fluorescence intensity was plotted against the total concentration of protein and TFO, and the following equation was used to obtain the dissociation constant (36,50).

$$y = y_o + \frac{B_{max}}{2R_o} \left[R_o + x + K_D - \sqrt{(R_o + x + K_D)^2 - 4R_o x} \right] \quad (1)$$

Here, y_o and y are the fluorescence intensities in the absence and presence of protein, respectively. The total concentration of protein or TFO is x . B_{max} is maximum fluorescence intensity change achieved upon addition of the ligand. R_o is the concentration of 2AP labelled RNA (1 μM). K_D is equilibrium dissociation constant.

2.5 Time-resolved fluorescence spectroscopy

Time-resolved fluorescence measurements were carried out at 20 $^{\circ}\text{C}$ by the time-correlated single photon counting (TCSPC) technique with a resolution of 10 ps (PicoQuant PicoHarp 300). The second harmonic of a titanium-sapphire laser (Chameleon, Coherent Inc.) at 320 nm (100 fs, 80 MHz) was used as the excitation source.

3. RESULTS AND DISCUSSION

3.1 RIG-I CTD binds to 2AP-Labelled 12-bp RNA duplexes and induces fluorescence intensity enhancement and fluorescence lifetime distribution changes

We incorporated 2AP into self-complementary 12-bp RNA duplexes at terminal and internal positions, respectively (Figure 5A,B). The steady state fluorescence measurements (Figure 6 and Figure 8A-D) reveal that: (1) 2AP fluorescence intensity increases significantly upon the binding of RIG-I CTD to both 12-bp RNAs (with 2AP positioned at or away from binding site), suggesting protein binding-induced changes in the local structure dynamics and/or environment of the (2AP)-U pairs. (2) The fluorescence intensity of the internal 2AP away from binding site increases more significantly than that of terminal 2AP upon RIG-I CTD binding to RNA, which might be due to a quenching effect by the local aromatic stacking interaction between terminal (2AP)-U pair and the phenylalanine residue F853 in RIG-I (Figure 2B). This local stacking interaction is observed in different crystal forms (32-34,36-38), suggesting its importance for the formation of RNA-protein complex. The dissociation constants obtained by fitting the titration curves (Figure 6) are similar to the value for the 12-bp RNA without 2AP previously obtained by isothermal titration calorimetry (ITC) in the same buffer (Table 1) (33).

Figure 6. Fluorescence intensity of 2-AP incorporated 12 bp RNAs increases upon titration with RIG-I CTD. The buffer contains 250 mM NaCl, 2 mM MgCl₂, and 25 mM HEPES at pH 7.5. The excitation wavelength is 320 nm to avoid excitation of fluorescent amino acid residues in RIG-I CTD. The emission intensities at 370 nm are used for fitting

We also run native polyacrylamide gel electrophoresis to confirm whether the 12-bp RNA binds to RIG-I CTD. As shown in lane 3 and lane 7 of Figure 7, we observe that both 12-bp RNA1 and 12-bp RNA2 can bind to RIG-I CTD.

Figure 7. Native PAGE for 12-bp RNA1 and 12-bp RNA2 with and without RIG-I CTD (C), RIG-I CTD K888E mutant (M), and full-length RIG-I (F) in an incubation buffer of 250 mM NaCl, 10 mM MgCl and 25 mM HEPES at pH 7.5. K888E mutant was expressed by replacing the Lysine (K) in the 888th of RIG-I with a Glutamic acid (E). In lane 1 of each gel, no protein was added (N). The RNAs and RIG-I CTD or RIG-I CTD K888E mutant were incubated at room temperature for 1 hour before loading. The full-length RIG-I and RNA were incubated at room temperature overnight before loading. The concentration of each RNA hairpin is 1 μ M. The concentration of RIG-I CTD, RIG-I CTD K888E mutant, and full-length RIG-I is 4 μ M. The gels were post-stained by EtBr.

The time-resolved fluorescence data (Figure 8E,F) reveal that there are four components (44,51,52) of fluorescence lifetimes for 2AP incorporated in 12-bp

RNA, and the amplitudes of the component with ~ 3 ns or ~ 6 ns for 12-bp RNA1 or 12-bp RNA2 increase by 20% or 33% upon binding to RIG-I CTD, respectively. The data suggest that the (2AP)-U pair has four microstates with local “breathing” dynamics at the microsecond time scale (53-57), and RIG-I CTD binding favors the relatively open conformation with relatively long fluorescence lifetime, and thus enhances steady-state fluorescence intensity. Thus, local conformation changes can be induced by both direct RIG-I CTD binding and by allosteric effect. Breathing dynamics within a duplex affects duplex (un)folding equilibrium (53-55). Thus, allosteric modulation of local dynamics may be utilized (by nature) to regulate functions involving RNA duplex (un)folding, such as ribozyme catalysis (14), RNA helicase/chaperone unwinding activity (54,55,58), and riboswitch conformational changes (59). Interestingly, T7 RNA polymerase binding to the promoter region of a DNA duplex also enhances 2AP fluorescence intensity by about 4 fold even though the (2AP)-T pair (at +4 position) is not melted (60). Ultrafast dynamics (e.g., fs-ns dynamics) may also change upon protein binding (45,61-63), which is not the focus of this study.

Figure 8. 2AP fluorescence study of RIG-I CTD binding to RNA duplexes incorporating (2AP)-U pairs (1 μ M). The buffer is 250 mM NaCl, 2 mM $MgCl_2$, and 25 mM HEPES at pH 7.5. (**A,B**) Steady-state fluorescence emission spectra

with the excitation wavelength at 320 nm. **(C,D)** Steady-state fluorescence excitation spectra with the emission wavelength at 370 nm. **(E,F)** Time-resolved 2AP fluorescence for 1 μ M 12-bp 2AP-labelled RNAs in the absence and presence of 10 μ M RIG-I CTD. The excitation wavelength was 320 nm, and the normalized fluorescence emission intensities at 380 nm were recorded and plotted. Fluorescence lifetime, τ , were determined from the fitting of the traces to a four-exponential decay of the form $y = y_0 + \text{IRF}(t_0) * \{A_1 \exp[-(t-t_0)/\tau_1] + A_2 \exp[-(t-t_0)/\tau_2] + A_3 \exp[-(t-t_0)/\tau_3] + A_4 \exp[-(t-t_0)/\tau_4]\}$, where A_1 , A_2 , A_3 , and A_4 are the amplitudes of the four components, respectively. $\text{IRF}(t_0)$ is the instrumental response function, and the fitting starts at t_0 . The average values of lifetimes are calculated based on the following equation.

$$\tau_{\text{ave}} = (A_1 * \tau_1 + A_2 * \tau_2 + A_3 * \tau_3 + A_4 * \tau_4) / (A_1 + A_2 + A_3 + A_4)$$

3.2 RIG-I CTD does not bind to or induce fluorescence changes of an RNA dumbbell structure incorporating with 2AP

In order to investigate whether the fluorescence enhancement of 2AP was induced by specific binding of RNA to RIG-I CTD, we synthesized an 11-bp 30-nt RNA with a dumbbell structure as a control (Figure 5L) (64). The RNA dumbbell structure is expected to have no binding with RIG-I CTD, because the dumbbell structure has two GAAA tetraloops but no solvent exposed base-paired (blunt) terminal end. As shown by native PAGE in lane 9-12 of Figure 9, we do not observe binding between the dumbbell structure and RIG-I CTD. Consistently,

2AP fluorescence intensity remains unchanged with the addition of RIG-I CTD (Figure 10). Thus, the 2AP fluorescence intensity enhancement is due to the specific binding of RIG-I CTD at RNA duplex blunt terminal ends.

Figure 9. Native PAGE for rHP6-13, rHP7-14, and dumbbell RNA with and without RIG-I CTD (C), RIG-I CTD K888E mutant (M), and full-length RIG-I (F) in an incubation buffer of 250 mM NaCl, 10 mM MgCl and 25 mM HEPES at pH 7.5. K888E mutant was expressed by replacing the Lysine (K) in the 888th of RIG-I with a Glutamic acid (E). In lane 1 of each gel, no protein was added (N). The RNAs and RIG-I CTD or RIG-I CTD K888E mutant were incubated at room temperature for 1 hour before loading. The full-length RIG-I and RNA were incubated at room temperature overnight before loading. The concentration of each RNA hairpin is 1 μ M. The concentration of RIG-I CTD, RIG-I CTD K888E mutant, and full-length RIG-I is 4 μ M. The gels were post-stained by EtBr.

Figure 10. Titration of RIG-I CTD into dumbbell RNA monitored by 2AP fluorescence in 250 mM NaCl, 10 mM MgCl₂, 25 mM HEPES at pH 7.5. The excitation wavelength is 320 nm to avoid excitation of fluorescent amino acid residues in RIG-I CTD. The emission intensities at 370 nm are used for fitting. The errors shown are standard deviations.

3.3 RIG-I CTD binds to RNA hairpins with varied stem length and induces distance-dependent 2AP fluorescence enhancement

In order to further investigate the allosteric effect on RNA local structural dynamics induced by RIG-I CTD binding, we incorporated the (2AP)-U pairs 5-14 bp away from the termini of a series of RNA hairpins (Figure 5). RIG-I CTD binds to the hairpins at the terminus but not the UUCG tetraloop, which facilitates the

measurement of the distance-dependent 2AP fluorescence enhancement induced by RIG-I CTD binding. For rHP1-5, rHP2-9, rHP3-11, and rHP4-12 (Figure 5), 2AP fluorescence intensities increase as RIG-I CTD concentration is increased until saturation (Figure 11A). Fitting of the binding curves gives K_D values of about 3 μ M. All the four hairpins (rHP1-5, rHP2-9, rHP3-11, and rHP4-12) have the same sequence near the terminus, thus giving essentially the same K_D value. Fluorescence intensity increases by two fold (from 9.0 to 26.9) for rHP1-5 and about one fold (from 7.6 to 17.3) for rHP2-9 (Figure 11). Only about 50% increase in 2AP fluorescence intensity is observed for rHP3-11 and rHP4-12 (Figure 11B). Our results further verify that RNA local structural dynamics changes upon binding to RIG-I CTD, resulting in the enhancement of 2AP fluorescence intensity in a distance-dependent manner (Figure 11). Interestingly, specific binding of a protein or a small molecule to a Watson-Crick DNA duplex may widen or narrow the width of DNA grooves within about 15 bp, and thus modulate the binding of a second protein or ligand to the DNA for gene regulation (65-68). Thus, allosteric modulation of local “breathing” dynamics may be important for the regulation of the assembly of multiple proteins on one RNA duplex (31,50,69-73).

2AP fluorescence intensity for rHP5-13 and rHP6-13 (with 2AP positioned 13 bp away from the terminus), and rHP7-14 (with 2AP positioned 14 bp away from the terminus), however, does not change significantly upon adding RIG-I CTD (Figure 11A,C). Native PAGE results suggest all of the RNA hairpins studied here (Figure

5) bind to RIG-I CTD (Figure 12). Thus, RIG-I CTD binds to rHP5-13, rHP6-13, and rHP7-14, but does not induce fluorescence intensity changes of 2AP. The local sequence environment of 2AP is the same for rHP1-5, rHP2-9, rHP3-11, rHP4-12, rHP6-13 and rHP7-14. Thus, the allosteric effect induced by RIG-I CTD binding is lost when 2AP is positioned 13 bp away. For rHP5-13, the (2AP)-U pair is the closing base pair of the UUCG tetraloop, which has different local sequence environment and may be intrinsically more dynamic compared to a (2AP)-U pair at internal positions. Consistently, the fluorescence intensity of free rHP5-13 is slightly higher than those of other hairpins (Figure 11A,C).

Figure 11. Titration of RIG-I CTD into RNA monitored by 2AP fluorescence. The excitation wavelength is 320 nm to avoid excitation of fluorescent amino acid residues in RIG-I CTD. The emission intensities at 370 nm are used for fitting. The errors shown are standard deviations. **(A,B)** Titration of RIG-I CTD into rHP1-5, rHP2-9, rHP3-11, rHP4-12, rHP6-13, and rHP7-14 (A), and rHP5-13 (B) in 250 mM NaCl, 10 mM MgCl_2 , 25 mM HEPES at pH 7.5. **(C)** Summary of distance-dependent fluorescence enhancement induced by RIG-I CTD binding. The 2AP fluorescence intensities in the presence of about 82 μM of RIG-I CTD (shown in panel B) were used for analysis. The fluorescence intensity of free RNA was normalized as 1.

Figure 12. Native PAGE for various RNAs with and without RIG-I CTD (C), RIG-I CTD K888E mutant (M), and full-length RIG-I (F) in an incubation buffer of 250 mM NaCl, 10 mM MgCl and 25 mM HEPES at pH 7.5. K888E mutant was expressed by replacing the Lysine (K) in the 888th of RIG-I with a Glutamic acid (E). In lane 1 of each gel, no protein was added (N). The RNAs and RIG-I CTD or RIG-I CTD K888E mutant were incubated at room temperature for 1 hour before loading. The full-length RIG-I and RNA were incubated at room temperature overnight before loading. The concentration of each RNA hairpin is 1 μ M. The concentration of RIG-I CTD, RIG-I CTD K888E mutant, and full-length RIG-I is

4 μ M. The gels were post-stained by EtBr. (A) rHP1-5, rHP2-9, and rHP2-9-AA. (B) rHP3-11, rHP4-12, and rHP5-13. (C) rHP6-13, rHP7-14, and dumbbell RNA.

3.4 Incorporating an A-A mismatch outside the binding site does not affect RIG-I CTD binding but abolishes 2AP fluorescence enhancement

Next, we made an RNA hairpin (rHP2-9-AA) (Figure 5F) by replacing an A-U pair away from RIG-I CTD binding site in rHP2-9 with an A-A mismatch. Remarkably, RIG-I CTD binds to rHP2-9-AA as suggested by the gel result (lane 11 of Figure 12A), but does not induce 2AP fluorescence change (Figure 13). It is likely that the allosteric effect is induced cooperatively through the relatively rigid Watson-Crick base-paired duplex region. A relatively dynamic A-A mismatch (74,75) may decouple the two Watson-Crick duplexes flanking the A-A mismatch, resulting in the loss of the allosteric effect induced by RIG-I CTD binding. Thus, allosteric 2AP fluorescence enhancement is not simply induced directly by the bound RIG-I CTD protein. The presence of mismatches and chemical modifications may introduce a perturbation of local dynamics and abolish the allosteric effect within a duplex. Thus, allosteric effect may be utilized (by nature) for detecting single mismatches and damaged/modified residues in nucleic acid duplexes (65,76). Compared to a 5'-phosphorylated RNA, non-phosphorylated RNA duplex binds to RIG-I CTD slightly differently (32-34). We observed 2AP fluorescence intensity enhancement for an RNA duplex with a 5'-triphosphate group (Figure 14). Thus, 2AP fluorescence enhancement upon binding to RIG-I CTD is not sensitive to structural variation due to the presence of a 5'-triphosphate group.

Figure 13. Titration of RIG-I CTD into rHP2-9-AA monitored by 2AP fluorescence in 250 mM NaCl, 10 mM MgCl₂, 25 mM HEPES at pH 7.5. The excitation wavelength is 320 nm to avoid excitation of fluorescent amino acid residues in RIG-I CTD. The emission intensities at 370 nm are used for fitting. The errors shown are standard deviations.

Figure 14. 2AP fluorescence titration of RIG-I CTD into 24-bp RNA1, 24-bp RNA2, 24-bp RNA3 and a 24-nt single-stranded RNA (ssRNA) in 250 mM NaCl, 10 mM MgCl₂, 25 mM HEPES at pH 7.5. The concentration of all RNAs is 1 μ M. RIG-I CTD binds to 24-bp RNA2 (with a 5' triphosphate group) more tightly than 24-bp RNA1 (with the apparent K_D 1.54 versus 4.70 μ M). 2AP fluorescence is not enhanced upon RIG-I CTD binding to 24-bp RNA3, which has two A-A mismatches. 2AP fluorescence is enhanced upon RIG-I CTD binding to 24-nt ssRNA, suggesting the 2AP residue may be indirect with RIG-I CTD. The error bars shown are standard deviations

3.5 RIG-I CTD K888E mutant does not bind to or induce fluorescence changes of RNA hairpins incorporating 2AP

A substitution of lysine 888 (K888) to glutamic acid (E) in RIG-I CTD has been shown to disrupt the binding with RNA (77). Our native PAGE results confirm that there is no binding between RNA hairpins and RIG-I CTD K888E mutant (Figure 9 and Figure 12). Consistently, 2AP fluorescence intensity remains unchanged with the addition of RIG-I CTD K888E mutant (Figure 15).

Figure 15. 2AP fluorescence titration of RIG-I CTD K888E mutant into rHP1-5, rHP2-9, 24-bp RNA1, and 24-bp RNA 2 in 250 mM NaCl, 10 mM MgCl₂, 25 mM HEPES at pH 7.5. K888E mutant was expressed by replacing the Lysine (K) in the 888th of RIG-I with a Glutamic acid (E). 2AP fluorescence is not enhanced upon adding RIG-I CTD K888E mutant, which is consistent with the fact that RIG-I

CTD K888E mutant does not bind to rHP1-5, rHP2-9, 24-bp RNA1, and 24-bp RNA2. The error bars shown are standard deviations.

Apparently, 2AP fluorescence enhancement is induced by specific binding of RIG-I CTD to RNA, but not merely solution environment changes upon the addition of a protein.

3.6 RIG-I CTD binding to RNA at various salt conditions measured by 2AP fluorescence changes and PAGE

A large range of dissociation constants ranging from nM to μ M (Table 1) has been reported (32,33,36,50,78) for RIG-I CTD binding to RNA. It is probable that the discrepancy results from variations in experimental condition and/or molecular constructs. For example, charge-charge (ionic) interactions contribute significantly to the binding of RIG-I CTD to RNA, and cause the binding strength sensitive to salt conditions. Hence, we measured the binding affinity between rHP1-5 and RIG-I CTD at various salt conditions by using 2AP fluorescence change as a probe, which has minimal perturbation to RNA structures. 2AP fluorescence intensity is insensitive to changes in salt condition (Figure 16), suggesting allostery is not due to long range electrostatic effect. However, the binding strength between RIG-I CTD is strongly dependent on salt condition. When the concentration of MgCl_2 is increased from 2 mM to 10 mM with a constant NaCl concentration at 250 mM, K_D increases from very low to 2.91 μ M (Figure 16B). When the concentration of NaCl is increased from 250 mM to 500 mM with a constant MgCl_2 concentration

at 10 mM, K_D increases moderately from 2.91 μM to 5.84 μM (Figure 16B). At 100 mM NaCl, 2 mM MgCl_2 , the K_D value is too low (nanomolar range) to be accurately measured with the concentration of 2AP-labelled RNA at 1 μM . The binding curve, however, allows for the estimation of binding ratio (RNA:protein) to be 1:1 (Figure 16A).

To verify the binding monitored by the 2AP fluorescence change, we carried out independent binding measurements by native PAGE by employing a Cy3-labelled rHP1 (rHP1-Cy3, Figure 5C and Figure 17) at a relatively low concentration (10 nM). To avoid the interference of the binding to RIG-I CTD and protein induced fluorescence changes of Cy3, we incorporated Cy3 in the tetraloop. The K_D values obtained from PAGE agree with that from the 2AP fluorescence titration method, except for the incubation buffer with 100 mM NaCl, 2 mM MgCl_2 (Figure 16 and Figure 17, Table 1). Thus, RIG-I CTD binding can be quantified by monitoring the allosteric changes in the local dynamics and fluorescence intensity of 2AP.

Figure 16. Binding analysis of RIG-I CTD and rHP1-5 (1 μM) monitored by 2AP fluorescence in various concentrations of NaCl and MgCl_2 . The errors shown are standard deviations. **(A)** Binding in 25 mM HEPES, pH 7.5 with varied concentrations of NaCl and MgCl_2 in linear X scale. The binding is too tight to be accurately measured at 100 mM NaCl (or 250 mM NaCl) and 2 mM MgCl_2 . The binding ratio (RNA:protein) is about 1:1 based on the linear fitting results using first eight and last five data points in the curves. **(B)** Binding in 25 mM HEPES, pH 7.5 with varied concentrations of NaCl and MgCl_2 in logarithmic X scale.

Figure 17. Native PAGE for the binding of RIG-I CTD to rHP1-Cy3 (10 nM) in 25 mM HEPES, pH 7.5 in various concentrations of NaCl and MgCl₂. The numbers shown on the top of the gel are the total concentrations of proteins in nM. The gels were imaged by Cy3 fluorescence with the excitation wavelength at 532 nm and emission wavelength at 580 nm. RIG-I CTD binding does not affect Cy3 fluorescence intensity for rHP1-Cy3. (A) 100 mM NaCl and 2 mM MgCl₂, (B) 100 mM NaCl and 10 mM MgCl₂, (C) 250 mM NaCl and 2 mM MgCl₂, (D) 250 mM NaCl and 10 mM MgCl₂, (E) 500 mM NaCl and 2 mM MgCl₂, (F) 500 mM NaCl

and 10 mM MgCl₂. (G) Binding analysis of the native PAGE results in 25 mM HEPES, pH 7.5 in various concentrations of NaCl and MgCl₂

3.7 Full-length RIG-I and TFO binding to RNA measured by 2AP fluorescence changes and PAGE

Recent crystal structures of the complexes of multidomain RIG-I and RNA duplex (36-38) suggest that RIG-I helicase and CTD domains bind synergistically to an RNA duplex with CTD partially capping one terminal end of the RNA duplex and helicase domain surrounding the adjacent duplex region (Figure 2A,C). The native PAGE result (Figure 18E,F) reveals that, in 250 mM NaCl and 10 mM MgCl₂ at pH 7.5, the K_D of rHP1-Cy3 binding to full-length RIG-I is 60 nM, comparable to the previous studies (Table 1) (36,50,78). Our results also show that full-length RIG-I can bind to an RNA hairpin (i.e., rHP1-Cy3) with only 8 bp, consistent with previous structural and functional assays (71,72). 2AP fluorescence intensity increases for rHP2-9, rHP2-9-AA, rHP5-13, rHP6-13, and rHP7-14 upon addition of full-length RIG-I (Figure 19). The binding ratio (RNA:protein) is close to 1:1, which is consistent with the native PAGE result (Figure 20). Apparently, full-length RIG-I can induce 2AP fluorescence change relatively independent of RNA duplex structure. Interestingly, full-length RIG-I does not bind to the dumbbell RNA (with an 11-bp coaxially stacked stem) as shown by the PAGE (lane 12 of Figure 9), and consistently, 2AP fluorescence intensity remains unchanged upon adding full-length RIG-I (Figure 19F).

Figure 18. Cy3 fluorescence of rHP1-Cy3 (Fig. 2C) in the solution and in the gel. (A-C) Steady-state Cy3 fluorescence of 1 μ M rHP1-Cy3 with and without RIG-I CTD (A), full-length RIG-I (B) and RIG-I CTD K888E mutant (C) in 250 mM NaCl, 10 mM MgCl_2 , 25 mM HEPES at pH 7.5 at 20 $^{\circ}\text{C}$. The excitation wavelength is at 530 nm. Binding of RIG-I CTD and full-length RIG-I and the presence of RIG-I CTD K888E mutant do not affect Cy3 fluorescence intensity, consistent with our PAGE studies (Fig. 5). (D,E,F) Native PAGE and analysis for the binding of full-length RIG-I to rHP1-Cy3 (10 nM) in 25 mM HEPES, pH 7.5, 250 mM NaCl, 10 mM MgCl_2 , and (D) with 1 mM DTT or (E) without DTT. Adding 1 mM DTT does not affect the dissociation constant (K_D) significantly (F). (G) Chemical structure of the Cy3 dye with the trade name Quasar 570.

Figure 19. Titration of full-length RIG-I into rHP2-9 (A), rHP2-9-AA (B), rHP5-13 (C), rHP6-13 (D), rHP7-14 (E), and dumbbell (F). RNA monitored by 2AP fluorescence in 250 mM NaCl, 10 mM MgCl_2 , 25 mM HEPES at pH 7.5. The concentration of all the RNAs is 1 μM . (A-E) The dissociation constants are too small to be accurately measured with 1

μM 2AP-labelled RNA. The crossing points of the two linear fits (First eight and last five data points are used to fit) at low and high protein concentrations are shown, suggesting the binding ratio (RNA:protein) is close to 1:1. (F) No binding curve was observed for dumbbell RNA upon addition of full-length RIG-I.

Figure 20. Native PAGE study of full-length RIG-I binding to 1 μ M RNAs in 250 mM NaCl, 10 mM MgCl₂, 25 mM HEPES at pH 7.5. The values shown on the top of the gel are the total concentrations (in nM) of full-length RIG-I. The full-length RIG-I-RNA complexes were incubated at room temperature overnight before

loading. The gel was post-stained by EtBr. Only one band of RIG-I-RNA complex was observed for rHP6-13, rHP7-14, and a 24-nt ssRNA, suggesting the binding ratio is 1:1. Consistently, the binding ratio (RNA:protein) is estimated to be 1:1. The band intensities for free RNA and RNA-protein complex are normalized for analysis. Two bands of RNA-protein complexes were observed for 12-bp RNA1 and 24-bp RNA1, suggesting a total of two copies of full-length RIG-I can bind to one duplex with 12 bp or 24 bp. **(A,B)** 16-bp rHP6-13. **(C,D)** 17-bp rHP7-14. **(E,F)** 12-bp RNA1. **(G,H)** 24-bp RNA1. **(I,J)** A 24-nt ssRNA (5'-GACGCAGACCCUGAAGUACAUCUU-3').

Finally, 2AP fluorescence intensity of rHP7-14 is enhanced moderately (Figure 21) upon binding to a negatively charged TFO to form a triplex (79). The (2AP)-U pair is positioned 1 bp away from the triplex. The K_D value obtained from the 2AP fluorescence enhancement is 1.91 μ M (Figure 21A,B), in agreement with the native PAGE result (1.83 μ M in Figure 21C,D). Thus, the allosteric effect within a Watson-Crick RNA duplex can be induced by both protein binding and tertiary interactions.

Figure 21. 2AP fluorescence titration and native PAGE experiments for the binding between a TFO (5'-UCUCUCUUCUUC-3', U: s²U^L) and rHP7-14 (Fig. 2K) in 250 mM NaCl, 2 mM MgCl₂, 25 mM HEPES at pH 7.5. The concentration of rHP7-14 is kept at 1 μM . (**A,B**) Binding monitored by 2AP fluorescence. (**C,D**)

Binding measured by native PAGE. The gel was post-stained by EtBr. The values shown on the top of the gel are the total concentrations (in μM) of the TFO. **(E)** Triplex structure formed between rHP7-14 and the TFO. **(F)** Chemical structure of the nucleoside LNA 2-thio U ($\text{s}^2\text{U}^{\text{L}}$).

The full-length RIG-I binding to RNAs is always too tight to be accurately measured. However, we may incorporate multiple 2AP residues in one double-strand RNA for measuring tight binding, and to increase the dynamic range of fluorescence intensity signals for competitive binding assay. Allosteric modulation of RNA local structural dynamics induced by ligand binding may also be utilized to study other RNA-ligand interactions.

Table 1. Binding dissociation constants of RIG-I CTD and full-length RIG-I to RNAs measured by various techniques.

Buffer	Method	RNA (5'-OH)	K_D (nM)	Protein
250 mM NaCl, 2 mM MgCl ₂ , 25 mM HEPES, pH 7.5	ITC (33) 25 °C	Self-complementary 12-bp RNA	1095.3	RIG-I CTD
	2AP fluorescence (this work) 20 °C	12-bp RNA1	585 ± 88	
		12-bp RNA2	889 ± 101	
100 mM NaCl, 2 mM MgCl ₂ , 25 mM HEPES, pH 7.5	2AP fluorescence (this work) 20 °C	rHP1-5	N/A ^a	RIG-I CTD
	Cy3 PAGE (this work) 4 °C	rHP1-Cy3	29 ± 11	
100 mM NaCl, 10 mM MgCl ₂ , 25 mM HEPES, pH 7.5	2AP fluorescence (this work) 20 °C	rHP1-5	1674 ± 285	RIG-I CTD
	Cy3 PAGE (this work) 4 °C	rHP1-Cy3	989 ± 203	
250 mM NaCl, 2 mM MgCl ₂ , 25 mM HEPES, pH 7.5	2AP fluorescence (this work) 20 °C	rHP1-5	247 ± 69	RIG-I CTD
	Cy3 PAGE (this work) 4 °C	rHP1-Cy3	771 ± 93	
250 mM NaCl, 10 mM MgCl ₂ , 25 mM HEPES, pH 7.5	2AP fluorescence (this work) 20 °C	rHP1-5	2908 ± 515	RIG-I CTD
			2515 ± 304	
	Cy3 PAGE (this work) 4 °C	rHP1-Cy3	60 ± 20	Full-length RIG-I
500 mM NaCl, 2 mM MgCl ₂ , 25 mM HEPES, pH 7.5	2AP fluorescence (this work) 20 °C	rHP1-5	1477 ± 100	RIG-I CTD
	Cy3 PAGE (this work) 4 °C	rHP1-Cy3	1266 ± 116	
500 mM NaCl, 10 mM MgCl ₂ , 25 mM HEPES, pH 7.5	2AP fluorescence (this work) 20 °C	rHP1-5	5836 ± 800	RIG-I CTD
	Cy3 PAGE (this work) 4 °C	rHP1-Cy3	5079 ± 795	
150 mM NaCl, 25 mM HEPES, 1.5 mM DTT, 0.1 mg/ml BSA, pH 7.4	PAGE with RNA labeled with ³² P (50) 4 °C	Non self-complementary 14-bp RNA (One strand was labeled with 5' α- ³² P)	5900 ± 750	RIG-I CTD
			20 ± 3	Full-length RIG-I

150 mM NaCl, 3 mM EDTA, 10 mM HEPES, 0.05% v/v P-20, pH 7.4	SPR (32) 25 °C	Non self-complementary 14-bp RNA (RIG-I CTD was immobilized.)	4.6 ± 0.6	RIG-I CTD
100 mM CH ₃ COONa, 10 mM Mg(CH ₃ COO) ₂ , 50 mM Tris-acetate, 5% glycerol, 5 mM DTT, 0.05% Tween 20, pH 7.5	Fluorescence anisotropy (36) 25 °C	Non self-complementary 14-bp RNA (One strand was labeled with 5'-fluorescein)	2.6 ± 1.2	RIG-I CTD
			1.2 ± 0.6	Full-length RIG-I
10 mM KCl, 5 mM MgCl ₂ , 20 mM Tris-HCl, 4% (w/v) glycerol, pH 7.5, 1 mg/ml glucose oxidase, 0.4% (w/v) D-glucose, 0.4 mg/ml catalase, 1% (v/v) 2-mercaptoethanol	Single-molecule fluorescence (78) 37 °C	Non-selfcomplementary 25-bp RNA (One strand is labelled with 3'-DY547, and the RNA was immobilized)	434.4 ± 48.7	Full-length RIG-I
	Fluorescence anisotropy (78) 37 °C		493.5 ± 35.7	

^aN/A: The dissociation constant is too small to be accurately measured with 2AP-labelled RNA at 1 μ M.

4. CONCLUSION

In this part, we have replaced Watson-Crick pair (A-U or G-C) with (2AP)-U pair in various RNA duplexes, and tested their binding to RIG-I CTD and full length RIG-I using steady state and time-resolved fluorescence assays. Binding was also confirmed by gel electrophoresis. RNA hairpin labeled with Cy3 was used to quantify the binding affinity by gel electrophoresis. The steady state 2AP fluorescence measurements reveal that 2AP fluorescence increases significantly upon binding of RIG-I CTD to RNA. Time-resolved fluorescence data suggest that 2AP has high local dynamics and RIG-I CTD binding favors the conformation with long fluorescence lifetime, and thus enhances fluorescence intensity. RIG-I CTD mutant K888E does not bind to RNA or induce 2AP fluorescence change, suggesting 2AP fluorescence changes are induced by specific RIG-I CTD binding to RNA. RIG-I CTD does bind to the 24-bp duplex with two A-A mismatches as suggested by the gel result, but does not induce 2AP fluorescence change. It indicates that RIG-I binding induces 2AP fluorescence changes in a distance and structure dependent manner. We also test the RIG-I CTD binding to rHP1 at various salt conditions measured by 2AP fluorescence changes and gel electrophoresis, and we found the binding is sensitive to salt conditions. Full-length RIG-I binding also induces 2AP fluorescence enhancement, but the binding affinity is

5. REFERENCES

1. Bevilacqua PC, Russell R (2008) Editorial overview: exploring the vast dynamic range of RNA dynamics. *Curr Opin Chem Biol* 12: 601-603.
2. Nguyen P, Qin PZ (2012) RNA dynamics: perspectives from spin labels. *Wiley Interdiscip Rev RNA* 3: 62-72.
3. Tinoco I, Bustamante C (1999) How RNA folds. *J Mol Biol* 293: 271-281.
4. Turner DH (2013) Fundamental Interactions in RNA: Questions Answered and Remaining. *Biopolymers* 99: 1097-1104.
5. Chen G, Kennedy SD, Turner DH (2009) A CA⁺ pair adjacent to a sheared GA or AA pair stabilizes size-symmetric RNA internal loops. *Biochemistry* 48: 5738-5752.
6. Walter NG, Burke JM, Millar DP (1999) Stability of hairpin ribozyme tertiary structure is governed by the interdomain junction. *Nat Struct Biol* 6: 544-549.
7. Paulsen RB, Seth PP, Swayze EE, Griffey RH, Skalicky JJ, Cheatham TE, Davis DR (2010) Inhibitor-induced structural change in the HCV IRES domain IIa RNA. *Proc Natl Acad Sci USA* 107: 7263-7268.
8. Bailor MH, Sun XY, Al-Hashimi HM (2010) Topology links RNA secondary structure with global conformation, dynamics, and adaptation. *Science* 327: 202-206.
9. Behrouzi R, Roh JH, Kilburn D, Briber RM, Woodson SA (2012) Cooperative tertiary interaction network guides RNA folding. *Cell* 149: 348-357.
10. Sattin BD, Zhao W, Travers K, Chu S, Herschlag D (2008) Direct measurement of tertiary contact cooperativity in RNA folding. *J Am Chem Soc* 130: 6085-6087.
11. Baird NJ, Westhof E, Qin H, Pan T, Sosnick TR (2005) Structure of a folding intermediate reveals the interplay between core and peripheral elements in RNA folding. *J Mol Biol* 352: 712-722.
12. McGinnis JL, Dunkle JA, Cate JHD, Weeks KM (2012) The mechanisms of RNA SHAPE chemistry. *J Am Chem Soc* 134: 6617-6624.
13. Butcher SE, Pyle AM (2011) The molecular interactions that stabilize RNA tertiary structure: RNA motifs, patterns, and networks. *Acc Chem Res* 44: 1302-1311.
14. Wilson TJ, Nahas M, Ha T, Lilley DM (2005) Folding and catalysis of the hairpin ribozyme. *Biochem Soc Trans* 33: 461-465.
15. Ha T, Zhuang X, Kim HD, Orr JW, Williamson JR, Chu S (1999) Ligand-induced conformational changes observed in single RNA molecules. *Proc Natl Acad Sci USA* 96: 9077-9082.
16. Brenner MD, Scanlan MS, Nahas MK, Ha T, Silverman SK (2010) Multivector fluorescence analysis of the xpt guanine riboswitch aptamer domain and the conformational role of guanine. *Biochemistry* 49: 1596-1605.
17. Stone MD, Mihalusova M, O'Connor C M, Prathapam R, Collins K, Zhuang X (2007) Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein. *Nature* 446: 458-461.
18. Caprara MG, Chatterjee P, Solem A, Brady-Passerini KL, Kaspar BJ (2007) An allosteric-feedback mechanism for protein-assisted group I intron splicing. *RNA* 13: 211-222.

19. Martick M, Scott WG (2006) Tertiary contacts distant from the active site prime a ribozyme for catalysis. *Cell* 126: 309-320.
20. Nelson JA, Uhlenbeck OC (2006) When to believe what you see. *Mol Cell* 23: 447-450.
21. Westhof E (2007) A tale in molecular recognition: the hammerhead ribozyme. *J Mol Recognit* 20: 1-3.
22. Veeraraghavan N, Bevilacqua PC, Hammes-Schiffer S (2010) Long-distance communication in the HDV ribozyme: insights from molecular dynamics and experiments. *J Mol Biol* 402: 278-291.
23. Haller A, Souliere MF, Micura R (2011) The dynamic nature of RNA as key to understanding riboswitch mechanisms. *Acc Chem Res* 44: 1339-1348.
24. Liberman JA, Wedekind JE (2012) Riboswitch structure in the ligand-free state. *Wiley Interdiscip Rev RNA* 3: 369-384.
25. McDowell SE, Jun JM, Walter NG (2010) Long-range tertiary interactions in single hammerhead ribozymes bias motional sampling toward catalytically active conformations. *RNA* 16: 2414-2426.
26. Wood S, Ferre-D'Amare AR, Rueda D (2012) Allosteric tertiary interactions preorganize the c-di-GMP riboswitch and accelerate ligand binding. *ACS Chem Biol* 7: 920-927.
27. Lee ER, Baker JL, Weinberg Z, Sudarsan N, Breaker RR (2010) An allosteric self-splicing ribozyme triggered by a bacterial second messenger. *Science* 329: 845-848.
28. Yoneyama M, Kikuchi M, Natsukawa T, Shinobu N, Imaizumi T, Miyagishi M, Taira K, Akira S, Fujita T (2004) The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat Immunol* 5: 730-737.
29. Satoh T, Kato H, Kumagai Y, Yoneyama M, Sato S, Matsushita K, Tsujimura T, Fujita T, Akira S, Takeuchi O (2010) LGP2 is a positive regulator of RIG-I- and MDA5-mediated antiviral responses. *Proceedings of the National Academy of Sciences of the United States of America* 107: 1512-1517.
30. Kawai T, Akira S (2008) Toll-like receptor and RIG-I-like receptor signaling. *Annals of the New York Academy of Sciences* 1143: 1-20.
31. Leung DW, Amarasinghe GK (2012) Structural insights into RNA recognition and activation of RIG-I-like receptors. *Curr Opin Struct Biol* 22: 297-303.
32. Lu C, Xu HY, Ranjith-Kumar CT, Brooks MT, Hou TY, Hu F, Herr AB, Strong RK, Kao CC, Li PW (2010) The structural basis of 5'-triphosphate double-stranded RNA recognition by RIG-I C-terminal domain. *Structure* 18: 1032-1043.
33. Wang YL, Ludwig J, Schuberth C, Goldeck M, Schlee M, Li HT, Juranek S, Sheng G, Micura R, Tuschl T, Hartmann G, Patel DJ (2010) Structural and functional insights into 5'-ppp RNA pattern recognition by the innate immune receptor RIG-I. *Nat Struct Mol Biol* 17: 781-U719.
34. Lu C, Ranjith-Kumar CT, Hao LJ, Kao CC, Li PW (2011) Crystal structure of RIG-I C-terminal domain bound to blunt-ended double-strand RNA without 5' triphosphate. *Nucleic Acids Res* 39: 1565-1575.
35. Kato H, Takahashi K, Fujita T (2011) RIG-I-like receptors: cytoplasmic sensors for non-self RNA. *Immunol Rev* 243: 91-98.
36. Jiang FG, Ramanathan A, Miller MT, Tang GQ, Gale M, Patel SS, Marcotrigiano J (2011) Structural basis of RNA recognition and activation by innate immune receptor RIG-I.

- Nature* 479: 423-U184.
37. Kowalinski E, Lunardi T, McCarthy AA, Louber J, Brunel J, Grigorov B, Gerlier D, Cusack S (2011) Structural basis for the activation of innate immune pattern-recognition receptor RIG-I by viral RNA. *Cell* 147: 423-435.
 38. Luo DH, Ding SC, Vela A, Kohlway A, Lindenbach BD, Pyle AM (2011) Structural insights into RNA recognition by RIG-I. *Cell* 147: 409-422.
 39. Nallagatla SR, Toroney R, Bevilacqua PC (2008) A brilliant disguise for self RNA 5'-end and internal modifications of primary transcripts suppress elements of innate immunity. *RNA biology* 5: 140-144.
 40. Hwang SY, Sun HY, Lee KH, Oh BH, Cha YJ, Kim BH, Yoo JY (2012) 5'-Triphosphate-RNA-independent activation of RIG-I via RNA aptamer with enhanced antiviral activity. *Nucleic Acids Res* 40: 2724-2733.
 41. Ranjith-Kumar CT, Murali A, Dong W, Srisathiyarayanan D, Vaughan R, Ortiz-Alacantara J, Bhardwaj K, Li XJ, Li PW, Kao CC (2009) Agonist and Antagonist Recognition by RIG-I, a Cytoplasmic Innate Immunity Receptor. *J Biol Chem* 284: 1155-1165.
 42. Sinkeldam RW, Greco NJ, Tor Y (2010) Fluorescent analogs of biomolecular building blocks: Design, properties, and applications. *Chem Rev* 110: 2579-2619.
 43. Ward DC, Reich E, Stryer L (1969) Fluorescence studies of nucleotides and polynucleotides. I. Formycin 2-aminopurine riboside 2,6-diaminopurine riboside and their derivatives. *J Biol Chem* 244: 1228-&.
 44. Lee BJ, Barch M, Castner EW, Volker J, Breslauer KJ (2007) Structure and dynamics in DNA looped domains: CAG triplet repeat sequence dynamics probed by 2-aminopurine fluorescence. *Biochemistry* 46: 10756-10766.
 45. Kadakkuzha BM, Zhao L, Xia TB (2009) Conformational distribution and ultrafast base dynamics of leadzyme. *Biochemistry* 48: 3807-3809.
 46. Guest CR, Hochstrasser RA, Sowers LC, Millar DP (1991) Dynamics of mismatched base pairs in DNA. *Biochemistry* 30: 3271-3279.
 47. Stivers JT (1998) 2-aminopurine fluorescence studies of base stacking interactions at abasic sites in DNA: metal-ion and base sequence effects. *Nucleic Acids Res* 26: 3837-3844.
 48. Law SM, Eritja R, Goodman MF, Breslauer KJ (1996) Spectroscopic and calorimetric characterizations of DNA duplexes containing 2-aminopurine. *Biochemistry* 35: 12329-12337.
 49. Nordlund TM, Andersson S, Nilsson L, Rigler R, Graslund A, McLaughlin LW (1989) Structure and dynamics of a fluorescent DNA oligomer containing the EcoRI recognition sequence: fluorescence, molecular dynamics, and NMR studies. *Biochemistry* 28: 9095-9103.
 50. Vela A, Fedorova O, Ding SC, Pyle AM (2012) The thermodynamic basis for viral RNA detection by the RIG-I innate immune sensor. *J Biol Chem* 287: 42564-42573.
 51. Hochstrasser RA, Carver TE, Sowers LC, Millar DP (1994) Melting of a DNA helix terminus within the active-site of a DNA-polymerase. *Biochemistry* 33: 11971-11979.
 52. Tsujikawa L, Strainic MG, Watrob H, Barkley MD, deHaseth PL (2002) RNA polymerase alters the mobility of an A-residue crucial to polymerase-induced melting of promoter

- DNA. *Biochemistry* 41: 15334-15341.
53. Jose D, Datta K, Johnson NP, von Hippel PH (2009) Spectroscopic studies of position-specific DNA "breathing" fluctuations at replication forks and primer-template junctions. *Proc Natl Acad Sci USA* 106: 4231-4236.
 54. Jose D, Weitzel SE, von Hippel PH (2012) Breathing fluctuations in position-specific DNA base pairs are involved in regulating helicase movement into the replication fork. *Proc Natl Acad Sci USA* 109: 14428-14433.
 55. von Hippel PH, Johnson NP, Marcus AH (2013) Fifty years of DNA "Breathing": Reflections on old and new approaches. *Biopolymers* 99: 923-954.
 56. Gherghe CM, Shajani Z, Wilkinson KA, Varani G, Weeks KM (2008) Strong correlation between SHAPE chemistry and the generalized NMR order parameter (S-2) in RNA. *J Am Chem Soc* 130: 12244-12245.
 57. Souliere MF, Haller A, Rieder R, Micura R (2011) A Powerful Approach for the Selection of 2-Aminopurine Substitution Sites to Investigate RNA Folding. *J Am Chem Soc* 133: 16161-16167.
 58. Cheng W, Dumont S, Tinoco I, Jr., Bustamante C (2007) NS3 helicase actively separates RNA strands and senses sequence barriers ahead of the opening fork. *Proc Natl Acad Sci USA* 104: 13954-13959.
 59. Reining A, Nozinovic S, Schlepckow K, Buhr F, Furtig B, Schwalbe H (2013) Three-state mechanism couples ligand and temperature sensing in riboswitches. *Nature* 499: 355-U135.
 60. Bandwar RP, Patel SS (2001) Peculiar 2-aminopurine fluorescence monitors the dynamics of open complex formation by bacteriophage T7 RNA polymerase. *J Biol Chem* 276: 14075-14082.
 61. Pal SK, Zhao L, Xia T, Zewail AH (2003) Site- and sequence-selective ultrafast hydration of DNA. *Proc Natl Acad Sci USA* 100: 13746-13751.
 62. Sontz PA, Muren NB, Barton JK (2012) DNA charge transport for sensing and signaling. *Acc Chem Res* 45: 1792-1800.
 63. Zhang Q, Sun XY, Watt ED, Al-Hashimi HM (2006) Resolving the motional modes that code for RNA adaptation. *Science* 311: 653-656.
 64. Meulen KAV, Davis JH, Foster TR, Record MT, Butcher SE (2008) Thermodynamics and folding pathway of tetraloop receptor-mediated RNA helical packing. *J Mol Biol* 384: 702-717.
 65. Hogan M, Dattagupta N, Crothers DM (1979) Transmission of allosteric effects in DNA. *Nature* 278: 521-524.
 66. Moretti R, Donato LJ, Brezinski ML, Stafford RL, Hoff H, Thorson JS, Dervan PB, Ansari AZ (2008) Targeted chemical wedges reveal the role of allosteric DNA modulation in protein-DNA assembly. *ACS Chem Biol* 3: 220-229.
 67. Kim S, Brostromer E, Xing D, Jin JS, Chong SS, Ge H, Wang SY, Gu C, Yang LJ, Gao YQ, Su XD, Sun YJ, Xie XS (2013) Probing allostery through DNA. *Science* 339: 816-819.
 68. Hancock SP, Ghane T, Cascio D, Rohs R, Di Felice R, Johnson RC (2013) Control of DNA minor groove width and Fis protein binding by the purine 2-amino group. *Nucleic Acids Res* 41: 6750-6760.
 69. Kolakofsky D, Kowalinski E, Cusack S (2012) A structure-based model of RIG-I

- activation. *RNA* 18: 2118-2127.
70. Berke IC, Li Y, Modis Y (2013) Structural basis of innate immune recognition of viral RNA. *Cell Microbiol* 15: 386-394.
 71. Luo D, Kohlway A, Vela A, Pyle AM (2012) Visualizing the determinants of viral RNA recognition by innate immune sensor RIG-I. *Structure* 20: 1983-1988.
 72. Kohlway A, Luo D, Rawling DC, Ding SC, Pyle AM (2013) Defining the functional determinants for RNA surveillance by RIG-I. *EMBO Rep* 14: 772-779.
 73. Patel JR, Jain A, Chou YY, Baum A, Ha T, Garcia-Sastre A (2013) ATPase-driven oligomerization of RIG-I on RNA allows optimal activation of type-I interferon. *EMBO reports* 14: 780-787.
 74. Chen G, Kennedy SD, Qiao J, Krugh TR, Turner DH (2006) An alternating sheared AA pair and elements of stability for a single sheared purine-purine pair flanked by sheared GA pairs in RNA. *Biochemistry* 45: 6889-6903.
 75. Yildirim I, Park H, Disney MD, Schatz GC (2013) A dynamic structural model of expanded RNA CAG repeats: a refined X-ray structure and computational investigations using molecular dynamics and umbrella sampling simulations. *J Am Chem Soc* 135: 3528-3538.
 76. Yi C, Chen B, Qi B, Zhang W, Jia G, Zhang L, Li CJ, Dinner AR, Yang CG, He C (2012) Duplex interrogation by a direct DNA repair protein in search of base damage. *Nat Struct Mol Biol* 19: 671-676.
 77. Cui S, Eisenacher K, Kirchhofer A, Brzozka K, Lammens A, Lammens K, Fujita T, Conzelmann KK, Krug A, Hopfner KP (2008) The C-terminal regulatory domain is the RNA 5'-triphosphate sensor of RIG-I. *Mol Cell* 29: 169-179.
 78. Myong S, Cui S, Cornish PV, Kirchhofer A, Gack MU, Jung JU, Hopfner KP, Ha T (2009) Cytosolic viral sensor RIG-I is a 5'-triphosphate-dependent translocase on double-stranded RNA. *Science* 323: 1070-1074.
 79. Zhou Y, Kierzek E, Loo ZP, Antonio M, Yau YH, Chuah YW, Geifman-Shochat S, Kierzek R, Chen G (2013) Recognition of RNA duplexes by chemically modified triplex-forming oligonucleotides. *Nucleic Acids Res* 41: 6664-6673.

LIST OF PUBLICATIONS

1. Zhou Y, Kierzek E, Loo ZP, Antonio M, Yau YH, Chuah YW, Geifman-Shochat S, Kierzek R, Chen G. (2013) Recognition of RNA duplexes by chemically modified triplex-forming oligonucleotides. *Nucleic Acids Res* 41: 6664-6673.
2. Gitli Devi, Yuan Zhou, and Gang Chen. (2014) RNA Triplexes – From Structural Principles to Biological and Biotech Applications. *WIREs RNA* reviewed.
3. Zhou Y, Tong CL, Narayanan, R, Kierzek E, Kierzek R, Chen G. (2014) Allostery within a Watson-Crick RNA Duplex and Its Application in Studying RNA Recognition by RIG-I. *Submitted*.