Combining single-molecule manipulation and peptide nucleic acid binding studies for unraveling how RNA structures regulate ribosomal frameshifting and alternative splicing

Yang, Lixia

2019

Yang, L. (2019). Combining single-molecule manipulation and peptide nucleic acid binding studies for unraveling how RNA structures regulate ribosomal frameshifting and alternative splicing. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/104424>

<https://doi.org/10.32657/10220/48612>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**COMBINING SINGLE-MOLECULE MANIPULATION
AND PEPTIDE NUCLEIC ACID BINDING STUDIES
FOR UNRAVELING HOW RNA STRUCTURES
REGULATE RIBOSOMAL FRAMESHIFTING
AND ALTERNATIVE SPLICING**

YANG LIXIA

SCHOOL OF PHYSICAL AND MATHEMATICAL SCIENCES

2019

**Combining single-molecule manipulation and peptide
nucleic acid binding studies for unraveling how RNA
structures regulate ribosomal frameshifting and
alternative splicing**

Yang Lixia

SCHOOL OF PHYSICAL AND MATHEMATICAL SCIENCES

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

2019

Statement of Originality

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

26 Dec 2018

.....

Date

Yang Lixia .

.....

Yang Lixia

Supervisor Declaration Statement

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

26 Dec, 2018

.....

Date

.....

Chen Gang

Authorship Attribution Statement

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

Chapter 2 is published as L.X. Yang, Z.S. Zhong, C.L. Tong, H. Jia, Y.R. Liu, G. Chen. Single-Molecule Mechanical Folding and Unfolding of RNA Hairpins: Effects of Single A-U to A·C Pair Substitutions and Single Proton Binding and Implications for mRNA Structure-Induced –1 Ribosomal Frameshifting. *Journal of the American Chemical Society* **140**, 8172-8184 (2018). DOI: 10.1021/jacs.8b02970.

The contributions of the co-authors are as follows:

- A/Prof Chen provided the initial project direction.
- I performed all the laboratory work. I also analyzed all the data.
- Dr Zhong, Dr Tong, Ms Jia, and Dr Liu helped construct the plasmids.
- I wrote the manuscript drafts. The manuscript was revised together with A/Prof Chen.

Chapter 4 is published as J.Z. Tan,= L.X. Yang,= A.A.L. Ong, J.H. Shi, Z.S. Zhong, M.L. Lye, S.Y. Liu, J. Lisowiec-Wachnicka, R. Kierzek, X. Roca, G. Chen. A Disease-Causing Intronic Point Mutation C19G Alters Tau Exon 10 Splicing via RNA Secondary Structure Rearrangement. *Biochemistry* **58**, 1565-1578 (2019). (= These authors contributed equally to this work). DOI: 10.1021/acs.biochem.9b00001.

The contributions of the co-authors are as follows:

- I performed all the single-molecule pulling experiments, the PAGE gel study, the thermal melting experiments and all data analysis.
- Dr Tan performed the cell culture splicing assay.
- Dr Tan and I prepared the manuscript. A/Prof and I revised the manuscript.
- Mr Ong performed the fluorescence measurement.
- Mr Shi, Dr Zhong, Ms Lye and Ms Liu assisted in the early stage of the work.
- Prof Lisowiec-Wachnicka, Prof Kierzek, and Prof Roca helped providing plasmids and materials.

26 Dec 2018

Yang Lixia .

.....

.....

Date

Yang Lixia

Abstract

RNA structures are involved in regulating many biological activities such as programmed ribosomal frameshifting and pre-mRNA alternative splicing. Studying the structural features, mechanical stabilities, unfolding and folding dynamics of native RNAs and their mutants, as well as the interaction with ligands such as peptide nucleic acids can help facilitate a deep and thorough understanding of the RNAs and the related biological functions. Here, we combined single-molecule manipulation using optical tweezers and peptide nucleic acids binding studies to unravel the biophysical properties of model RNA hairpins and disease-related RNA structures, and their correlations with minus-one ribosomal frameshifting and pre-mRNA alternative splicing activities. Minus-one ribosomal frameshifting produces multiple functional proteins from one viral mRNA through ribosomal backward slippage by one nucleotide. The alternative splicing of exons in a pre-mRNA by spliceosome produces different protein isoforms through the inclusion or exclusion of exons.

Project 1. A wobble A·C pair can be protonated at near physiological pH to form a more stable wobble A⁺·C pair. Here, we constructed an RNA hairpin (rHP) and three mutants with one A-U base pair substituted with an A·C mismatch on the top (near the loop, U22C), middle (U25C) and bottom (U29C) positions of the stem, respectively. Our results on single-molecule mechanical (un)folding using optical tweezers reveal the destabilization effect of A-U to A·C pair substitution, and protonation-dependent enhancement of mechanical stability facilitated

through an increased folding rate, or decreased unfolding rate, or both. Our data show that protonation may occur rapidly upon the formation of apparent mechanical folding transition state. Furthermore, we measured the bulk -1 ribosomal frameshifting efficiencies of the hairpins by a cell-free translation assay. For the mRNA hairpins studied, -1 frameshifting efficiency correlates with mechanical unfolding force at equilibrium and folding rate at around 15 pN. U29C has a frameshifting efficiency similar to that of rHP ($\sim 2\%$). Accordingly, the bottom 2-4 base pairs of U29C may not form under a stretching force at pH 7.3, which is consistent with the fact that the bottom base pairs of the hairpins may be disrupted by ribosome at the slippery site. U22C and U25C have a similar frameshifting efficiency ($\sim 1\%$), indicating that both unfolding and folding rates of an mRNA hairpin in a crowded environment may affect frameshifting. Our data indicate that mechanical (un)folding of RNA hairpins may mimic how mRNAs unfold and fold in the presence of translating ribosomes.

Project 2. Base triples in RNA pseudoknots are crucial in maintaining RNA structure, stability, and biological functions. Minor-groove base triples formed between stem 1 and loop 2 of SRV-1 mRNA frameshifting pseudoknot are essential in determining the thermodynamics and mechanical properties of the pseudoknot as well as stimulating the minus-one ribosomal frameshifting. How tertiary base triple formation affects the stability of stem 1 and stem 2 and thus ribosomal frameshifting efficiency is not well understood. We designed a PNA that is expected to be able to invade stem 1 of the SRV-1 pseudoknot to mimic stem 1 unwinding process by a translating ribosome. In addition, we designed a

PNA for invading stem 2 in SRV-1 pseudoknot. Our non-denaturing polyacrylamide gel electrophoresis data revealed that a single mutation in loop 2 disrupting the base triple formation in SRV-1 pseudoknot results in tighter binding by both PNAs. Thus, tertiary stem 1-loop 2 base triple interactions in SRV-1 pseudoknot can stabilize the secondary structural components stem 1 and stem 2. We observed that invasion efficiencies of both PNAs positively correlated with the mechanical stability and the frameshifting efficiency of the SRV-1 pseudoknot, and inversely correlated with their unfolding rates. Our work suggests that both stem 1 and stem 2 stabilities in a pseudoknot structure are important for stimulating ribosomal frameshifting.

Project 3. Frontotemporal dementia with Parkinsonism linked to chromosome-17 (FTDP-17) is caused by mutations in the microtubule-associated protein tau (*MAPT*) gene. Alternative splicing of the *MAPT* cassette exon 10 produces tau isoforms with 4 microtubule-binding repeat domains (4R) upon exon inclusion, or 3 repeats (3R) upon exon skipping. In human neurons, deviations from the physiological 4R:3R ratio of ~1:1 lead to FTDP-17. Certain FTDP-17-associated mutations affect a regulatory hairpin that sequesters the exon 10 5' splice site (5'ss). These mutations tend to increase the 4R:3R ratio by destabilizing the hairpin, thereby improving 5'ss recognition by U1 small nuclear ribonucleoprotein. Interestingly, a single C-to-G mutation at the 19th nucleotide in intron 10 (C19G or +19G) reduces exon 10 inclusion significantly from 56% to 1%, despite the disruption of a G-C base pair in the bottom stem of the hairpin. Here, we show by biophysical characterization including thermal melting and

single-molecule mechanical unfolding using optical tweezers, that the +19G mutation alters the structure of the bottom stem, resulting in the formation of a new bottom stem with enhanced stability. Some of the disease-causing point mutations in the top stem may also result in local structural rearrangement within the top stem. We exploited the effect of the +19G mutation on the known FTDP-17-linked hairpin top stem-destabilizing mutants. The cell culture alternative splicing patterns of a series of minigenes containing the disease-causing mutations and artificial mutations reveal that the splicing activities of the mutants with destabilizing mutations on the top stem can be compensated in position-dependent manner by a +19G mutation in the bottom stem. A local 5-base pair stem surrounding the 5'ss site (involving residues -2 to +3 and +12 to +15) is critical in determining the alternative splicing pattern, and a point mutation within this local stem causes an increased exon 10 inclusion from 56% to above 95%, with the splicing level not compensated by a stabilizing mutation such as +19G in the bottom stem. We observed an excellent correlation between exon 10 inclusion level and the mechanical unfolding rate at 10 pN, consistent with the fact that the unfolding of the splice site hairpins is aided by helicases. Thus, mechanical unfolding directly reveals the structural changes upon single-nucleotide mutations and can mimic how the splice site hairpins are unfolded in the cell for the recognition by U1 snRNA.

Acknowledgements

I fulfilled four years of study and research at Nanyang Technological University, Singapore. At this very moment, I have so much gratitude to give for all the reliable support, inspiring guidance, and enthusiastic encouragement that accompany with me along this wonderful journey.

My greatest gratitude is devoted to my supervisor Professor Chen Gang. He's been instructive and supportive on every aspect of my study and research. His kindness to all the students, his deep understanding and insights towards research projects and his keenness to create and discover have set a very good model for me to learn from.

I'm truly grateful to all the members in our research group. I give special acknowledgement to Ms. Desiree-Faye Toh Kaixin for her synthesizing the PNAs used in PNA invasion project and her coaching in non-denaturing polyacrylamide gel electrophoresis. Special appreciation also goes to Dr. Tan Jiazi for his work on cell culture splicing assay. I particularly thank our former group member Dr. Zhong Zhensheng for teaching me the single-molecule mechanical manipulation using optical tweezers.

The endless love and support from my parents have got me through the ups and downs of my life. My parents always respect my choices and decisions, encourage me to explore, and believe in me. I can never thank them enough. All these years apart, I wish I'd be there for them more often just like they've been here for me the whole time. Finally, I thank my beloved fiancé and best friend,

Dr. Wang Shaomeng, who's been loving, gentle, and wise accompanying me.
I'm so lucky to be with him and I long for the future yet to be explored together
with him.

Table of Contents

Abstract	1
Acknowledgements	5
Table of Contents	7
List of Abbreviations	11
Chapter 1 Introduction	13
1.1 Ribonucleic Acid (RNA) Structures and Functions	13
1.2 Peptide Nucleic Acid (PNA)	17
1.3 Minus-one Programmed Ribosomal Frameshifting	19
1.4 Alternative Splicing	23
1.5 Mechanical Properties of RNA Structures	24
1.6 Single-molecule Manipulation by Optical Tweezers	26
1.6.1 Principle of Optical Tweezers	26
1.6.2 MiniTweezers	29
1.6.3 The Experimental Setup	31
1.6.4 The Force Measurement and Calibration of MiniTweezers	34
1.6.5 The Distance Measurement and Calibration of MiniTweezers	36
1.6.6 The Measurement Modes and Data Analysis	38
References	44

Chapter 2 Single-Molecule Mechanical Folding and Unfolding of RNA Hairpins: Effects of A·C mismatch and Single Proton Binding and Implications for mRNA Structure-Induced –1 Ribosomal Frameshifting 54

2.1	Introduction	54
2.2	Materials and Methods	58
2.2.1	Thermal Melting Experiments	58
2.2.2	Optical Tweezers Experiments	59
2.2.3	<i>In vitro</i> Ribosomal Frameshifting Assay	61
2.3	Results and Discussion	62
2.3.1	Thermal Unfolding	62
2.3.2	Mechanical (Un)Folding Dynamics	67
2.3.3	Minus-one Ribosomal Frameshifting Efficiency Positively Correlates with Critical Force and Folding Rate of the mRNA Hairpins	82
2.4	Conclusions	86
	References	90

Chapter 3 Base Triple Formation between Loop 2 and Stem 1 in SRV-1 Pseudoknot Inhibits Antisense PNA Invasion of Stem 1 and Stem 2 98

3.1	Introduction	98
3.2	Materials and Methods	103
3.2.1	PNA Oligomer Synthesis	103
3.2.2	Non-denaturing Polyacrylamide Gel Electrophoresis	104
3.2.3	Data Analysis	105
3.3	Results and Discussion	107

3.3.1	Invasion of Stem 1 and Stem 2 in SRV-1 Pseudoknots	107
3.3.2	Correlations among Stem Invasion Efficiency, Mechanical Unfolding Force, Mechanical Unfolding Kinetics, and Frameshifting Efficiency	115
3.4	Conclusions	117
	References	118

Chapter 4 Single-Molecule Mechanical Folding and Unfolding of Disease-causing Tau pre-mRNA Exon 10 Splice Site Hairpins: Characterization of Secondary Structure Rearrangement and the Implications for Alternative Splicing

4.1	Introduction	122
4.2	Materials and Methods	126
4.2.1	RNA Secondary Structure Prediction	126
4.2.2	<i>MAPT</i> Exon 10 Hybrid Minigenes Construction and Splicing Assay in Cells	127
4.2.3	Thermal Melting Experiments	127
4.2.4	Non-denaturing Polyacrylamide Gel Electrophoresis	128
4.2.5	Single-molecule Optical Tweezers Experiments	129
4.3	Results and Discussion	130
4.3.1	Thermal Melting of Expanded 30-mer Splice Site Hairpin Constructs Reveals a Stabilization Effect of +19G Mutation	130
4.3.2	Cell Culture pre-mRNA Splicing Assay Shows that a Stabilizing Mutation in the Bottom Stem Results in Inhibition of the 5'ss Recognition on the Top Stem	134
4.3.3	Disease-causing Destabilizing Mutations on the Top Stem Rescue the Splicing of +19G Mutant	138

4.3.4	Correlations Between Splicing Activity and T_m and Predicted Folding Free Energy	140
4.3.5	Single-molecule RNA Hairpin Force Measurements Directly Reveal +19G Hairpin Structural Rearrangement	144
4.4	Conclusions	154
	References	156
Chapter 5	Conclusion and Future Work	161
	References	164
Appendixes		165
Appendix 1.	The parameters obtained from mechanical pulling experiments for rHP, U22C, U25C, and U29C at varied pH	165
Appendix 2.	The parameters obtained from mechanical pulling experiments for tau pre-mRNA splice site RNA hairpins of WT, +14U, +19G and +14U+19G	173
List of Publications		176

List of Abbreviations

ΔX	end-to-end extension change
$\Delta n^{\ddagger, \text{FU}}$	Number of nucleotides disrupted at unfolding transition state
$\Delta n^{\ddagger, \text{UF}}$	Number of nucleotides folded at folding transition state
3R	Tau isoforms with three microtubule-binding repeats
4R	Tau isoforms with four microtubule-binding repeats
5'ss	5' splice site
μM	micromolar
μm	micrometer
bp	base pair
-1 FS	Minus-one frameshifting
CCD	Charge-coupled device
DDX5	DEAD-box helicase 5
DNA	Deoxyribonucleic acid
dsDNA	double-stranded DNA
dsRNA	double-strand RNA
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetraacetic acid
EWLC	Extensible wormlike chain
F_c	Critical force
FEC	Force-extension curve
FLuc	Firefly luciferase
FTDP-17	Frontotemporal dementia and parkinsonism linked to chromosome 17
G_o	Free energy for unfolding reaction at zero force
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	High-performance liquid chromatography
I	Intermediate hairpin structure
LED	Light-emitting diode
M	Molar
MALDI-TOF	Matrix-assisted laser desorption/ionization-Time of flight

<i>MAPT</i>	Microtubule-associated protein tau
MES	2-(N-morpholino)ethanesulfonic acid
ms	millisecond
N	Native hairpin
nm	nanometer
NMR	Nuclear magnetic resonance spectroscopy
npbs	non-polarizing beam splitter
nt	nucleotide
OBJ	Objective lenses
PAGE	Polyacrylamide gel electrophoresis
pbs	polarizing beam splitter
PCR	Polymerase chain reaction
PD	Photodiode detector
pN	piconewton
PNA	Peptide nucleic acid
PSD	Position sensitive detector
RLuc	Renilla luciferase
RNA	Ribonucleic acid
RP-HPLC	Reverse phase high-performance liquid chromatography
RT-PCR	Reverse transcription polymerase chain reaction
SFPQ	Splicing factor proline- and glutamine-rich
ssRNA	single-strand RNA
sulfo-NHS	N-hydroxysulfosuccinimide
T_m	Melting temperature
Tris	Tris(hydroxymethyl)aminomethane
tRNA	transfer RNA
U	Unfolded single-stranded random coil
U1 snRNA	U1 small nuclear RNA
U1 snRNP	U1 small nuclear ribonucleoprotein
W_{Fc}	Work for unfolding reaction at equilibrium at critical force

Chapter 1

Introduction

1.1 Ribonucleic Acid (RNA) Structures and Functions

RNAs are involved in many biological functions such as viral replication, RNA transcription, protein synthesis, ribosomal frameshifting, and pre-mRNA alternative splicing.¹⁻⁶ RNA is composed of ribonucleotides covalently linked by phosphodiester bonds. The monomer nucleotide consists of a five-carbon ribose sugar, with carbons numbered 1' to 5', a nitrogenous base of adenine (A), uracil (U), cytosine (C) or guanine (G) attached to the 1'-carbon through N-glycosidic bond, and a phosphate group connecting the neighbor nucleotides by 5' to 3' phosphodiester bonds (**Figure 1.1A**).

A and G pair with U and C to form stable Watson–Crick pairs aided by two and three hydrogen bonds, respectively (**Figure 1.1B**). G-C pair, with an extra hydrogen bond forming, is apparently more stable than A-U pair. Noncanonical base pairs or mismatches can also have hydrogen bonds formed, contributing to the stability and diversity of the RNA structures and RNA-mediated recognition. Although most of the base mismatches, e.g. A·C, G·U, U·C, are evidently less stable than canonical base pairs following Watson–Crick pairing rules, they can actually provide a greater degree of flexibility in RNA conformation, which increases the adaptability to accommodate interaction partners such as proton, ions, and proteins.⁷⁻¹²

Figure 1.1. Chemical structures of RNA. (A) Chemical structure of RNA nucleotide. The nucleobase shown is uracil (U). (B) Chemical structures of RNA base pairs A-U and G-C. Hydrogen bonds formed are indicated with dashed lines.

Proper folding of RNAs is crucial for carrying out many biologically important functions. RNAs adopt various secondary structures and more complex tertiary structures through base pairing, hydrogen bonding, stacking, conformational entropy releasing, counterion binding, and solvation effect. RNA secondary structure consists of stems, bulges, branches, junctions, and various loops. The interactions of these secondary structural components result in the formation of the RNA tertiary structure.^{1,13,14}

RNA hairpin is composed of single-stranded loop closed with A-form double helix stem stabilized by canonical base pairing as well as stacking.¹³ RNA double helix typically adopts A-form with 2.63 Å rise per nucleotide and 11 base pairs per turn. RNA double helix has a deep yet narrow major groove with Hoogsteen edge exposed to solution, and a shallow yet wide minor groove with ribose sugar edge exposed. Both major and minor groove provide access for ligand binding

and tertiary interactions. The ribose ring mainly adopts two conformations of C3'-endo pucker (3'-carbon out of the ribose ring plane on the base side) and C2'-endo pucker (2'-carbon out of the ribose ring plane on the base side). In RNA helix, ribose adopts C3'-endo pucker while in unpaired single strand ribose rapidly switch between the C3'-endo and C2'-endo conformations. The secondary structures and thermodynamics of short RNAs can be well predicted by a nearest-neighbor model with free energy minimization.¹⁵⁻¹⁷ As one of the most important secondary structure motifs, RNA hairpins often mediate RNA tertiary structure folding, serve as target binding domains for regulatory proteins, affect alternative splicing, and regulate protein translation in ribosomal frameshifting.

Unpaired single strand could interact with a Watson–Crick double helix in the major or minor groove through hydrogen bonds, forming tertiary structures such as base triples.¹⁸⁻²¹ In the major-groove triple helix, Hoogsteen hydrogen bonds form between a pyrimidine-rich third strand and a double helix in the major groove, for example, U·A-U and C⁺·G-C (with Hoogsteen hydrogen bonds marked as ‘·’, and Watson–Crick hydrogen bonds marked as ‘-’), with no significant distortion of the A-form conformation. In the minor-groove base triples, adenosine or other nucleotide interact with canonical base pair through hydrogen bonds involving 2'-OH of the ribose in the minor groove such as the A-minor motif.²²

RNA base triples exist in many biologically important pseudoknots.^{19-21,23} RNA pseudoknot usually composes of two helical stems and two single strand loops. The unpaired loop can interact with the stem region and form tertiary base

triples.²⁴ Such tertiary interactions help stabilize the pseudoknot structure and further modulate biological functions.²⁵⁻²⁷

RNAs fold into complex three-dimensional structures to carry out their important functions involving interactions with other molecules.^{1,13,28} Studies have revealed the three-dimensional (3D) structures of RNAs at high resolution, e.g., highly stable RNA hairpins analyzed using NMR,^{29,30} a SRV-1 RNA pseudoknot determined by NMR,²³ the crystal structure of tRNA,^{31,32} (**Figure 1.2**) and the crystal structure of rRNAs in ribosome,³³ as well as the interactions with mRNA and tRNA.³⁴ The 3D structures of RNAs can be characterized through computational simulation in combination with experimental constraints. Relatively stable 3D structures of RNAs have been predicted by many modeling methods developed,³⁵⁻⁴⁰ for example, small RNA hairpins,⁴¹ RNA pseudoknots,⁴² tRNA,⁴³ ribosomal RNAs,⁴⁴ and RNA G-quadruplex.⁴⁵

Figure 1.2. Examples of three-dimensional structures of (A) RNA hairpin (PDB ID: 2KOC),³⁰ (B) SRV-1 RNA pseudoknot (PDB ID: 1E95),²³ and (C) tRNA (PDB ID: 1EVV).³²

Chemical modifications of RNA introduce new groups to nucleotides, thus changing the structure, stability, nucleases resistance, folding pathway, and binding to other functional DNA, RNA, or proteins.⁴⁶ Chemical modifications of RNA mainly include nucleobase modifications, ribose sugar modifications, and phosphodiester backbone modifications. The nucleobase modifications can increase or decrease duplex thermal stability and reduce immune stimulation. For example, pseudouridine (5-ribosyluracil) is a modified uracil with the linking C–N bond between uracil and ribose changed to a C–C bond.⁴⁷ Sugar modifications can change the sugar conformation, alter its puckering preferences, and the binding affinity and duplex conformation of RNA. 2'-O-methyl (2'-OMe), 2'-O-methoxyethyl (2'-OMOE), and locked nucleic acid (LNA) substitutions are important and commonly used modifications in sugar modifications. They can increase the binding affinity with complementary RNA sequences, reduce immunostimulatory properties of siRNAs, and increase nuclease stability.^{46,48} The phosphodiester backbone modifications introduce modifications to the phosphodiester linkage of RNA such as phosphorothioate linkage with nonbridging phosphate oxygen in phosphodiester linkage substituted with sulfur.⁴⁹

1.2 Peptide Nucleic Acid (PNA)

PNA was initially synthesized and studied by Nielsen et al. back in 1991 in mimicry of DNA.⁵⁰ The phosphodiester backbone of DNA or RNA is modified with N-(2-aminoethyl)-glycine units linked by peptide bonds. PNA has its nucleobases of adenine (A), uracil (U), cytosine (C), guanine (G) as well as thymine (T) attached to the peptide-like acyclic, achiral and neutral backbone

through methylene carbonyl linkages (**Figure 1.3A**).⁵⁰⁻⁵² Further modifications could be incorporated into the backbone^{53,54} and bases⁵⁵⁻⁵⁷ of the PNA to enhance the binding selectivity, affinity, nuclease resistance, cellular uptake, etc.

PNA hybridizes to complementary RNA as well as DNA and PNA through Watson–Crick base pairing (**Figure 1.3B**).^{52,58} With the neutral backbone, PNA-RNA bound complexes are more stable than RNA-RNA duplexes.⁵⁹ PNA has its N-terminus and C-terminus, and binds to complementary strands in both parallel (N-terminus of PNA aligned with 5' end of RNA) and antiparallel (C-terminus of PNA aligned with 5' end of RNA) orientations, with antiparallel orientations more favoured. PNA binds to RNA sequence-specifically and a single mismatch in PNA-binding complex generally destabilizes the complex structure drastically.^{52,60-62}

Figure 1.3. Chemical structures of PNA. (A) Chemical structure of PNA monomer. The base could be adenine, cytosine, guanine, thymine, uracil or other modified bases. (B) Chemical structure of PNA-RNA Watson–Crick U-A and C-G base pairs.

PNA serves as a very useful ligand targeting RNA structures with high selectivity and affinity. Studies have shown that duplex forming PNA recognizes RNA component of human telomerase and inhibits telomerase activity *in vitro*⁶³ and in cell⁶⁴, binds to pre-mRNA to alter splicing pattern *in vitro*,⁶⁵ and inhibits translation *in vitro* and in cell through binding to the terminus of the 5'-untranslated region of mRNA and blocking binding of the translation machinery,⁶² targeting the coding region of mRNA to arrest translation elongation,^{66,67} and invading ribosomal RNA helix to disable ribosome activity.⁶⁸

1.3 Minus-one Programmed Ribosomal Frameshifting

Minus-one programmed ribosomal frameshifting (−1 FS) is found in viruses, bacteria, and higher eukaryotes as a gene-expression regulation mechanism by which multiple proteins are produced using one mRNA or mRNA half-life.⁶⁹⁻⁷² During the minus-one ribosomal frameshifting, the ribosome moving along a mRNA would be stalled and pause over a slippery sequence in a metastable conformational state when encountered by a downstream RNA structure, such as hairpin,^{73,74} pseudoknot,^{75,76} and G-quadruplex.^{77,78} The slippery sequence and the downstream structure is often separated by a single-stranded spacer of 2–8 nucleotides. The mechanical tension generated may cause the ribosome at the slippery site to slip one nucleotide backward so as to release the tension and continue translation in a new reading frame (**Figure 1.4**).^{69,79} The slippery sequence in viruses and eukaryotes normally adopts the sequences of X XXY YYZ (codons in the 0 frame are underlined), of which X could be any nucleotide,

Y is either A or U, and Z could be A, C or U, e.g., U UUU UUA and G GGA AAC.

Figure 1.4. Schematic of minus-one ribosomal frameshifting. (Left) Ribosome paused over the slippery site of mRNA at 0 frame. (Right) Ribosome slips one nucleotide backward and continues translation at -1 frame. The slippery sequence (X XXY YYZ) is shown in red. The spacer with varied length in shown in green.

During translation, ribosome binds to mRNA for the initiation of translation with the reading frame defined by the start codon (AUG). Multiple ribosomes may have access to the same mRNA forming polysomes. When ribosome encounters a downstream stem-loop or pseudoknot structure, the mRNA entrance channel of ribosome is blocked due to the fact that the size of a stem (2 nm in diameter) or pseudoknot structure is larger than the diameter the mRNA entrance channel (1.5 nm) so that the ribosome would be stalled, and the rotation of the ribosomal subunits hindered (metastable conformational state). When the ribosome is in a metastable conformational state, both backward slippage (-1 FS, -1 frame) and reverse rotation upon simultaneous downstream RNA structure unwinding (no FS, 0 frame) are possible. During the -1 FS, the ribosome, spanning around 30 nt of mRNA,^{80,81} will interact with and unwind the 5'-end of the downstream

mRNA structures at the mRNA entrance tunnel of the small subunit of ribosome.⁸⁰⁻⁸² The unfolding of the RNA structure could be due to the ribosome biasing the thermal fluctuations towards the unfolded state or the mechanical force generated during ribosome translocation.⁸³ The ribosomal proteins S3, S4 and S5 in prokaryotic ribosomes contribute to the mRNA helicase activity in unwinding downstream RNA structures. The force exerted by ribosome during translation is about 12-20 pN.^{84,85} Thus, the mRNA structures may be pre-melted by ribosome at the mRNA entry site when ribosome is initially paused at the slippery site, followed by consequent unzipping of the rest of the hairpin structure. This is comparable with the case of mechanical pulling where force is directly applied to the 5'-end and 3'-end of the RNA structure, unzipping the RNA structure sequentially.

The minus-one programmed ribosomal frameshifting could be ranging from as low as 1% to as high as 80%.⁷⁰ Previous studies have shown that the mechanical stability of the frameshift site mRNA structure positively correlates with minus-one ribosomal frameshifting efficiency.^{26,74,86,87}

RNA structures of hairpins, pseudoknots, and G-quadruplexes can all effectively stimulate ribosomal frameshifting. Generally, the frameshifting efficiency caused by RNA hairpins is lower than that induced by pseudoknots and G-quadruplexes, possibly because RNA hairpins are less stable than pseudoknots and G-quadruplexes.^{24,88,89} Hairpins in nature can stimulate high levels of frameshifting, though.⁹⁰⁻⁹² Frameshift-inducing hairpins of similar base pair

composition as the SRV-1 gag-pro pseudoknot have been also reported to sufficiently substitute the pseudoknot in stimulating frameshifting.⁷³

Studies have revealed the mechanical stabilities are positively correlated with minus-one ribosomal frameshifting efficiency for RNA hairpin, pseudoknot, and G-quadruplex.^{26,74,86} It has been widely accepted that the RNA structures serve as mechanical roadblocks for pausing ribosome over slippery site and stimulating backward slippage, with the mechanical stabilities of such structures critical to -1 FS.⁹³ Both unfolding and folding rates have also been found correlated with -1 FS.^{26,74,94} Other factors such as slippery sequences, spacer length, and trans-acting elements can also modulate frameshifting.^{70,95-100}

Other than -1 ribosomal frameshifting, there also exists +1 ribosomal frameshifting and unusual -2 ribosomal frameshifting where the ribosome moves forward 1 nt and continues translation in the +1 reading frame, or moves back 2 nt and continues translation in the -2 reading frame. -2 FS generates a product with an additional amino acid compared to +1 FS. +1 FS is widely involved in the regulation of release factor 2, antizyme, and cellular polyamine levels.¹⁰¹⁻¹⁰³ +1 FS was also found in influenza A virus at a C UUU CGU slippery sequence.¹⁰⁴ Unusual -2 FS was found in porcine reproductive and respiratory syndrome virus (PRRSV) at a conserved G GUU UUU slippery sequence followed by an unstructured motif 11 nt downstream, and in the HIV gag/pol signal at a U UUU UUA slippery sequence with a mutated spacer shortened by 1-2 nt.^{100,105,106} Studies have found that spacer length, downstream mRNA structure, and trans-acting protein can impact the efficient -2 FS, while slippery sequence is less

important for determining -2 FS.^{100,105,106} The mechanisms of $+1$ or -2 ribosomal frameshifting are more complicated and may vary from case to case.

1.4 Alternative Splicing

Pre-mRNA alternative splicing in eukaryotes allows the exons and introns to be excluded or included giving rise to multiple mRNAs and thus a variety of protein isoforms with diverse biological functions (**Figure 1.5**).^{107,108} During alternative splicing, a certain exon in the pre-mRNA sequence can be either included or excluded in the transcript sequences aided by the spliceosome. Alternative splicing may be modulated by *cis*-regulatory element of RNA secondary structures at the splice site. Alternative splicing may also be regulated by many *trans*-acting factors, e.g., RNA-binding proteins that recognize *cis*-factors and facilitate the unwinding of the RNA secondary structures. Changes in exon splicing could cause human diseases such as cancers,^{109,110} and frontotemporal dementia with Parkinsonism linked to chromosome-17 (FTDP-17).¹¹¹⁻¹¹⁴

Figure 1.5. Schematic of pre-mRNA alternative splicing. During alternative splicing, a certain exon in the pre-mRNA sequence can be either included or excluded to generate different mature mRNA sequences and protein isoforms.

Mutations in the *cis*-regulatory RNA structures at the splice site can alter pre-mRNA splicing, protein sequence, and protein structure.^{115,116} For example, in FTDP-17, studies revealed a regulatory hairpin in microtubule-associated tau pre-mRNA that sequesters the exon 10 5' splice site,¹¹⁵⁻¹¹⁸ and mutations within the hairpin region can result in the disruption a base pair, and tend to increase exon 10 inclusion.¹¹⁵⁻¹¹⁹ In the meanwhile, however, some mutation causes a significant decrease in exon 10 inclusion.¹²⁰ Force is found to be involved in pre-mRNA splicing by spliceosome.^{108,121} The mechanical stabilities of the *cis*-regulatory RNA structures are critical in regulating alternative splicing. The regulatory elements and the mechanism of splicing need to be investigated for facilitating the understanding of splicing process, and the diagnostics and treatment of aberrant-splicing-induced diseases.

1.5 Mechanical Properties of RNA Structures

RNAs adopt specific secondary and tertiary structures to perform biological activities. Apparently, the way RNA folds and unfolds and the folding and unfolding kinetics are critical for understanding the structural and catalytic functions of the RNA. In many cellular activities, force is extensively involved in RNA-regulating processes such as ribosome unwinding mRNA structure during translation,^{83,122} transcription by RNA-dependent RNA polymerase,¹²³⁻¹²⁵ and pre-mRNA splicing by spliceosome.^{108,121}

The thermodynamics of RNA secondary structures can be characterized by bulk UV thermal melting studies in equilibrium and well predicted by a nearest-neighbor model.¹⁵⁻¹⁷ However, RNAs in the cell may not behave in equilibrium

and proteins may affect RNA structures and dynamics significantly. With RNA-dependent enzymes and RNA binding proteins present, the properties of RNAs *in vivo* could deviate away from those in thermodynamic equilibrium at near physiological conditions. Thus, a direct measurement of single-molecule conformational change and mechanical properties under mechanical force at near physiological temperatures and buffer conditions would provide additional insights into RNA function.

The structural transitions, mechanical stability, and unfolding and folding kinetics under tension can be well probed by applying mechanical force on individual RNA structures. Exerting force to the termini of RNA structures makes the end-to-end extension change serve as a well-defined reaction coordinate. Single-molecule force spectroscopy of RNA structures will provide with answers to questions such as at what force it unfolds, what conformation it adopts alongside unfolding and folding, how fast it folds and unfolds, etc. With force directly applied to the termini of RNA structures, mechanical unfolding can mimic the way ribosomes unwind mRNA structure with the help of helicase components exerting forces on the structure during translation. The force exerted by base pairing and stacking involved in stabilizing RNA structures and the force exerted by ribosome and spliceosome are in the pico-Newton (pN) range and the end-to-end distance during RNA unfolding falls in the nm to μm range.

A variety of single-molecules techniques have been developed for measuring the mechanical properties of RNA structures. For example, Atomic force microscopy (AFM) can be used to unfold RNA in near physiological temperature and solution

conditions.¹²⁶⁻¹²⁸ Single-molecule fluorescence resonance energy transfer (smFRET) can allow one to measure kinetics of RNA folding and conformational changes.¹²⁹⁻¹³² Nanopore force-spectroscopy has been assembled for unzipping and extracting kinetics of DNA using electric field.¹³³⁻¹³⁵ Magnetic tweezers have been used to probe the stability as well as kinetics of DNA G-quadruplexes and other structures.^{128,136} Optical tweezers have been extensively applied to study the mechanical properties RNA structures.^{25,26,74,137,138} Optical tweezers allows researchers to directly manipulate biomolecules with the size ranging from nm to μm by applying 0.1 to more than 100 pN force, with high time and spatial resolution.^{25,26,74,125,128,137-139}

1.6 Single-molecule Manipulation by Optical Tweezers

Optical tweezers were invented by Arthur Ashkin in 1986,¹⁴⁰ and Ashkin was awarded the Nobel Prize in Physics in 2018 for his contributions to the pioneering work in developing optical tweezers for applications in studying biological systems. Since then, optical tweezers have served as a very useful tool for probing the mechanical properties of protein, DNA, RNA, and other single micromolecules.^{25,74,137,138,141-143}

1.6.1 Principle of Optical Tweezers

Optical tweezers employ gradient force generated by highly focused laser beams with Gaussian distribution to form a stable potential well (optical trap) so that it can steadily capture and manipulate viruses, cells, dielectric particles and apply external forces on the single molecules of interest attached to the particles.¹⁴⁰

The key of the mechanism of optical tweezers is light momentum change. When a beam of laser passes through one transparent particle, most of the light is refracted. The momentum of the light changes after the refraction (**Figure 1.6A**). Due to the law of conservation of momentum, the particle would accordingly gain an equal momentum change with opposite direction. According to the formula of $\vec{F} = \frac{d\vec{p}}{dt}$, a force is applied on the particle in the same direction as the momentum change of the particle, or in the opposite direction as the momentum change of the light.

Figure 1.6. Mechanism of optical tweezers. (A) Light momentum change scheme. (B) Force applied on the particle in a uniform optical field. (C) Force applied on the particle in a gradient optical field.

In a uniform light field, as shown in **Figure 1.6B**, the force caused by the overall momentum changes would be in the same direction as the travelling direction of the laser light because of the uniform distribution of light intensity. The particle

would be pushed along the light travelling direction. In the case of a non-uniform or gradient optical field, however, the force generated would deviate towards the direction with maximum light intensity (shown in **Figure 1.6C**). In general, the transparent particle inside the optical trap would be driven to wherever the highest light intensity is. With that elaborated, we then use a highly focused Gaussian beam by microscope objective with high numerical aperture to create such a gradient optical field with the maximum light intensity located at the focus. A gradient force generated by focused Gaussian laser field will drag the particle towards the laser focus whenever the particle is off the trap center (**Figure 1.7**). Single-molecule manipulation can be achieved by the laser trap control. Apart from the gradient force generated, there's also scattering force caused by small fraction of laser reflection and absorption, pushing the bead away from the laser focus. With two counter-propagating laser beams focused and well-aligned, such scattering forces can be offset.

Figure 1.7. Gradient force generated by a highly focused Gaussian beam. When the bead is off the focus center of the focused Gaussian beam, it will be dragged towards the trap center by the gradient force.

1.6.2 MiniTweezers

MiniTweezers^{83,144-147} is an optical tweezers instrument with a compact optical and mechanical design that generates dual-laser optical traps with highly focused laser beams. It can exert mechanical force up to 150 pN with sub-piconewton resolution and manipulate single molecules in three dimensions while the end-to-end extensions are controlled and monitored by piezo motors and position sensitive detectors with sub-nanometer distance resolution and 1 millisecond time resolution.

The instrument layout of MiniTweezers is shown in **Figure 1.8**. Two diode laser beams with wavelength of 845 nm (Lumic LU0845M150-1G06F10A) are steered by fiber wigglers in orthogonal directions with the glass optical fiber bent near its tip. An external light lever amplification system is used to amplify piezo motion (controlling the optical trap position), in which around 5% of the laser light is reflected by a pellicle mirror, focused by a lens at a position sensitive detector (PSD) (UDT DL-10) and detected, amplifying the piezo movement to about 10 times. Most of the laser light pass through the pellicle, the polarizing beam splitters (pbs), the quarter-wave plates ($\lambda/4$ plate), and then got focused by 60X water-immersed microscope objective lenses (OBJ) (Olympus UPlanApo-60x-Water) at foci. Optical traps are generated and aligned at their common foci. The exiting beams are collected and collimated by the opposite objective lens, and directed to different detectors by polarizing beam splitters for force measurements. The position sensitive detector (PSD) measures forces in x and y direction while a photodiode detector (PD) measures force in z direction. The flow chamber is mounted in between the two objective lenses, and can move in

x, y and z directions driven by piezo motors (Thorlabs Z606) mounted together. CCD camera connected to a monitor is used to monitor and achieve real-time control. Background light for imaging is generated by a blue LED.

Figure 1.8. The Minitweezers layout. Here, each of the two counter-propagating fiber-coupled diode lasers (845 nm) individually passes through pellicle mirror, polarizing beam-splitter (pbs), $\lambda/4$ quarter-wave plate, water immersed 60x objective lens (OBJ), another water immersed 60x objective lens (OBJ), $\lambda/4$ quarter-wave plate, polarizing beam-splitter (pbs), another polarizing beam-splitter (pbs), relay lens, non-polarizing beam-splitter (npbs), and finally reach position-sensitive photo detector (PSD) and photodiode detector (PD). The chamber movement is driven by piezo motors. PSDs in blue color measure extension. PSDs in grey color measure forces in x and y directions. PD in cyan measure force in z direction. All related information about Minitweezers can be well found on the website of TweezersLAB (<http://tweezerslab.unipr.it/cgi-bin/mt/home.pl>). The MiniTweezers instrument is isolated with a soundproof box

and hung from the ceiling with a bungee cord from in a quiet room located at ground floor to minimize the environmental influences such as heat fluctuation and noise vibration.

1.6.3 The Experimental Setup

Biological macromolecules of interest can be attached to a polystyrene bead trapped by the optical traps to achieve manipulation. With one terminus of the target molecule mounted to a polystyrene bead or other transparent particle manipulated by optical tweezers, and the other terminus of the molecule attached to a glass slip, fixed base, or another polystyrene bead fixed by a micropipette or controlled by another optical trap, the target molecule can be manipulated by moving the optical trap.

In MiniTweezers, dual optical traps are well aligned together to control and exert mechanical force. During the experiment, molecule of interest is attached to one streptavidin-coated polystyrene bead, which is sucked onto a micropipette that is connected to a piezoelectric stage, and an anti-digoxigenin-coated polystyrene bead by linkers or handles (**Figure 1.9**). The handles can either be a dsDNA or a RNA/DNA hybrid.¹⁴⁸⁻¹⁵¹ RNA/DNA hybrid handles are chosen when studying RNA structures. The RNA/DNA hybrids serve as two double helix handles flanking the RNA structure, with the 5'-side handle (handle A) labelled with biotin and 3'-side handle (handle B) labelled with digoxigenin at the terminal ends respectively, so that the handle-RNA-handle structure can then be mounted to corresponding coated polystyrene beads. During the single-molecule optical tweezers experiment, handle B would be annealed to the anti-digoxigenin-coated

bead (about 3 to 5 μm in diameter) at room temperature for 10 min. The annealed products will flow inside the chamber and be trapped by the dual-laser optical traps. In the same time, handle A of the construct would attach to a streptavidin-coated polystyrene bead (about 1.7 μm in diameter) and be held onto a micropipette by suction. The force can then be loaded and unloaded through controlling the pipette or the optical traps.

Figure 1.9. Schematic of experimental setup using optical traps. The RNA structure of interest (shown in blue) is flanked by two DNA/RNA hybrid handles (shown in black). The construct is annealed to a streptavidin-coated bead held by a micropipette through non-covalent biotin-streptavidin interaction and an anti-digoxigenin-coated bead through non-covalent digoxigenin-anti-digoxigenin interaction manipulated by the optical traps. The optical traps are generated using two counter-propagating laser beams focused by 60x water-immersed objective lenses.

The constructs for the mechanical pulling are synthesized based on biological polymerase chain reaction (PCR) and transcription (**Figure 1.10**). Target DNA sequences are inserted into specific vectors to construct the recombinant plasmids. With properly designed primer, the ~1000-nucleotide (nt) single-strand RNA (ssRNA) molecules are synthesized through PCR using the recombinant plasmids as templates and subsequent *in vitro* transcription. dsDNA for handle A is synthesized by PCR and post-labelled with biotin through the incorporation of the biotin-modified dUTP at the 3'-end using T4 DNA polymerase. dsDNA for handle B is labelled using digoxigenin-labelled reverse primer at the 5'-end during PCR (see Figure 1.10 for the schematic). The ssRNA containing the target sequences flanked by ~500-nt sequences on either side, is then annealed with the dsDNAs for handle A and handle B to form RNA/DNA hybrid construct.

Figure 1.10. Schematic of RNA construct synthesis. Handle A: The plasmids containing the target sequences are used as templates for PCR amplification using primers HAF (the forward primer) and HAR (the reverse primer). The 3'-terminus of handle A is post labelled with biotin through biotin modified dUTP. Handle B: The plasmids containing the target sequences are used as templates for PCR amplification using primers HBF (the forward primer) and HBR-dig (the reverse primer with digoxigenin labelled at the 5'-end). The 5'-terminus of the handle B is labelled with digoxigenin, which is introduced through digoxigenin-

labelled reverse primer during PCR. RNA of interest: The plasmids containing the target sequences are used as templates for PCR amplification using primers HAF-Promotor (the forward primer with promotor sequence) and HAR (the reverse primer). *In vitro* transcription is performed to generate the RNA. Annealing: RNA molecule is then annealed together with labelled Handle A and Handle B to form the construct for pulling.

1.6.4 The Force Measurement and Calibration of MiniTweezers

With the dual-trap design, force in MiniTweezers is directly measured based on light momentum change with sub-piconewton resolution. When a bead is trapped in the common center of the optical traps, no external force is imposed. When the bead in traps is dragged by an external force, it will be off the center. The light refracted asymmetrically then exits the trap with a change in momentum, which is collected and collimated again by the opposite lens and reaches the photodetectors for measurement of the power and offset of this light to determine the change in the momentum flux. The external force is inferred by the momentum conservation $F = dP/dt$, and can be further calculated from $F = nW/c$, where n is the refraction index of the buffer, W is the light power and c is the light speed. The external force applied is calculated by the direct measurements of PSDs calibrated by sensitivity factors.

$$F_i = s_i D_i \quad (1.1)$$

Here, F_i is the external force, s_i is the sensitivity factors, and D_i is the electronic difference signals. Subscript i stands for x , y , or z direction.

The sensitivity factors calibration is performed for accurate force measurements. With a direct measurement of the light momentum change, the force is calibrated independent of bead size, shape, refraction index, location, and laser power fluctuation. Other methods for calibration of force sensors include fluid-drag force test based on Stokes' Law,¹⁵² and testing a certain dsDNA with known overstretching force in a specific buffer.¹⁵³

During the Stokes test, the calibration spherical bead with a known radius (r) and a very narrow size distribution moves in distilled water inside the chamber with a certain velocity (v) upon the fluid drag force from the moving of the fluid chamber. The drag force (F_d) will be proportional to the fluid viscosity and the bead radius according to Stokes' Law,

$$F_d = 6\pi\eta rv \quad (1.2)$$

where η is the dynamic viscosity of the fluid, which is water in our case.

With the calibration spherical bead held still by the traps (with the trap force against the drag force), the trap force loaded on the bead will be the same value with the Stokes drag force in opposite direction. The apparent bead radius can then be calculated and then compared to the known size test to determine if the force is well calibrated. If the calculated size deviates from the known size, the force sensitivities of the lasers in each direction will be rescaled. Typical errors from Stokes test are about 10%.

Measuring the overstretching force of a certain dsDNA is also useful when determining if the force is well calibrated. In this test, a certain dsDNA with 50% CG base pairs and 50% AT base pairs can be used with the stretching force

invariant in buffers with known salts concentrations. Typically, a lambda DNA (48,502 bp) has a stretching force between 65 and 70 pN in buffer that contains 500 mM NaCl, 10 mM Tris, 1 mM EDTA, pH 7.5.

1.6.5 The Distance Measurement and Calibration of MiniTweezers

The trap position is steered by a fiber wiggler coupled with high resolution piezo motor stage. The movement of the trap is amplified using the light lever system and measured by PSDs with sub-nanometer distance resolution with time resolution of 1 millisecond. The direct measurements from PSDs are the amplified signal and need to be calibrated to retrieve the real distances of traps by adjusting the light lever sensitivities per laser.

To calibrate the distance, changes in the light-lever readings and changes in the motorized-stage shaft encoders are compared. During the distance calibration, a bead is sucked onto the tip of the pipette, which is fixed to the motorized stage and moved a known distance by the motors. In the same time, the traps are also tracking the movement of the bead through constant-force feedback, as a result of which the trap positions are locked to the motor positions and move the same distance. The amplified signal directly measured from PSDs can then be calibrated to the real distance through adjusting the light lever sensitivities per laser.

During the pulling experiments, single-molecule tether forms between the bead fixed onto the pipette and another bead controlled using optical traps. The distance directly measured is the movement change of the optical traps. The bead inside the trap will, however, have a displacement relative to the trap center with

external force exerted. The extension of the molecule tether can be obtained by subtracting the bead displacement from the directly measured trap distance.

$$X_{tether} = X_{trap} - X_{bead} \quad (1.3)$$

Here X_{tether} is the extension of the molecule tether, X_{trap} is the directly measured trap distance, and X_{bead} is the bead displacement.

Optical trap is analogous to a harmonic oscillator as a spring obeying Hooke's Law (**Equation 1.4**) with force generated proportional to the displacement of the bead and dependent on trap stiffness.

$$F = \varepsilon X_{bead} \quad (1.4)$$

Here ε is the trap stiffness, analogous to a spring constant, and X_{bead} is the bead displacement.

The bead displacement X_{bead} can then be determined from the Hooke's Law with the force and trap stiffness directly measured. The trap stiffness is measured with the help of the dual-trap setup, where the forces and extensions of both optical traps are measured. During the trap stiffness measurement, the force as a sum effect of the two traps is maintained as constant by the electronic feedback. The molecule tether extension will not change with the sum force kept invariant. The individual traps, however, move with a known distance but in opposite directions so that the sum force will not be affected. For example, when one trap moves to the left, the other trap moves simultaneously to the right. Individual force changes will be generated accordingly due to the movement of each trap. The stiffness of each trap then can be determined from the distance moved and force change

measured for individual traps using Hooke's Law. The effective stiffness of the dual-trap system can then be obtained from the sum of the two stiffness values.

1.6.6 The Measurement Modes and Data Analysis

There're mainly three different measuring modes in single-molecule force-induced manipulation: force-ramp mode, constant-force mode, and passive mode. During mechanical (un)folding of the RNA structure, force is applied to the two termini of the target RNA in all measuring modes.

In force-ramp measurement, the force is loaded and unloaded at a constant speed (pN/s) with the displacement of the trapped bead increased or decreased at a constant speed (nm/s). The end-to-end extension increases upon RNA unfolding and decreases when RNA molecule refolds. Force-extension curves (FECs) will be monitored, of which transitions are featured by certain rupture forces and end-to-end extension changes.

The end-to-end extension change (ΔX) of RNA molecule during transition can be converted to the number of nucleotides unfolded during pulling by an extensible wormlike chain (EWLC) model (**Equation 1.5**) due to the semi-flexible properties of RNA with the diameter of the RNA structure considered.¹⁵⁴⁻

156

$$F = \frac{k_B T}{P} \left[\frac{1}{4 \left(1 - \frac{\Delta X}{L_c} + \frac{F}{K} \right)^2} - \frac{1}{4} + \frac{\Delta X}{L_c} - \frac{F}{K} \right] \quad (1.5)$$

Here F is the stretching force, k_B is Boltzmann's constant, T is the absolute temperature. P , L_c , K are the persistent length, contour length and stretch modulus

with the values of 1 nm, 0.59 nm/nt and 1500 pN for ssRNA, respectively.^{137,157-159} For the RNA/DNA hybrid handle, L_c is 0.3 nm/bp, while P and K are extracted from least-square fitting.¹³⁷ For ssDNA, the stretch modulus is 800 pN, the persistence length is 0.75 nm, and the contour length is 0.56 nm/bp.¹⁵³ The persistence length of ssDNA increases as the sodium ion concentration decreases while the elastic stretch modulus remained approximately constant.¹⁶⁰ ssRNA is similar to ssDNA in ion dependence, and becomes more flexible with increasing ion concentrations of Na^+ and Mg^{2+} .¹⁶¹

In constant-force mode, the force is maintained as constant through a feedback of 5 ms during which the averaged force in the 5-ms interval is compared to the desired force, and the displacement will be adjusted accordingly with the piezoelectric stage movement to compensate for any difference. During the constant-force experiment, the force is set to desired values so that the RNA molecules can hop between the unfolded and folded states and the end-to-end extension of the sample is monitored over time. The molecules will only hop in a narrow force range about 2 pN. Any higher or lower force out of this range would render the molecule unfolded or folded completely due to the tilted transition energy barrier. The constant-force experiment directly reveals the real-time dynamics of (un)folding of the RNA molecule and allows simultaneous unfolding and folding measurement. During transitions where large force range detection needed, the loaded force can jump between the desired lower and higher constant values to probe the unfolding and folding transitions separately. For example, when there's a large hysteresis between the unfolding and folding pathways of complex pseudoknot structures, force-jump method can be applied

to unfold the RNA structure at higher force and refold it at lower force repeatedly.^{138,162}

In constant-force mode, the end-to-end extension change (ΔX) is obtained from the extension trajectory at a certain force through the extension distributions fitted to Gaussian functions:¹⁶³

$$y = a_1 e^{-[(X-X^F)/c_1]^2} + a_2 e^{-[(X-X^U)/c_2]^2} \quad (1.6)$$

where X is the extension, y is the count for each binned extension; a_1/a_2 and c_1/c_2 are the amplitudes and width of the peaks, respectively; X^U and X^F are the averaged extensions at unfolded and folded state, respectively.

Extension change ΔX is calculated from the distance between the two peaks.

$$\Delta X = X^U - X^F \quad (1.7)$$

To obtain the standard deviation of extension change, the distribution of the difference between any two neighbor extensions (X_{diff}) is fitted to one Gaussian distribution:

$$y = a e^{-[(X_{\text{diff}} - b)/c]^2} \quad (1.8)$$

where y is the count for each binned X_{diff} ; a and c are the amplitudes and width of the peak, respectively; b is the average of extension X_{diff} . The standard deviation of ΔX then equals to $c/2$.

Unlike the force-ramp method, constant-force experiments directly unveil the force-dependent kinetics. The extensions monitored over time in constant-force mode are fitted with hidden Markov modelling¹⁶⁴ to obtain the lifetimes of each state with unbiased separation of noise from state-to-state transitions. The lifetimes fit a single exponential decay distribution $P(t)$.

$$P(t) = e^{-kt} \quad (1.9)$$

Here $P(t)$ is the cumulative probability that the unfolded or folded state remains over time t . k is the rate constant.

The lifetimes at several forces measured are fitted to generate apparent rate constants at a series of forces. For a simple two-state transition with first order kinetics, the force-dependent unfolding and folding rates are then fitted to the Bell model (**Equation 1.10**).¹⁶²

$$\ln k(F) = \ln k_0 \frac{F\Delta X^\ddagger}{k_B T} \quad (1.10)$$

Here F is the force, k_0 is the rate with no external force, $\Delta X^\ddagger$ is the distance between the transition state and unfolded or folded state. Bell's model assumes that the external force only lowers the energy barrier without affecting the barrier distance. It is limited at low to intermediate force regime when the $\Delta X^\ddagger$ is independent of force. The critical forces of the RNA hairpins studied can be extracted through fitting the data to the Bell's model. It is not accurate in obtaining the intrinsic transition rate k_0 (at zero force) using Bell's model.^{165,166}

The critical force F_c at which the unfolding and folding rates equal each other is obtained from the cross point of the Bell model fittings of the force-dependent unfolding and refolding rates. At critical force, the force-dependent entropic conformational free energy difference between the unfolded and the folded states at critical force ($\Delta\phi_{F_c}$)¹⁶⁷ can be calculated from

$$\Delta\phi_{F_c} = - \int_0^{F_c} \Delta X(f) df \quad (1.11)$$

$\Delta X(f)$ is the extension difference of the RNA between the unfolded and folded states at force f . During constant-force mode, $\Delta\phi_{F_c}$ can be obtained from the critical force F_c multiplying by end-to-end extension change under critical force (ΔX_c): $\Delta\phi_{F_c} = F_c \cdot \Delta X(F_c)$.

The free energy change at zero force (ΔG_0) can be obtained by subtracting the work of stretching ssRNA at equilibrium ($W_{\text{stretching}}$) from $\Delta\phi_{F_c}$.

$$\Delta G_0 = F_c \cdot \Delta X(F_c) - W_{\text{stretching}} \quad (1.12)$$

$W_{\text{stretching}}$ can be obtained from the integration of the EWLC curve from 0 pN to critical force.¹³⁷ The mechanical unfolding or folding transition state positions are converted from the distances between transition state and folded state ($\Delta X^{\ddagger, \text{FU}}$) or unfolded state ($\Delta X^{\ddagger, \text{UF}}$), respectively. The number of nucleotides unfolded in unfolding transition state and folded in folding transition state can be converted from $\Delta X^{\ddagger}$ of the unfolding and folding processes, respectively.

In passive mode, the position of the bead inside the trap is maintained during which the DNA/RNA handles are stretched and a constant tension on the molecule is maintained while the force of the molecule is monitored. The force of the molecule would decrease upon RNA unfolding. There is no electronic force feedback, achieving 1-ms time resolution suitable for measuring fast kinetics.¹⁶³ Similar to that in the constant force mode, the molecules will hop in a relative narrow range and the real-time unfolding and folding dynamics are revealed simultaneously in the passive-force experiment. Similarly, distributions of the force are fitted to Gaussian functions (**Equation 1.13**) for the folding and unfolding processes.¹⁶³

$$y = a_1 e^{-[(F-F^U)/c_1]^2} + a_2 e^{-[(F-F^F)/c_2]^2} \quad (1.13)$$

Here y is counts for each binned force; a_1/a_2 and c_1/c_2 are the amplitudes and width of the peaks, respectively; and F^U and F^F are the forces at unfolded and folded states, respectively.

The force trajectories in passive mode are also fit to hidden Markov modeling based algorithm¹⁶⁴ to obtain the lifetimes of the unfolded state and folded state. Further extractions, not to be repeated here, of the thermodynamics, mechanical critical forces as well as transition state positions are the same with those in the constant-force mode.

References

- (1) I. Tinoco, Jr., C. Bustamante, How RNA folds. *J Mol Biol* **1999**, *293*, 271-281.
- (2) P. Nissen, J. Hansen, N. Ban, P. B. Moore, T. A. Steitz, The structural basis of ribosome activity in peptide bond synthesis. *Science* **2000**, *289*, 920-930.
- (3) P. G. Higgs, RNA secondary structure: physical and computational aspects. *Q Rev Biophys* **2000**, *33*, 199-253.
- (4) T. R. Cech, J. A. Steitz, The Noncoding RNA Revolution-Trashing Old Rules to Forge New Ones. *Cell* **2014**, *157*, 77-94.
- (5) M. L. Anko, K. M. Neugebauer, Long noncoding RNAs add another layer to pre-mRNA splicing regulation. *Mol Cell* **2010**, *39*, 833-834.
- (6) N. Romero-Barrios, M. F. Legascue, M. Benhamed, F. Ariel, M. Crespi, Splicing regulation by long noncoding RNAs. *Nucleic Acids Res* **2018**, *46*, 2169-2184.
- (7) X. Roca, M. Akerman, H. Gaus, A. Berdeja, C. F. Bennett, A. R. Krainer, Widespread recognition of 5' splice sites by noncanonical base-pairing to U1 snRNA involving bulged nucleotides. *Genes Dev* **2012**, *26*, 1098-1109.
- (8) M. T. Record, Jr., M. L. Lohman, P. De Haseth, Ion effects on ligand-nucleic acid interactions. *J Mol Biol* **1976**, *107*, 145-158.
- (9) E. M. Moody, J. T. J. Lecomte, P. C. Bevilacqua, Linkage between proton binding and folding in RNA: A thermodynamic framework and its experimental application for investigating pK(a) shifting. *Rna-a Publication of the Rna Society* **2005**, *11*, 157-172.
- (10) T. E. Edwards, T. M. Okonogi, S. T. Sigurdsson, Investigation of RNA-protein and RNA-metal ion interactions by electron paramagnetic resonance spectroscopy. The HIV TAR-Tat motif. *Chem Biol* **2002**, *9*, 699-706.
- (11) D. E. Draper, A guide to ions and RNA structure. *RNA* **2004**, *10*, 335-343.
- (12) R. Yamagami, J. L. Bingaman, E. A. Frankel, P. C. Bevilacqua, Cellular conditions of weakly chelated magnesium ions strongly promote RNA stability and catalysis. *Nat Commun* **2018**, *9*, 2149.
- (13) N. B. Leontis, A. Lescoute, E. Westhof, The building blocks and motifs of RNA architecture. *Curr Opin Struct Biol* **2006**, *16*, 279-287.
- (14) Tinoco, Nucleic acid structures, energetics, and dynamics. *J Phys Chem* **1996**, *100*, 12.
- (15) J. L. Chen, A. L. Dishler, S. D. Kennedy, I. Yildirim, B. Liu, D. H. Turner, M. J. Serra, Testing the Nearest Neighbor Model for Canonical RNA Base Pairs: Revision of GU Parameters. *Biochemistry* **2012**, *51*, 3508-3522.
- (16) T. Xia, J. SantaLucia, Jr., M. E. Burkard, R. Kierzek, S. J. Schroeder, X. Jiao, C. Cox, D. H. Turner, Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry* **1998**, *37*, 14719-14735.
- (17) J. Zuber, B. J. Cabral, I. McFadyen, D. M. Mauger, D. H. Mathews, Analysis of RNA Nearest Neighbor Parameters Reveals Interdependencies and Quantifies the Uncertainty in RNA Secondary Structure Prediction. *RNA* **2018**.
- (18) G. Devi, Y. Zhou, Z. S. Zhong, D. F. K. Toh, G. Chen, RNA triplexes: from structural principles to biological and biotech applications. *Wiley Interdiscip Rev RNA* **2015**, *6*, 111-128.

- (19) C. A. Theimer, C. A. Blois, J. Feigon, Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Mol Cell* **2005**, *17*, 671-682.
- (20) L. Su, L. Q. Chen, M. Egli, J. M. Berger, A. Rich, Minor groove RNA triplex in the crystal structure of a ribosomal frameshifting viral pseudoknot. *Nat Struct Biol* **1999**, *6*, 285-292.
- (21) P. L. Nixon, A. Rangan, Y. G. Kim, A. Rich, D. W. Hoffman, M. Hennig, D. P. Giedroc, Solution structure of a luteoviral P1-P2 frameshifting mRNA pseudoknot. *J Mol Biol* **2002**, *322*, 621-633.
- (22) P. Nissen, J. A. Ippolito, N. Ban, P. B. Moore, T. A. Steitz, RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. *Proc Natl Acad Sci U S A* **2001**, *98*, 4899-4903.
- (23) P. J. Michiels, A. A. Versleijen, P. W. Verlaan, C. W. Pleij, C. W. Hilbers, H. A. Heus, Solution structure of the pseudoknot of SRV-1 RNA, involved in ribosomal frameshifting. *J Mol Biol* **2001**, *310*, 1109-1123.
- (24) D. P. Giedroc, C. A. Theimer, P. L. Nixon, Structure, stability and function of RNA pseudoknots involved in stimulating ribosomal frameshifting. *J Mol Biol* **2000**, *298*, 167-185.
- (25) G. Chen, K. Y. Chang, M. Y. Chou, C. Bustamante, I. Tinoco, Triplex structures in an RNA pseudoknot enhance mechanical stability and increase efficiency of -1 ribosomal frameshifting. *Proc Natl Acad Sci U S A* **2009**, *106*, 12706-12711.
- (26) Z. Zhong, L. Yang, H. Zhang, J. Shi, J. J. Vandana, D. T. Lam, R. C. Olsthoorn, L. Lu, G. Chen, Mechanical unfolding kinetics of the SRV-1 gag-pro mRNA pseudoknot: possible implications for -1 ribosomal frameshifting stimulation. *Sci Rep* **2016**, *6*, 39549.
- (27) M. Y. Chou, K. Y. Chang, An intermolecular RNA triplex provides insight into structural determinants for the pseudoknot stimulator of -1 ribosomal frameshifting. *Nucleic Acids Res* **2010**, *38*, 1676-1685.
- (28) R. A. Laskowski, J. M. Thornton, Understanding the molecular machinery of genetics through 3D structures. *Nat Rev Genet* **2008**, *9*, 141-151.
- (29) C. Cheong, G. Varani, I. Tinoco, Jr., Solution structure of an unusually stable RNA hairpin, 5'GGAC(UUCG)GUCC. *Nature* **1990**, *346*, 680-682.
- (30) S. Nozinovic, B. Furtig, H. R. Jonker, C. Richter, H. Schwalbe, High-resolution NMR structure of an RNA model system: the 14-mer cUUCGg tetraloop hairpin RNA. *Nucleic Acids Res* **2010**, *38*, 683-694.
- (31) H. Shi, P. B. Moore, The crystal structure of yeast phenylalanine tRNA at 1.93 Å resolution: a classic structure revisited. *RNA* **2000**, *6*, 1091-1105.
- (32) L. Jovine, S. Djordjevic, D. Rhodes, The crystal structure of yeast phenylalanine tRNA at 2.0 Å resolution: cleavage by Mg(2+) in 15-year old crystals. *J Mol Biol* **2000**, *301*, 401-414.
- (33) B. S. Schuwirth, M. A. Borovinskaya, C. W. Hau, W. Zhang, A. Vila-Sanjurjo, J. M. Holton, J. H. Cate, Structures of the bacterial ribosome at 3.5 Å resolution. *Science* **2005**, *310*, 827-834.
- (34) M. Selmer, C. M. Dunham, F. V. t. Murphy, A. Weixlbaumer, S. Petry, A. C. Kelley, J. R. Weir, V. Ramakrishnan, Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **2006**, *313*, 1935-1942.
- (35) S. E. McDowell, N. Spackova, J. Sponer, N. G. Walter, Molecular dynamics simulations of RNA: an in silico single molecule approach. *Biopolymers* **2007**, *85*, 169-184.

- (36) X. J. Lu, W. K. Olson, 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res* **2003**, *31*, 5108-5121.
- (37) R. Das, D. Baker, Automated de novo prediction of native-like RNA tertiary structures. *Proc Natl Acad Sci U S A* **2007**, *104*, 14664-14669.
- (38) Z. Xia, D. P. Gardner, R. R. Gutell, P. Ren, Coarse-grained model for simulation of RNA three-dimensional structures. *J Phys Chem B* **2010**, *114*, 13497-13506.
- (39) J. Sponer, G. Bussi, M. Krepl, P. Banas, S. Bottaro, R. A. Cunha, A. Gil-Ley, G. Pinamonti, S. Poblete, P. Jurecka, N. G. Walter, M. Otyepka, RNA Structural Dynamics As Captured by Molecular Simulations: A Comprehensive Overview. *Chem Rev* **2018**, *118*, 4177-4338.
- (40) M. Popena, M. Szachniuk, M. Antczak, K. J. Purzycka, P. Lukasiak, N. Bartol, J. Blazewicz, R. W. Adamiak, Automated 3D structure composition for large RNAs. *Nucleic Acids Res* **2012**, *40*, e112.
- (41) G. Zuo, W. Li, J. Zhang, J. Wang, W. Wang, Folding of a small RNA hairpin based on simulation with replica exchange molecular dynamics. *J Phys Chem B* **2010**, *114*, 5835-5839.
- (42) S. Cao, D. P. Giedroc, S. J. Chen, Predicting loop-helix tertiary structural contacts in RNA pseudoknots. *RNA* **2010**, *16*, 538-552.
- (43) P. Auffinger, E. Westhof, H-bond stability in the tRNA(Asp) anticodon hairpin: 3 ns of multiple molecular dynamics simulations. *Biophys J* **1996**, *71*, 940-954.
- (44) A. Malhotra, R. K. Tan, S. C. Harvey, Prediction of the three-dimensional structure of Escherichia coli 30S ribosomal subunit: a molecular mechanics approach. *Proc Natl Acad Sci U S A* **1990**, *87*, 1950-1954.
- (45) M. Havrila, P. Stadlbauer, P. Kuhrova, P. Banas, J. L. Mergny, M. Otyepka, J. Sponer, Structural dynamics of propeller loop: towards folding of RNA G-quadruplex. *Nucleic Acids Res* **2018**, *46*, 8754-8771.
- (46) G. F. Deleavey, M. J. Damha, Designing chemically modified oligonucleotides for targeted gene silencing. *Chem Biol* **2012**, *19*, 937-954.
- (47) M. Charette, M. W. Gray, Pseudouridine in RNA: What, where, how, and why. *lmb Life* **2000**, *49*, 341-351.
- (48) Y. S. Sanghvi, A status update of modified oligonucleotides for chemotherapeutics applications. *Curr Protoc Nucleic Acid Chem* **2011**, Chapter 4, Unit 4 1 1-22.
- (49) U. Flierl, T. L. Nero, B. Lim, J. F. Arthur, Y. Yao, S. M. Jung, E. Gitz, A. Y. Pollitt, M. T. K. Zaldivia, M. Jandrot-Perrus, A. Schafer, B. Nieswandt, R. K. Andrews, M. W. Parker, E. E. Gardiner, K. Peter, Phosphorothioate backbone modifications of nucleotide-based drugs are potent platelet activators. *J Exp Med* **2015**, *212*, 130-138.
- (50) P. E. Nielsen, M. Egholm, R. H. Berg, O. Buchardt, Sequence-Selective Recognition of DNA by Strand Displacement with a Thymine-Substituted Polyamide. *Science* **1991**, *254*, 1497-1500.
- (51) M. Egholm, O. Buchardt, P. E. Nielsen, R. H. Berg, Peptide Nucleic-Acids (Pna) - Oligonucleotide Analogs with an Achiral Peptide Backbone. *J Am Chem Soc* **1992**, *114*, 1895-1897.
- (52) M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden, P. E. Nielsen, PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules. *Nature* **1993**, *365*, 566-568.
- (53) G. Haima, A. Lohse, O. Buchardt, P. E. Nielsen, Peptide nucleic acids (PNAs) containing thymine monomers derived from chiral amino acids: Hybridization

- and solubility properties of D-lysine PNA. *Angew Chem Int Edit* **1996**, *35*, 1939-1942.
- (54) O. Muse, T. Zengeya, J. Mwaura, D. Hnedzko, D. W. McGee, C. T. Grewer, E. Rozners, Sequence Selective Recognition of Double-Stranded RNA at Physiologically Relevant Conditions Using PNA-Peptide Conjugates. *ACS Chem Biol* **2013**, *8*, 1683-1686.
 - (55) D.-F. K. Toh, G. Devi, K. M. Patil, Q. Qu, M. Maraswami, Y. Xiao, T. P. Loh, Y. Zhao, G. Chen, Incorporating a guanidine-modified cytosine base into triplex-forming PNAs for the recognition of a C-G pyrimidine–purine inversion site of an RNA duplex. *Nucleic Acids Res* **2016**, *44*, 9071-9082.
 - (56) K. M. Patil, D. F. K. Toh, Z. Yuan, Z. Y. Meng, Z. Y. Shu, H. P. Zhang, A. A. L. Ong, M. S. Krishna, L. Y. Lu, Y. P. Lu, G. Chen, Incorporating uracil and 5-halouracils into short peptide nucleic acids for enhanced recognition of A-U pairs in dsRNAs. *Nucleic Acids Res* **2018**, *46*, 7506-7521.
 - (57) T. Zengeya, P. Gupta, E. Rozners, Triple-Helical Recognition of RNA Using 2-Aminopyridine-Modified PNA at Physiologically Relevant Conditions. *Angew Chem Int Edit* **2012**, *51*, 12593-12596.
 - (58) S. C. Brown, S. A. Thomson, J. M. Veal, D. G. Davis, NMR solution structure of a peptide nucleic acid complexed with RNA. *Science* **1994**, *265*, 777-780.
 - (59) P. E. Nielsen, M. Egholm, An introduction to peptide nucleic acid. *Curr Issues Mol Biol* **1999**, *1*, 89-104.
 - (60) K. K. Jensen, H. Orum, P. E. Nielsen, B. Norden, Kinetics for hybridization of peptide nucleic acids (PNA) with DNA and RNA studied with the BIAcore technique. *Biochemistry* **1997**, *36*, 5072-5077.
 - (61) G. L. Igloi, Variability in the stability of DNA-peptide nucleic acid (PNA) single-base mismatched duplexes: real-time hybridization during affinity electrophoresis in PNA-containing gels. *Proc Natl Acad Sci U S A* **1998**, *95*, 8562-8567.
 - (62) D. F. Doyle, D. A. Braasch, C. G. Simmons, B. A. Janowski, D. R. Corey, Inhibition of gene expression inside cells by peptide nucleic acids: effect of mRNA target sequence, mismatched bases, and PNA length. *Biochemistry* **2001**, *40*, 53-64.
 - (63) J. C. Norton, M. A. Piatyszek, W. E. Wright, J. W. Shay, D. R. Corey, Inhibition of human telomerase activity by peptide nucleic acids. *Nat Biotechnol* **1996**, *14*, 615-619.
 - (64) R. Villa, M. Folini, S. Lualdi, S. Veronese, M. G. Daidone, N. Zaffaroni, Inhibition of telomerase activity by a cell-penetrating peptide nucleic acid construct in human melanoma cells. *FEBS Lett* **2000**, *473*, 241-248.
 - (65) J. G. Karras, M. A. Maier, T. Lu, A. Watt, M. Manoharan, Peptide nucleic acids are potent modulators of endogenous pre-mRNA splicing of the murine interleukin-5 receptor-alpha chain. *Biochemistry* **2001**, *40*, 7853-7859.
 - (66) N. Dias, S. Dheur, P. E. Nielsen, S. Gryaznov, A. Van Aerschot, P. Herdewijn, C. Helene, T. E. Saison-Behmoaras, Antisense PNA tridecamers targeted to the coding region of Ha-ras mRNA arrest polypeptide chain elongation. *J Mol Biol* **1999**, *294*, 403-416.
 - (67) N. Dias, C. Senamaud-Beaufort, E. Forestier El, C. Auvin, C. Helene, T. Ester Saison-Behmoaras, RNA hairpin invasion and ribosome elongation arrest by mixed base PNA oligomer. *J Mol Biol* **2002**, *320*, 489-501.
 - (68) M. Kulik, A. Markowska-Zagrajek, M. Wojciechowska, R. Grzela, T. Witula, J. Trylska, Helix 69 of Escherichia coli 23S ribosomal RNA as a peptide nucleic acid target. *Biochimie* **2017**, *138*, 32-42.

- (69) P. J. Farabaugh, Programmed translational frameshifting. *Microbiol Rev* **1996**, *60*, 103-134.
- (70) N. Caliskan, F. Peske, M. V. Rodnina, Changed in translation: mRNA recoding by -1 programmed ribosomal frameshifting. *Trends Biochem Sci* **2015**, *40*, 265-274.
- (71) P. V. Baranov, R. F. Gesteland, J. F. Atkins, Recoding: translational bifurcations in gene expression. *Gene* **2002**, *286*, 187-201.
- (72) A. T. Belew, A. Meskauskas, S. Musalgaonkar, V. M. Advani, S. O. Sulima, W. K. Kasprzak, B. A. Shapiro, J. D. Dinman, Ribosomal frameshifting in the CCR5 mRNA is regulated by miRNAs and the NMD pathway. *Nature* **2014**, *512*, 265-+.
- (73) C. H. Yu, M. H. Noteborn, C. W. A. Pleij, R. C. L. Olsthoorn, Stem-loop structures can effectively substitute for an RNA pseudoknot in -1 ribosomal frameshifting. *Nucleic Acids Res* **2011**, *39*, 8952-8959.
- (74) L. X. Yang, Z. S. Zhong, C. L. Tong, H. Jia, Y. R. Liu, G. Chen, Single-Molecule Mechanical Folding and Unfolding of RNA Hairpins: Effects of Single A-U to A.C Pair Substitutions and Single Proton Binding and Implications for mRNA Structure-Induced -1 Ribosomal Frameshifting. *J Am Chem Soc* **2018**, *140*, 8172-8184.
- (75) S. Matsumoto, N. Caliskan, M. V. Rodnina, A. Murata, K. Nakatani, Small synthetic molecule-stabilized RNA pseudoknot as an activator for -1 ribosomal frameshifting. *Nucleic Acids Res* **2018**, *46*, 8079-8089.
- (76) J. Tholstrup, L. B. Oddershede, M. A. Sorensen, mRNA pseudoknot structures can act as ribosomal roadblocks. *Nucleic Acids Res* **2012**, *40*, 303-313.
- (77) T. Endoh, N. Sugimoto, Unusual -1 ribosomal frameshift caused by stable RNA G-quadruplex in open reading frame. *Anal Chem* **2013**, *85*, 11435-11439.
- (78) C. H. Yu, M. P. Teulade-Fichou, R. C. Olsthoorn, Stimulation of ribosomal frameshifting by RNA G-quadruplex structures. *Nucleic Acids Res* **2014**, *42*, 1887-1892.
- (79) J. D. Dinman, Mechanisms and implications of programmed translational frameshifting. *Wiley Interdiscip Rev RNA* **2012**, *3*, 661-673.
- (80) G. Yusupova, L. Jenner, B. Rees, D. Moras, M. Yusupov, Structural basis for messenger RNA movement on the ribosome. *Nature* **2006**, *444*, 391-394.
- (81) G. Z. Yusupova, M. M. Yusupov, J. H. Cate, H. F. Noller, The path of messenger RNA through the ribosome. *Cell* **2001**, *106*, 233-241.
- (82) S. Takyar, R. P. Hickerson, H. F. Noller, mRNA helicase activity of the ribosome. *Cell* **2005**, *120*, 49-58.
- (83) X. Qu, J. D. Wen, L. Lancaster, H. F. Noller, C. Bustamante, I. Tinoco, Jr., The ribosome uses two active mechanisms to unwind messenger RNA during translation. *Nature* **2011**, *475*, 118-121.
- (84) T. Liu, A. Kaplan, L. Alexander, S. Yan, J. D. Wen, L. Lancaster, C. E. Wickersham, K. Fredrick, H. Noller, I. Tinoco, C. J. Bustamante, Direct measurement of the mechanical work during translocation by the ribosome. *Elife* **2014**, *3*, e03406.
- (85) J. D. Wen, L. Lancaster, C. Hodges, A. C. Zeri, S. H. Yoshimura, H. F. Noller, C. Bustamante, I. Tinoco, Following translation by single ribosomes one codon at a time. *Nature* **2008**, *452*, 598-603.
- (86) T. M. Hansen, S. N. S. Reihani, L. B. Oddershede, M. A. Sorensen, Correlation between mechanical strength of messenger RNA pseudoknots and ribosomal frameshifting. *Proc Natl Acad Sci U S A* **2007**, *104*, 5830-5835.

- (87) G. Chen, K. Y. Chang, M. Y. Chou, C. Bustamante, I. Tinoco, Jr., Triplex structures in an RNA pseudoknot enhance mechanical stability and increase efficiency of -1 ribosomal frameshifting. *Proc Natl Acad Sci U S A* **2009**, *106*, 12706-12711.
- (88) L. Green, C. H. Kim, C. Bustamante, I. Tinoco, Jr., Characterization of the mechanical unfolding of RNA pseudoknots. *J Mol Biol* **2008**, *375*, 511-528.
- (89) I. Brierley, N. J. Rolley, A. J. Jenner, S. C. Inglis, Mutational analysis of the RNA pseudoknot component of a coronavirus ribosomal frameshifting signal. *J Mol Biol* **1991**, *220*, 889-902.
- (90) C. Gaudin, M. H. Mazauric, M. Traikia, E. Guittet, S. Yoshizawa, D. Fourmy, Structure of the RNA signal essential for translational frameshifting in HIV-1. *J Mol Biol* **2005**, *349*, 1024-1035.
- (91) J. Lucchesi, K. Makelainen, A. Merits, T. Tamm, K. Makinen, Regulation of -1 ribosomal frameshifting directed by cocksfoot mottle sobemovirus genome. *Eur J Biochem* **2000**, *267*, 3523-3529.
- (92) B. Larsen, R. F. Gesteland, J. F. Atkins, Structural probing and mutagenic analysis of the stem-loop required for Escherichia coli dnaX ribosomal frameshifting: programmed efficiency of 50%. *J Mol Biol* **1997**, *271*, 47-60.
- (93) O. Namy, S. J. Moran, D. I. Stuart, R. J. Gilbert, I. Brierley, A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting. *Nature* **2006**, *441*, 244-247.
- (94) J. D. Lopinski, J. D. Dinman, J. A. Bruenn, Kinetics of ribosomal pausing during programmed -1 translational frameshifting. *Mol Cell Biol* **2000**, *20*, 1095-1103.
- (95) H. Kontos, S. Naphthine, I. Brierley, Ribosomal pausing at a frameshifter RNA pseudoknot is sensitive to reading phase but shows little correlation with frameshift efficiency. *Mol Cell Biol* **2001**, *21*, 8657-8670.
- (96) D. P. Giedroc, P. V. Cornish, Frameshifting RNA pseudoknots: structure and mechanism. *Virus Res* **2009**, *139*, 193-208.
- (97) D. B. Ritchie, D. A. Foster, M. T. Woodside, Programmed -1 frameshifting efficiency correlates with RNA pseudoknot conformational plasticity, not resistance to mechanical unfolding. *Proc Natl Acad Sci U S A* **2012**, *109*, 16167-16172.
- (98) I. Brierley, A. J. Jenner, S. C. Inglis, Mutational analysis of the "slippery-sequence" component of a coronavirus ribosomal frameshifting signal. *J Mol Biol* **1992**, *227*, 463-479.
- (99) S. Cao, S. J. Chen, Predicting ribosomal frameshifting efficiency. *Phys Biol* **2008**, *5*, 016002.
- (100) Z. Lin, R. J. Gilbert, I. Brierley, Spacer-length dependence of programmed -1 or -2 ribosomal frameshifting on a U6A heptamer supports a role for messenger RNA (mRNA) tension in frameshifting. *Nucleic Acids Res* **2012**, *40*, 8674-8689.
- (101) W. J. Craigie, C. T. Caskey, Expression of peptide chain release factor 2 requires high-efficiency frameshift. *Nature* **1986**, *322*, 273-275.
- (102) I. P. Ivanov, J. F. Atkins, Ribosomal frameshifting in decoding antizyme mRNAs from yeast and protists to humans: close to 300 cases reveal remarkable diversity despite underlying conservation. *Nucleic Acids Res* **2007**, *35*, 1842-1858.
- (103) L. Kurian, R. Palanimurugan, D. Godderz, R. J. Dohmen, Polyamine sensing by nascent ornithine decarboxylase antizyme stimulates decoding of its mRNA. *Nature* **2011**, *477*, 490-494.
- (104) A. E. Firth, B. W. Jagger, H. M. Wise, C. C. Nelson, K. Parsawar, N. M. Wills, S. Naphthine, J. K. Taubenberger, P. Digard, J. F. Atkins, Ribosomal frameshifting

- used in influenza A virus expression occurs within the sequence UCC_UUU_CGU and is in the +1 direction. *Open Biol* **2012**, *2*, 120109.
- (105) Y. Fang, E. E. Treffers, Y. Li, A. Tas, Z. Sun, Y. van der Meer, A. H. de Ru, P. A. van Veelen, J. F. Atkins, E. J. Snijder, A. E. Firth, Efficient -2 frameshifting by mammalian ribosomes to synthesize an additional arterivirus protein. *Proc Natl Acad Sci U S A* **2012**, *109*, E2920-2928.
 - (106) Y. Li, E. E. Treffers, S. Naphine, A. Tas, L. Zhu, Z. Sun, S. Bell, B. L. Mark, P. A. van Veelen, M. J. van Hemert, A. E. Firth, I. Brierley, E. J. Snijder, Y. Fang, Transactivation of programmed ribosomal frameshifting by a viral protein. *Proc Natl Acad Sci U S A* **2014**, *111*, E2172-2181.
 - (107) Y. Wang, J. Liu, B. O. Huang, Y. M. Xu, J. Li, L. F. Huang, J. Lin, J. Zhang, Q. H. Min, W. M. Yang, X. Z. Wang, Mechanism of alternative splicing and its regulation. *Biomed Rep* **2015**, *3*, 152-158.
 - (108) N. A. Faustino, T. A. Cooper, Pre-mRNA splicing and human disease. *Genes Dev* **2003**, *17*, 419-437.
 - (109) J. Lukas, D. Q. Gao, M. Keshmeshian, W. H. Wen, D. Tsao-Wei, S. Rosenberg, M. F. Press, Alternative and aberrant messenger RNA splicing of the mdm2 oncogene in invasive breast cancer. *Cancer Res* **2001**, *61*, 3212-3219.
 - (110) J. D. Fackenthal, L. A. Godley, Aberrant RNA splicing and its functional consequences in cancer cells. *Dis Model Mech* **2008**, *1*, 37-42.
 - (111) B. Ghetti, A. L. Oblak, B. F. Boeve, K. A. Johnson, B. C. Dickerson, M. Goedert, Invited review: Frontotemporal dementia caused by microtubule-associated protein tau gene (MAPT) mutations: a chameleon for neuropathology and neuroimaging. *Neuropathol Appl Neurobiol* **2015**, *41*, 24-46.
 - (112) C. Li, J. Gotz, Tau-based therapies in neurodegeneration: opportunities and challenges. *Nat Rev Drug Discov* **2017**, *16*, 863-883.
 - (113) M. G. Spillantini, J. R. Murrell, M. Goedert, M. R. Farlow, A. Klug, B. Ghetti, Mutation in the tau gene in familial multiple system tauopathy with presenile dementia. *Proc Natl Acad Sci U S A* **1998**, *95*, 7737-7741.
 - (114) I. D'Souza, P. Poorkaj, M. Hong, D. Nochlin, V. M. Lee, T. D. Bird, G. D. Schellenberg, Missense and silent tau gene mutations cause frontotemporal dementia with parkinsonism-chromosome 17 type, by affecting multiple alternative RNA splicing regulatory elements. *Proc Natl Acad Sci U S A* **1999**, *96*, 5598-5603.
 - (115) E. Buratti, F. E. Baralle, Influence of RNA secondary structure on the pre-mRNA splicing process. *Mol Cell Biol* **2004**, *24*, 10505-10514.
 - (116) A. Grover, H. Houlden, M. Baker, J. Adamson, J. Lewis, G. Prihar, S. Pickering-Brown, K. Duff, M. Hutton, 5' splice site mutations in tau associated with the inherited dementia FTDP-17 affect a stem-loop structure that regulates alternative splicing of exon 10. *J Biol Chem* **1999**, *274*, 15134-15143.
 - (117) C. P. Donahue, C. Muratore, J. Y. Wu, K. S. Kosik, M. S. Wolfe, Stabilization of the tau exon 10 stem loop alters pre-mRNA splicing. *J Biol Chem* **2006**, *281*, 23302-23306.
 - (118) L. Varani, M. Hasegawa, M. G. Spillantini, M. J. Smith, J. R. Murrell, B. Ghetti, A. Klug, M. Goedert, G. Varani, Structure of tau exon 10 splicing regulatory element RNA and destabilization by mutations of frontotemporal dementia and parkinsonism linked to chromosome 17. *Proc Natl Acad Sci U S A* **1999**, *96*, 8229-8234.
 - (119) J. Zhou, Q. Yu, T. Zou, Alternative splicing of exon 10 in the tau gene as a target for treatment of tauopathies. *BMC Neurosci* **2008**, *9 Suppl 2*, S10.

- (120) P. M. Stanford, C. E. Shepherd, G. M. Halliday, W. S. Brooks, P. W. Schofield, H. Brodaty, R. N. Martins, J. B. J. Kwok, P. R. Schofield, Mutations in the tau gene that cause an increase in three repeat tau and frontotemporal dementia. *Brain* **2003**, *126*, 814-826.
- (121) J. P. Staley, C. Guthrie, Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **1998**, *92*, 315-326.
- (122) C. L. Chen, H. B. Zhang, S. L. Broitman, M. Reiche, I. Farrell, B. S. Cooperman, Y. E. Goldman, Dynamics of translation by single ribosomes through mRNA secondary structures. *Nat Struct Mol Biol* **2013**, *20*, 582-+.
- (123) H. Yin, M. D. Wang, K. Svoboda, R. Landick, S. M. Block, J. Gelles, Transcription against an applied force. *Science* **1995**, *270*, 1653-1657.
- (124) E. A. Abbondanzieri, W. J. Greenleaf, J. W. Shaevitz, R. Landick, S. M. Block, Direct observation of base-pair stepping by RNA polymerase. *Nature* **2005**, *438*, 460-465.
- (125) M. H. Larson, R. Landick, S. M. Block, Single-molecule studies of RNA polymerase: one singular sensation, every little step it takes. *Mol Cell* **2011**, *41*, 249-262.
- (126) J. Zlatanova, S. M. Lindsay, S. H. Leuba, Single molecule force spectroscopy in biology using the atomic force microscope. *Prog Biophys Mol Bio* **2000**, *74*, 37-61.
- (127) H. J. Butt, B. Cappella, M. Kappl, Force measurements with the atomic force microscope: Technique, interpretation and applications. *Surf Sci Rep* **2005**, *59*, 1-152.
- (128) K. C. Neuman, A. Nagy, Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat Methods* **2008**, *5*, 491-505.
- (129) D. K. Sasmal, L. E. Pulido, S. Kasal, J. Huang, Single-molecule fluorescence resonance energy transfer in molecular biology. *Nanoscale* **2016**, *8*, 19928-19944.
- (130) E. D. Holmstrom, D. J. Nesbitt, Single-molecule fluorescence resonance energy transfer studies of the human telomerase RNA pseudoknot: temperature-/urea-dependent folding kinetics and thermodynamics. *J Phys Chem B* **2014**, *118*, 3853-3863.
- (131) L. Ying, J. J. Green, H. Li, D. Klenerman, S. Balasubramanian, Studies on the structure and dynamics of the human telomeric G quadruplex by single-molecule fluorescence resonance energy transfer. *Proc Natl Acad Sci U S A* **2003**, *100*, 14629-14634.
- (132) A. A. Deniz, T. A. Laurence, G. S. Beligere, M. Dahan, A. B. Martin, D. S. Chemla, P. E. Dawson, P. G. Schultz, S. Weiss, Single-molecule protein folding: diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2. *Proc Natl Acad Sci U S A* **2000**, *97*, 5179-5184.
- (133) A. F. Sauer-Budge, J. A. Nyamwanda, D. K. Lubensky, D. Branton, Unzipping kinetics of double-stranded DNA in a nanopore. *Phys Rev Lett* **2003**, *90*, 238101.
- (134) J. Mathe, H. Visram, V. Viasnoff, Y. Rabin, A. Meller, Nanopore unzipping of individual DNA hairpin molecules. *Biophys J* **2004**, *87*, 3205-3212.
- (135) O. K. Dudko, J. Mathe, A. Szabo, A. Meller, G. Hummer, Extracting kinetics from single-molecule force spectroscopy: nanopore unzipping of DNA hairpins. *Biophys J* **2007**, *92*, 4188-4195.

- (136) H. J. You, J. Y. Wu, F. W. Shao, J. Yan, Stability and Kinetics of c-MYC Promoter G-Quadruplexes Studied by Single-Molecule Manipulation. *J Am Chem Soc* **2015**, *137*, 2424-2427.
- (137) J. Liphardt, B. Onoa, S. B. Smith, I. Tinoco, Jr., C. Bustamante, Reversible unfolding of single RNA molecules by mechanical force. *Science* **2001**, *292*, 733-737.
- (138) G. Chen, J. D. Wen, I. Tinoco, Jr., Single-molecule mechanical unfolding and folding of a pseudoknot in human telomerase RNA. *RNA* **2007**, *13*, 2175-2188.
- (139) T. L. Min, P. J. Mears, L. M. Chubiz, C. V. Rao, I. Golding, Y. R. Chemla, High Resolution, Long Term Characterization of Bacterial Motility Using Optical Tweezers. *Biophys J* **2010**, *98*, 416a-416a.
- (140) A. Ashkin, J. M. Dziedzic, J. E. Bjorkholm, S. Chu, Observation of a single-beam gradient force optical trap for dielectric particles. *Opt Lett* **1986**, *11*, 288.
- (141) A. Ashkin, J. M. Dziedzic, Optical trapping and manipulation of viruses and bacteria. *Science* **1987**, *235*, 1517-1520.
- (142) H. J. Geertsema, K. E. Duderstadt, A. M. van Oijen, Single-molecule observation of prokaryotic DNA replication. *Methods Mol Biol* **2015**, *1300*, 219-238.
- (143) K. M. Herbert, W. J. Greenleaf, S. M. Block, Single-molecule studies of RNA polymerase: motoring along. *Annu Rev Biochem* **2008**, *77*, 149-176.
- (144) S. B. Smith, Y. Cui, C. Bustamante, Optical-trap force transducer that operates by direct measurement of light momentum. *Methods Enzymol* **2003**, *361*, 134-162.
- (145) N. Bosaeus, A. H. El-Sagheer, T. Brown, S. B. Smith, B. Akerman, C. Bustamante, B. Norden, Tension induces a base-paired overstretched DNA conformation. *Proc Natl Acad Sci U S A* **2012**, *109*, 15179-15184.
- (146) C. Cecconi, E. A. Shank, F. W. Dahlquist, S. Marqusee, C. Bustamante, Protein-DNA chimeras for single molecule mechanical folding studies with the optical tweezers. *Eur Biophys J* **2008**, *37*, 729-738.
- (147) J. M. Huguet, C. V. Bizarro, N. Forns, S. B. Smith, C. Bustamante, F. Ritort, Single-molecule derivation of salt dependent base-pair free energies in DNA. *Proc Natl Acad Sci U S A* **2010**, *107*, 15431-15436.
- (148) J. P. Berezney, O. A. Saleh, Locked nucleic acid oligomers as handles for single molecule manipulation. *Nucleic Acids Res* **2014**, *42*, e150.
- (149) V. S. Jadhav, D. Bruggemann, F. Wruck, M. Hegner, Single-molecule mechanics of protein-labelled DNA handles. *Beilstein J Nanotech* **2016**, *7*, 138-148.
- (150) C. Cecconi, E. A. Shank, S. Marqusee, C. Bustamante, DNA molecular handles for single-molecule protein-folding studies by optical tweezers. *Methods Mol Biol* **2011**, *749*, 255-271.
- (151) I. Tinoco, Jr., R. L. Gonzalez, Jr., Biological mechanisms, one molecule at a time. *Genes Dev* **2011**, *25*, 1205-1231.
- (152) K. Svoboda, S. M. Block, Biological Applications of Optical Forces. *Annu Rev Biophys Biomol Struct* **1994**, *23*, 247-285.
- (153) S. B. Smith, Y. Cui, C. Bustamante, Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science* **1996**, *271*, 795-799.
- (154) J. Ubbink, T. Odijk, Polymer- and salt-induced toroids of hexagonal DNA. *Biophys J* **1995**, *68*, 54-61.

- (155) D. Bensimon, A. J. Simon, V. V. Croquette, A. Bensimon, Stretching DNA with a receding meniscus: Experiments and models. *Phys Rev Lett* **1995**, *74*, 4754-4757.
- (156) C. Bustamante, J. F. Marko, E. D. Siggia, S. Smith, Entropic elasticity of lambda-phage DNA. *Science* **1994**, *265*, 1599-1600.
- (157) F. Vanzi, Y. Takagi, H. Shuman, B. S. Cooperman, Y. E. Goldman, Mechanical studies of single ribosome/mRNA complexes. *Biophys J* **2005**, *89*, 1909-1919.
- (158) I. Tinoco, Jr., Force as a useful variable in reactions: unfolding RNA. *Annu Rev Biophys Biomol Struct* **2004**, *33*, 363-385.
- (159) K. Neupane, H. Yu, D. A. Foster, F. Wang, M. T. Woodside, Single-molecule force spectroscopy of the add adenine riboswitch relates folding to regulatory mechanism. *Nucleic Acids Res* **2011**, *39*, 7677-7687.
- (160) J. R. Wenner, M. C. Williams, I. Rouzina, V. A. Bloomfield, Salt dependence of the elasticity and overstretching transition of single DNA molecules. *Biophys J* **2002**, *82*, 3160-3169.
- (161) H. Chen, S. P. Meisburger, S. A. Pabit, J. L. Sutton, W. W. Webb, L. Pollack, Ionic strength-dependent persistence lengths of single-stranded RNA and DNA. *Proc Natl Acad Sci U S A* **2012**, *109*, 799-804.
- (162) P. T. Li, D. Collin, S. B. Smith, C. Bustamante, I. Tinoco, Jr., Probing the mechanical folding kinetics of TAR RNA by hopping, force-jump, and force-ramp methods. *Biophys J* **2006**, *90*, 250-260.
- (163) J. D. Wen, M. Manosas, P. T. Li, S. B. Smith, C. Bustamante, F. Ritort, I. Tinoco, Jr., Force unfolding kinetics of RNA using optical tweezers. I. Effects of experimental variables on measured results. *Biophys J* **2007**, *92*, 2996-3009.
- (164) S. A. McKinney, C. Joo, T. Ha, Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophys J* **2006**, *91*, 1941-1951.
- (165) O. K. Dudko, G. Hummer, A. Szabo, Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys Rev Lett* **2006**, *96*, 108101.
- (166) S. Guo, Q. Tang, M. Yao, H. You, S. Le, H. Chen, J. Yan, Structural-elastic determination of the force-dependent transition rate of biomolecules. *Chem Sci* **2018**, *9*, 5871-5882.
- (167) H. You, S. Guo, S. Le, Q. Tang, M. Yao, X. Zhao, J. Yan, Two-State Folding Energy Determination Based on Transition Points in Nonequilibrium Single-Molecule Experiments. *J Phys Chem Lett* **2018**, *9*, 811-816.

Chapter 2

Single-Molecule Mechanical Folding and Unfolding of RNA

Hairpins: Effects of A·C mismatch and Single Proton Binding

and Implications for mRNA Structure-Induced –1 Ribosomal Frameshifting

2.1 Introduction

RNAs form complex and dynamic structures for various biological functions such as serving as templates for the synthesis of proteins and nucleic acids, and as regulators or catalytic centers for transcription, pre-mRNA splicing, and protein synthesis.¹ RNA secondary structures are typically composed of double-stranded stems and various loops, which can further assemble into tertiary structures.²⁻⁴ The RNA stems are stabilized by canonical base pairs including A-U, G-C, and G·U.^{3,5,6} Formation of various non-canonical base pairs such as G·A, U·U, U·C, and A·C pairs⁷⁻¹⁴ contributes to the complex structures and dynamics of RNA. Non-canonical base pairs, in turn, are often involved in RNA tertiary interactions and facilitate binding to proteins, small molecules, and ions.

Shifting of pK_a values and protonation dependent conformational dynamics in RNAs¹⁵⁻²³ are important for the catalytic activities of ribozymes,^{18,19,24-26} small molecule-RNA interactions,^{27,28} protein-RNA assembly,²⁹⁻³¹ and regulation of protein expression.^{32,33} A wobble A·C pair (*cis* Watson–Crick/Watson–Crick³⁴)

(**Figure 2.1a**) is structurally similar to a wobble G·U pair,^{35,36} with the N1 nitrogen atom of A often having a pK_a near neutrality primarily driven by the formation of a wobble A⁺·C pair with one extra hydrogen bond.^{16,18,19,37-39}

The sequence and structural contexts and environmental factors affect the formation of wobble or non-wobble A·C pairs. For example, an “open” A·C mismatch with no hydrogen bonding interactions and a wobble A·C pair are found in the acceptor stem and anticodon loop of tRNAs, respectively.⁴⁰⁻⁴² A wobble A·C pair also exists in ribozymes such as *Neurospora* Varkud satellite (VS) ribozyme,^{26,43} and the artificial leadzyme with the self-cleavage activity stimulated by micromolar concentrations of lead ions.¹⁸ An RNA duplex containing an A·C pair (compared to A·U, A·A, and A·G pairs) is a more favored substrate for the RNA editing enzyme adenosine deaminase acting on RNA (ADAR).^{44,45} A wobble A·C pair found in the pre-catalytic conformation of intramolecular stem-loop of U6 snRNA may be critical for regulating the binding and positioning of a catalytic magnesium ion and the dynamic assembly of spliceosomal complexes.⁴⁶⁻⁴⁹ A substitution of an A·U pair with an A·C pair resulting from genetic mutations in the microtubule-associated tau pre-mRNA splice site hairpin may cause dysregulation of alternative splicing and thus brain disorders such as frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP-17).^{50,51}

Figure 2.1. Wobble A·C and A⁺·C base pairs and 32-mer hairpins studied in this paper. (a) Chemical structures of wobble A·C and A⁺·C base pairs. (b) RNA constructs. U22C, U25C and U29C are mutants with the single A·C mismatch at the top, middle, and bottom position of the stem, respectively. The mutated residues are shown in red colour. The open and closed arrows indicate the approximate apparent two-state mechanical unfolding and folding transition state positions, respectively, revealed in this work.

Thermal unfolding studies for relatively small model RNAs and DNAs reveal that pH dependent wobble A⁺·C pair formation enhances folding stability.^{9,16,19,52} Single-molecule manipulation using optical tweezers allows the characterization of RNA mechanical (un)folding at near-physiological buffer conditions at room temperature, and may provide deeper insights into the complex folding free energy landscapes⁵³⁻⁵⁵ and the effects of protonation. In addition, mechanical

(un)folding may better mimic how RNA structures (un)fold in physiological conditions.^{13,56-69} For example, during ribosomal translation, mRNA structures in the coding regions need to be unfolded into single strand form at the mRNA entrance tunnel of the small subunit of ribosome.^{70,71} The force exerted by ribosome during translation elongation is about 12–20 pN.^{72,73}

Minus-one programmed ribosomal frameshifting is a mechanism utilized by cells and viruses to regulate the expression of multiple proteins from one mRNA as well as mRNA half-life.⁷⁴⁻⁷⁸ Minus-one programmed ribosomal frameshifting can be stimulated by an mRNA structure (hairpin or pseudoknot) located downstream a slippery site (e.g., U UUU UUA), where ribosome may be temporarily stalled. A backward slippage of ribosomal reading frame by one nucleotide may release the tension and result in the translation at –1 frame.^{66,67,78-85} The mechanical stability of the downstream mRNA structure shows positive correlation with –1 frameshifting efficiency.^{66,67,81} However, other factors may also affect –1 frameshifting efficiency.⁸⁶

Here, we combined traditional ensemble UV-absorbance-detected thermal melting and single-molecule mechanical (un)folding methods to unveil how the structural stability and (un)folding dynamics of a series of 32-mer RNA hairpins (**Figure 2.1b**) are affected by (i) a single A-U to A·C pair substitution at varied stem positions and (ii) the protonation of the single A·C pairs by varying the buffer pH. Furthermore, we used an *in vitro* translation assay to test how the single A-U to A·C pair substitutions in mRNA hairpins affect –1 ribosomal frameshifting.⁷⁴⁻⁷⁸

2.2 Materials and Methods

2.2.1 Thermal Melting Experiments

We constructed a 32-mer RNA hairpin (rHP) and three mutants with one A-U base pair substituted with an A·C mismatch on the top (U22C), middle (U25C) and bottom (U29C) position of the stem, respectively, for the thermal melting experiments (**Figure 2.1b**). The 32-mer RNA oligonucleotides (**Figure 2.1b**) were purchased from Sigma with HPLC-purification. The 32-mer hairpins have the sequences of 5'-UAGAGAGAGAAAGUUUCGACUUUCUCUCUCUA-3' (rHP), 5'-UAGAGAGAGAGAAAGUUUCGACUCUCUCUCUCUA-3' (U22C), 5'-UAGAGAGAGAGAAAGUUUCGACUUUCCCUCUCUA-3' (U25C), and 5'-UAGAGAGAGAGAAAGUUUCGACUUUCUCUCCCUA-3' (U29C). All thermal melting studies for the RNA hairpins were carried out using a Shimadzu 2550 spectrophotometer. The absorption at 260 nm was recorded. The six thermal melting buffers with the pH ranging from 5.0 to 8.0 are 200 mM NaCl, 20 mM NaOAc, 0.1 mM EDTA, pH 5.0; 200 mM NaCl, 20 mM MES, 0.1 mM EDTA, pH 5.5 or 6.0; and 200 mM NaCl, 20 mM HEPES, 0.1 mM EDTA, pH 6.5, 7.3, or pH 8.0. NaOAc, MES and HEPES are used because their pK_a values change insignificantly when temperature varies. The sample concentrations used in the UV-absorbance-detected thermal melting studies were 3 μ M. The temperature was firstly held at 95 °C for 15 minutes, then decreased from 95 °C to 20 °C and increased back to 95 °C at a ramp rate of 0.5 °C/min. The thermal melting experiments were repeated three times. The normalized melting curves were obtained by dividing the measured absorption values at 260 nm by the lowest

adsorption value in each curve. The first derivatives of the thermal curves were fit to Gaussian functions to extract the thermal melting temperatures.

2.2.2 Optical Tweezers Experiments

The single-molecule constructs suitable for optical tweezer experiment were constructed as previously reported.^{61,66} The DNA sequences were inserted into p2luc vector as reported previously.⁸⁷ The mutated sequences for the constructs with an A·C mismatch were constructed by site directed mutagenesis. The constructed vectors were transformed into *E. coli* competent cells (DH5 α). The plasmids were purified from the bacterial cell cultures. The sequences of all the constructs were verified by Applied Biosystems highest capacity-based genetic analyzer platforms (1st BASE).

With properly designed primers, the duplex DNAs of about 1170 base pairs (bp) containing a T7 promoter were amplified by PCR. *In vitro* transcription by T7 RNA polymerase was carried out to transcribe the RNAs. A 583-bp DNA handle B and a 521-bp DNA handle A flanking the 32-mer sequences of interest were also generated by PCR. The RNA containing the 32-nucleotide (nt) native hairpin sequences, flanked by a 583-nt upstream sequence together with a 4-nt linker and a 521-nt downstream sequence together with a 4-nt linker, were then annealed with complementary DNA strands to form RNA/DNA hybrid handle B and handle A, respectively. The 3' end of the DNA strand of the RNA/DNA hybrid handle A was labelled by introducing biotin-modified dUTP (Biotin-16-dUTP, Roche) using T4 DNA polymerase. The 5' end of the DNA strand of handle B was labelled by digoxigenin-labelled primer during PCR. The streptavidin coated

polystyrene beads with a diameter of $\sim 1.8 \mu\text{m}$ were purchased from Spherotech. The anti-digoxigenin coated beads were synthesized by cross-linking the anti-digoxigenin antibody to the protein G-coated beads with a diameter of $\sim 3 \mu\text{m}$. The protein G-coated beads were synthesized by coating the recombinant protein G (Thermo Scientific) on the carboxyl coated polystyrene beads (Spherotech) using EDC (Sigma) and sulfo-NHS (Thermo Scientific). Well-annealed DNA/RNA hybrids are then linked to both beads to form a tether for the mechanical pulling experiments.

The single-molecule mechanical pulling experiments were performed using a dual-beam Minitweezers with a dual-laser optical trap generated by two well aligned and highly focused laser beams.^{73,88,89} During the pulling experiment, handle B was attached to an anti-digoxigenin coated polystyrene bead through digoxigenin-anti-digoxigenin interaction and handle A was attached to a streptavidin coated polystyrene bead through biotin-streptavidin interaction. The anti-digoxigenin-coated polystyrene bead was trapped in an optical trap and the streptavidin-coated polystyrene bead was held on a micropipette by suction. We used three different pulling modes (force-ramp, constant-force, and passive modes) for the characterization of the mechanical (un)folding dynamics. The force was varied by moving the trap at 100 nm/s during the force-ramp experiments. In constant-force experiments, the force was maintained at a pre-set value by an electronic feedback with a response time of 5 ms . In a passive mode, there is no electronic feedback. The constant trap stiffness was about 0.07 pN/nm . The data acquisition rate was 1000 Hz . All the mechanical pulling experiments were carried out at a constant temperature of $22 \pm 1 \text{ }^\circ\text{C}$. The optical

tweezers experiments were carried out in four buffers: (i) 200 mM NaCl, 20 mM MES, 0.1 mM EDTA, pH 5.0; (ii) 200 mM NaCl, 20 mM HEPES, 0.1 mM EDTA, pH 6.5; (iii) 200 mM NaCl, 20 mM HEPES, 0.1 mM EDTA, pH 7.3; and (iv) 200 mM NaCl, 20 mM HEPES, 0.1 mM EDTA, pH 8.0.

2.2.3 *In vitro* Ribosomal Frameshifting Assay

The plasmids used for the frameshifting assay are the same as those for mechanical pulling studies. Briefly, we inserted in between the *Renilla* luciferase (RLuc) and firefly luciferase (FLuc) reporter genes of p2luc vector^{87,90,91} the target DNA sequences of experimental rHP (5'-GAT CCT TTT TTA GGG TAG AGA GAG AAA GTT TCG ACT TTC TCT CTC TAG AGC T-3'), and a positive control rHP (5'-GAT CCA CTT CTT AGG GTA GAG AGA GAA AGT TTC GAC TTT CTC TCT CTA GAG CT-3'). The mutants were made by site-directed mutagenesis and have the sequences of U22C (5'-GAT CCT TTT TTA GGG TAG AGA GAG AAA GTT TCG ACT CTC TCT CTC TAG AGC T-3'), U25C (5'-GAT CCT TTT TTA GGG TAG AGA GAG AAA GTT TCG ACT TTC CCT CTC TAG AGC T-3'), and U29C (5'-GAT CCT TTT TTA GGG TAG AGA GAG AAA GTT TCG ACT TTC TCT CCC TAG AGC T-3').

The plasmids were linearized using PmlI restriction enzyme (NEB), followed by *in vitro* transcription to generate mRNA templates. 8.75 μ L nuclease treated Rabbit Reticulocyte Lysate mixture (Promega) was added into mRNA with a final volume of 12.5 μ L. The final concentration of mRNA is 0.1 μ M. The mixture was incubated at 30 °C for the *in vitro* translation reaction for 90 minutes.

The *in vitro* translation was quenched by placing the tubes on ice for 20 minutes. The pH of the translation mixture is within 7.2–7.6.

The luminescence of the translation products was quantified by using the dual-luciferase reporter assay (Promega).^{87,90,92,93} The luminescence was measured using Tecan infinite M200 microplate reader on Nunclon™ flat-bottom 96-well black microplate with an integration time of 3 seconds at room temperature. The reading was taken immediately after the addition of 50 µL of the respective substrates. The *in vitro* translation reactions were conducted three times. The –1 frameshifting efficiency (–1 FS) was calculated using the equation, $-1 \text{ FS (\%)} = \frac{\text{FLuc(E)}/\text{RLuc(E)}}{\text{FLuc(C)}/\text{RLuc(C)}} \times 100\%$.⁹⁰ Here, FLuc (E) and RLuc (E) are the luminescence reading values of FLuc and RLuc, respectively, for the experimental constructs. FLuc (C) and RLuc (C) are the luminescence reading values of FLuc and RLuc, respectively, for the positive control construct of rHP.

2.3 Results and Discussion

2.3.1 Thermal Unfolding

Hairpins U25C and U29C have a single A·C mismatch, respectively, at the middle and bottom part of the stem of the hairpin flanked by two G-C base pairs, whereas hairpin U22C has an A·C mismatch at the top part of the stem of the hairpin flanked by two A-U base pairs (**Figure 2.1b**). According to a nearest-neighbour model, at 1 M NaCl, pH 7, 37 °C (calculated using RNAstructure 5.3),^{3,94-96} a single A-U to A·C substitution induces a free energy destabilization

effect for hairpins U22C, U25C, and U29C by 3.7, 4.8 and 4.8 kcal/mol, respectively.

Figure 2.2. Thermal denaturation experiments reveal limited pH dependency of thermal stability. (a) Representative normalized thermal melting curves (heating) monitored at 260 nm for rHP (red), U22C (cyan), U25C (magenta) and U29C (blue) at 200 mM NaCl, pH 7.3. The two thermal melting transitions of U22C are indicated with cyan arrows. (b) Melting (heating) temperatures (T_m) of rHP, U22C, U25C and U29C at 200 mM NaCl, pH 5.0, 5.5, 6.0, 6.5, 7.3, and 8.0. Errors of melting temperatures are reported as standard deviations. (c) Proposed one-step thermal unfolding and folding pathways for rHP, U25C and U29C. (d) Proposed two-step thermal unfolding and folding pathways for U22C. Due to the relatively large size of the hairpins, the thermal melting transitions are relatively

broad and may involve more intermediate states. Hence, the thermodynamic parameters were not extracted from the thermal melting curves.

The thermal melting data at 200 mM NaCl reveal that the heating and cooling curves are superimposable at pH 6.0, 6.5, 7.3 and 8.0, indicating equilibrated (un)folding transitions. A single thermal unfolding transition was observed for hairpins rHP, U25C, and U29C (**Figures 2.1b and 2.2a,c**). Hairpin U22C, however, shows a two-step thermal unfolding from native hairpin structure to an intermediate, and then to single-stranded random coil (**Figure 2.2a,d**). By comparing the thermal melting curves of the four hairpins, we can conclude that the top part of the stem adjacent to the A11·C22 pair of hairpin U22C thermally melts first (with a $T_{m,1}$ value of ~ 53 °C), followed by the thermal melting of the remaining bottom stem (with a $T_{m,2}$ value of ~ 79 °C) (**Figures 2.1b and 2.2d**). Consistent with the prediction by the RNAstructure 5.3 program,^{3,94-96} we observed that, at 200 mM NaCl, pH 7.3, the T_m values of hairpins U22C, U25C, and U29C are lower than that of rHP (**Figure 2.2a,b and Table 2.1**). However, we observed position-dependent destabilization effect, with U25C being the least stable at pH 7.3, although U25C has the same nearest-neighbor combinations of base pairs as U29C.

Table 2.1. Thermal melting temperatures at 200 mM NaCl with varied pH.^a

RNA molecule	Reaction	pH 5.0	pH 5.5	pH 6.0	pH 6.5	pH 7.3	pH 8.0
rHP	N→U ^a	83.9±0.2	84.1±0.7	84.5±0.3	84.6±0.3	84.3±0.6	83.9±0.4
U22C	N→I	53.3±0.1	53.2±0.4	53.2±0.4	53.0±0.5	52.9±0.1	52.9±0.2
	I→U	79.7±0.3	79.3±0.3	79.9±0.1	79.3±0.1	79.4±0.2	79.6±0.1
U25C	N→U	80.7±0.3	80.5±0.2	80.2±0.0	79.5±0.2	77.7±0.1	76.8±0.1
U29C	N→U	81.8±0.2	81.1±0.2	81.4±0.1	80.9±0.2	80.3±0.2	80.3±0.2

^a The N–U reaction is the melting of the native hairpin to single-stranded random coil. The N–I and I–U reactions are the melting of the native hairpin to the intermediate structure, and the intermediate structure to single-stranded random coil, respectively. All the experimental data are reported as mean $\pm$ standard deviation from three repeated experiments.

Previous thermal unfolding studies for relatively short model RNAs and DNAs have revealed that pH dependent wobble A⁺·C pair formation enhances folding stability.^{9,16,19,52} We carried out thermal melting studies of the relatively large RNA hairpins with the pH ranging from 5.0 to 8.0 (**Figure 2.2b**). The thermal melting results reveal that pH doesn't influence the stability of rHP, with the T_m of rHP remaining at around 84 °C. Decreasing the pH from 8.0 to 5.0, however, results in an increase of the T_m values of U25C and U29C by 3.9 °C and 1.5 °C, respectively. Surprisingly, neither $T_{m,1}$ (53 °C) nor $T_{m,2}$ (79 °C) of U22C shows a significant change upon varying the pH.

The relatively broad thermal melting transitions of the relatively large hairpin constructs studied here may prevent the revelation of the effect of pH on structural stability. It's also probable that UV-absorbance-detected thermal melting studies are only informative for relatively short RNAs with T_m values close to physiologically relevant temperatures. In addition, the pK_a values of a wobble A⁺·C pair and other structures are often highly temperature dependent, which may complicate the analysis of the pH dependent thermal melting data.^{9,16,19,97} For example, due to the fact that the pK_a value of the A residue in a wobble A⁺·C pair decreases with increasing temperature,¹⁶ increasing temperature is expected to weaken the protonation dependent stabilization effect of a wobble A⁺·C pair. Interestingly, we observed hysteresis between thermal unfolding and folding transitions at pH 5.5 and 5.0. We repeated the heating and

cooling cycles, and observed superimposable heating and cooling curves, respectively, although the hysteresis exists between heating and cooling traces. Thus, we speculate that the purine rich part of the RNAs may form non-native multimeric structures stabilized by protonated A·A and A⁺·A pairs at relatively low pH^{20,21} before folding into native hairpin structures. We didn't extract the thermodynamic parameters from the thermal melting curves due to the reasons discussed above.

Figure 2.3. Optical tweezers experiment setup and representative mechanical (un)folding force-extension curves (FECs) and proposed mechanical (un)folding

pathways. (a) Schematic of experimental setup for mechanical (un)folding of RNA hairpins using Minitweezers (not to scale). (b) Representative FECs. Force-loading and force-unloading curves are shown in colors and black, respectively. The curves are smoothed to 200 Hz. rHP has relatively higher unfolding forces than the three mutants. U22C has a relatively large hysteresis between unfolding and folding. (c,d) Proposed one-step mechanical (un)folding pathways. The terminal 4 base pairs of U29C do not form at pH 8.0.

2.3.2 Mechanical (Un)Folding Dynamics

We next employed single-molecule mechanical (un)folding experiments using high resolution optical tweezers to investigate at room temperature (**Figure 2.3a**) how a single A-U to A-C substitution and pH affect the stability and dynamics of the RNA hairpins. To facilitate the single-molecule mechanical pulling experiments, the hairpin constructs were extended on the 5' and 3' ends with DNA/RNA hybrid handles (**Figures 2.1b and 2.3a**). In a single-molecule force-ramp experiment, a varying force is applied to the terminal ends of RNA to facilitate the mechanical folding and unfolding of the RNA structures (**Figure 2.3**). One-step mechanical (un)folding transitions (with unfolding indicated by an abrupt extension increase and force decrease) were observed for all hairpins at varied pH's (**Figure 2.3b,c**). The force-extension curves (FECs) of the force-ramp data suggest that rHP shows transitions near 17 pN with an end-to-end extension change ΔX around 12.5 nm (**Figure 2.3b**), which is consistent with the formation of a 32-mer hairpin as estimated by an extensible wormlike chain (EWLC) model (Equation 2.1).⁹⁸⁻¹⁰²

$$F = \frac{k_B T}{P} \left[\frac{1}{4 \left(1 - \frac{\Delta X}{L_c} + \frac{F}{K} \right)^2} - \frac{1}{4} + \frac{\Delta X}{L_c} - \frac{F}{K} \right] \quad (2.1)$$

Here k_B is Boltzmann's constant, T is the absolute temperature, P is persistent length, L_c is the contour length, K is stretch modulus, ΔX is the end-to-end distance and F is the stretching force. The L_c , P and K of single-stranded RNA are 0.59 nm per nucleotide, 1 nm, and 1500 pN, respectively.^{56,60,63,100}

To further characterize how the single mutations and pH affect mechanical (un)folding kinetics, we used a constant-force mode with the force maintained at a pre-set value by adjusting the pipette bead position (**Figures 2.3a and 2.6**).^{57,64} In a constant-force mode, we observed that at 16.7 pN, rHP has the folded state more populated, and the population distribution and (un)folding dynamics essentially show no change upon increasing the pH from 6.5 to 8.0 (**Figure 2.4a**), suggesting that varying pH does not affect the folding free energy landscape of rHP, consistent with the force-ramp results (**Figure 2.3b**). A relatively lower force (14.9 pN) is needed to obtain a similar population distribution at pH 6.5 for the three hairpin mutants with single mutations (**Figure 2.4b–d**), which clearly suggests that the single U-to-C mutations destabilize the hairpins. Upon varying the pH from 6.5 to 8.0, the population of the unfolded state (with a larger extension) increases for all three mutants (**Figure 2.4b–d**). Apparently, increasing pH lowers the mechanical stability (unfolding force) of the three mutants (U22C, U25C and U29C). The one-step mechanical unfolding pathway of hairpin U22C at 22 °C (**Figure 2.3c**) is in contrast to a two-step unfolding pathway observed in the thermal unfolding (temperature-ramp, 0.5 °C/min) experiment (**Figure 2.2a,d**). The relatively less stable top part of the stem of

hairpin U22C thus thermally melts at a relatively low temperature (~ 53 °C, **Figure 2.2a,d**).

In addition, we used a passive mode to study the mechanical (un)folding kinetics, with the pipette bead kept at a fixed position, allowing the bead in the optical trap to move upon RNA (un)folding (**Figures 2.3a and 2.7**).⁶⁴ There is no electronic feedback in passive mode, which makes it suitable for measuring relatively fast kinetics.^{64,88} In the real-time (un)folding traces of passive mode, the high force and low force states correspond to a folded state and an unfolded state, respectively (**Figure 2.7**). The passive-mode and constant-force mode data give essentially the same kinetic parameters for various molecular tethers (**Figures 2.7 and 2.8**).

In a constant-force mode, the force was maintained at a constant value near transition force range which can be estimated through the rupture force in force-ramp curves. The constant-force traces directly reveal one-step (un)folding of the studied hairpin structures, with higher extension trajectory representing the unfolded states while lower extension trajectory representing the folded states (see **Figure 2.4**). The end-to-end extension change (ΔX) were obtained from the extension value distributions fit to Gaussian functions.⁶⁴ The extension change can be well estimated by an extensible wormlike chain (EWLC) model (Equation 2.1).⁹⁸⁻¹⁰¹

The ΔX values can reveal if the RNAs form the designed native structures. The ΔX value (~ 12.5 nm) of rHP, U22C, and U25C does not vary with pH (**Figures**

2.3b and 2.4a–c, Table 2.2), which, based on EWLC model, corresponds to about 33 nucleotides, which is close to the 32 nucleotides present in all hairpin constructs. Interestingly, hairpin U29C unfolds at below 16 pN with a significantly shortened ΔX of ~10 nm at pH 8.0 (**Figures 2.3b and 2.4d, Table 2.2**). Furthermore, the ΔX of U29C decreases with increasing pH: ~12 nm at pH 5.0, ~11 nm at pH 6.5, and ~10 nm at pH 7.3 and 8.0 (**Table 2.2**), suggesting that the bottom part of the stem of U29C is increasingly destabilized upon increasing pH. Even at pH 5.0, ΔX of U29C is still about 0.5 nm shorter than that expected for a 32-mer hairpin. Thus, at pH 7.3 and pH 8.0, there are 24 nucleotides folded for U29C hairpin, as calculated based on EWLC model (with the inter-nucleotide distance considered as 0.44 nm at 15 pN) (**Figures 2.1b and 2.3d**). The results suggest that substitution of the A4-U29 pair with A4·C29 pair in hairpin U29C causes the destabilization of the bottom part of the stem, resulting in the bottom base pairs unable to fully form in the presence of the mechanical stretching force applied on the terminal ends.^{103,104} Consistently, increasing pH results in an increasing number of base pairs unable to form in U29C hairpin, presumably due to the increased dynamics upon the loss of a hydrogen bond in the wobble A4·C29 pair.

We observed relatively fast unfolding and folding dynamics (designated as hopping) for rHP, U25C, and U29C (**Figures 2.3b and 2.4**). However, hairpin U22C shows significantly slower (un)folding dynamics, presumably due to the fact that the A11-U22 to A11·C22 mutation results in a decreased folding rate with the destabilization mutation near the hairpin loop. It has been previously observed in thermal and mechanical (un)folding studies that destabilizing

mutations of the base pairs near or closing the hairpin loop decrease the folding rate of nucleic acid hairpins.^{53-55,57,69,105,106} The highly dynamic unfolding and folding processes (hopping) as observed in the FECs (**Figure 2.3b**) make it hard to quantitatively measure the transition forces and the effect of pH.

Figure 2.4 . Representative traces of mechanical (un)folding at a constant-force mode at pH 6.5 and 8.0. Only the extension curves are shown. The curves are smoothed to 200 Hz. The extension distributions are fit to two-peak Gaussian distributions (binned to 1 nm). A one-step mechanical unfolding transition is observed. The dashed cyan and green lines at the peaks of the Gaussian functions correspond to the unfolded states and folded states, respectively. (a) rHP at 16.7 pN. rHP folding and unfolding rates show no significant pH dependence. (b) U22C at 14.9 pN. U22C folding rate decreases with increasing pH. (c) U25C at 14.9 pN. With increasing pH, U25C folding rate decreases, whereas unfolding rate increases. (d) U29C at 14.9 pN. U29C shows a decreased extension change

at pH 8.0. No significant difference in kinetics was observed between pH 6.5 and pH 8.0. However, unfolding rate significantly decreases at pH 5.0 (see **Figure 2.5d**).

Figure 2.5. Force-dependent unfolding (FU, open markers) and folding (UF, filled markers) rates at varied pH's obtained by constant-force mode. The critical forces (F_c) of the hairpins were obtained based on the crossing points of the force-dependent unfolding and refolding rates curves. The R^2 values of the linear fits range from 0.871 to 0.999. (a) rHP. No significant pH dependence is observed for rHP folding and unfolding rates. (b) U22C. pH dependence is observed for U22C folding but not unfolding. (c) U25C. pH dependence is observed for both unfolding and folding. (d) U29C. Unfolding rate is significantly slowed at pH 5.0.

The quantitative analysis of the pH- and force-dependent (un)folding kinetics was done using the constant-force and passive mode data. The extension versus time trajectories in a constant-force mode (**Figure 2.4**) are fit to a hidden Markov model¹⁰⁷ to extract the lifetimes of the unfolded and folded states at each fixed force (**Figure 2.6a**). A single exponential fit is then applied to extract the mean lifetime of the folded state (τ_{folded} or $k_{\text{unfolding}}$) and the unfolded state (τ_{unfolded} or k_{folding}), respectively (**Figure 2.6b**). To extract the mechanical unfolding and folding transition state positions (**Figure 2.5 and Appendix Table A1.1**), the force-dependent rates are further fit to the Bell model (Equation 2.2).^{64,108}

$$\ln k(F) = \ln(k_0) + \frac{F\Delta X^\ddagger}{k_B T} \quad (2.2)$$

Here, k_0 is the rate coefficient at zero force, $\Delta X^\ddagger$ is the distance from unfolded or folded state to the transition state, k_B is Boltzmann constant, and T is absolute temperature. The force trajectories in passive mode (**Figure 2.7**) at each force were also fit to hidden Markov model¹⁰⁷ to extract the lifetimes of the mechanically unfolded and folded states. Overall, the constant-force mode data (**Appendix Table A1.1**) and passive mode data (**Appendix Table A1.2**) give essentially the same kinetic parameters.

Figure 2.6. Extraction of (un)folding rates from constant-force data. (a) Representative extension versus time curve (blue) at 16.7 pN for rHP at pH 7.3

by constant-force mode, and fitting trajectory to hidden Markov model¹⁰⁷ (black).

(b) Histogram of lifetime distributions of folded state and unfolded state fit to single exponential equation based on one-step reaction model for extracting unfolding and folding rates, respectively. The mean lifetime of the folded state (τ_{folded}) and the unfolded state (τ_{unfolded}) obtained are 0.192 ± 0.006 s ($R^2 = 0.995$) and 0.341 ± 0.004 s ($R^2 = 0.999$), respectively. The uncertainties are reported as 95% confidence interval for the single-exponential fit.

Figure 2.7. Representative passive-mode traces observed for (a) rHP, (b) U22C,

(c) U25C, and (d) U29C at pH 6.5 and pH 8.0. The data bandwidth is 1000 Hz. The force (binned to 0.05 pN) and extension (binned to 1 nm) distributions are fit to Gaussian functions.

Two individual tethers of U22C and U29C were measured at constant-force mode. The force-dependent kinetics data obtained from two tethers were pooled together and analyzed. The unfolding and refolding rates obtained from different tethers of the same sample showed no significant difference. The pooled results revealed the same kinetic parameters as those from the individual tethers. Two individual tethers of U29C were measured at passive mode. Similarly, the force-dependent kinetics obtained from the two tethers were pooled together and analyzed. The pooled results of the passive-mode data again showed no significant difference compared with the individual tethers (**Figure 2.8**).

Figure 2.8. Comparison of force-dependent unfolding and folding rates measured for different tethers. (a,b) U22C and U29C studied by constant-force mode. (c) U29C studied by passive mode. Unfolding and refolding rates obtained from the pooled and individual tethers show no significant difference. The R^2 values of the linear fits range from 0.871 to 0.994.

The critical force F_c is defined as the force at which the unfolding rate equals to the folding rate. The critical forces of the hairpins were obtained based on the crossing points of the force-dependent unfolding and refolding rates curves (**Figure 2.5**). At all four pH's (5.0, 6.5, 7.3 and 8.0), the critical forces of rHP

(16.7, 17.2, 16.5, and 16.8 pN) are higher than U22C (15.4, 15.2, 14.7, and 14.5 pN), U25C (16.4, 15.4, 14.7, and 14.4 pN), and U29C (16.2, 15.3, 15.0, and 15.1 pN) (**Figure 2.9a, Table 2.2 and Appendix Table A1.1**). Clearly, a single A-U to A·C substitution destabilizes the mechanical stability of the hairpin.

The critical force of rHP remains at around 17 pN with a small change when pH is varied (**Figures 2.5a and 2.9a, Table 2.2**), suggesting that the mechanical stability of rHP is independent of pH. The critical forces of the three mutants, however, increase by over 1 pN when pH is lowered from 8.0 to 5.0 (**Figures 2.5b–d and 2.9a, Table 2.2**). At pH 8.0, the critical forces of U22C, U25C and U29C are significantly lower (by 2.3, 2.4 and 1.7 pN) than that of rHP. However, at pH 5.0, the critical forces of U25C and U29C are only slightly lower (by 0.3 and 0.5 pN, respectively) than that of rHP. Our data suggest that single-molecule mechanical (un)folding can reveal the stabilizing effect of a single proton binding at the N1 nitrogen atom of adenine in a wobble A·C pair to form a wobble A⁺·C pair with two hydrogen bonds (**Figure 2.1a**) in hairpins U22C, U25C and U29C.^{16,18,37,38}

Table 2.2. Parameters for mechanical (un)folding by constant-force experiments.^a

RNA molecule	pH	ΔX_c (nm)	# of folded nt ^c	F_c (pN)	W_{Fc} (kcal mol ⁻¹)	ΔG_o (kcal mol ⁻¹)
rHP	5.0	12.5 ± 1.4	32.8 ± 3.2	16.7 ± 0.1	30.0 ± 3.4	21.1 ± 3.3
	6.5	12.8 ± 1.1	33.4 ± 2.5	17.2 ± 0.1	31.5 ± 2.7	22.4 ± 2.7
	7.3	12.7 ± 1.1	33.3 ± 2.5	16.5 ± 0.1	30.2 ± 2.6	21.3 ± 2.6
	8.0	12.7 ± 1.1	33.3 ± 2.5	16.8 ± 0.1	30.7 ± 2.7	21.7 ± 2.7
U22C	5.0	12.6 ± 1.2	33.7 ± 2.8	15.4 ± 0.1	27.9 ± 2.7	19.5 ± 2.7
	6.5	12.4 ± 1.0	33.2 ± 2.3	15.2 ± 0.1	27.3 ± 2.2	19.0 ± 2.2
	7.3	12.5 ± 1.1	33.7 ± 2.6	14.7 ± 0.2	26.4 ± 2.3	18.3 ± 2.3
	8.0	12.6 ± 1.1	34.2 ± 2.6	14.5 ± 0.1	26.3 ± 2.3	18.3 ± 2.3
U25C	5.0	12.4 ± 1.0	32.7 ± 2.3	16.4 ± 0.2	29.3 ± 2.4	20.4 ± 2.4

	6.5	12.6 ± 1.1	33.7 ± 2.5	15.4 ± 0.1	27.9 ± 2.4	19.5 ± 2.4
	7.3	12.5 ± 1.0	33.8 ± 2.3	14.7 ± 0.1	26.5 ± 2.1	18.4 ± 2.1
	8.0	12.7 ± 1.0	34.5 ± 2.3	14.4 ± 0.1	26.1 ± 2.1	18.1 ± 2.1
U29C	5.0	12.2 ± 1.2	32.5 ± 2.7	16.2 ± 0.1	28.4 ± 2.8	19.7 ± 2.8
	6.5	11.1 ± 1.3	30.3 ± 3.0	15.3 ± 0.2	24.4 ± 2.9 ^b	16.1 ± 2.9
	7.3	10.2 ± 1.2	28.4 ± 2.8	15.0 ± 0.2	22.0 ± 2.6 ^b	13.8 ± 2.6
	8.0	9.7 ± 1.0	27.1 ± 2.3	15.1 ± 0.1	21.1 ± 2.2 ^b	12.8 ± 2.2

^a The critical force F_c is defined as the force at which the unfolding and folding rates are equal. W_{Fc} is the work for unfolding reaction at equilibrium at critical force ($W_{Fc} = N_A F_c \Delta X_c$). N_A is Avogadro constant. ΔG_0 is the free energy for unfolding reaction at zero force at 22 °C ($\Delta G_0 = W_{Fc} - W_{stretching}$). $W_{stretching}$ is the work for stretching single-stranded RNA from zero force to critical force at equilibrium, which can be calculated through the integration of the EWLC curve (Equation 2.1). All uncertainties are reported as standard errors.

^b The equilibrium unfolding work (W_{Fc}) decreases with increasing pH for all hairpins except for rHP. The W_{Fc} value for U29C decreases more significantly with increasing pH, mainly due to the decreasing ΔX_c value with the bottom base pairs not forming (see **Figure 2.3d**). The F_c values decrease with increasing pH for all hairpins except rHP.

^c The number of nucleotides folded are calculated based on the measured extension change (ΔX_c) values, with the diameter of RNA helix (2 nm) considered. The number of folded nucleotides of U29C decreases with increasing pH.

The critical force of U22C is about 1 pN lower than U25C and U29C at pH 5.0, which might be due to the fact that the A·C mismatch in U22C is located at a position adjacent to the hairpin loop and affects mechanical folding significantly. In addition, the A·C mismatch in U22C is flanked by two A-U pairs, whereas U25C and U29C have their A·C mismatch flanked by two G-C pairs. The nearest-neighbor flanking base pairs may affect the apparent pK_a differently.^{9,16,18,27,52} We expect that destabilizing mutations near the loop including the loop-closing base pair may generally destabilize a hairpin by decreasing mechanical folding rate, and the folding rate can be rescued if a potential protonation site exists to stabilize the base pair.

Detailed force-dependent kinetic studies reveal that proton binding affects the mechanical (un)folding kinetics of the A·C pair-containing hairpins in a position-dependent manner (**Figure 2.5**). Mechanical folding rate of hairpin U22C (with an A·C pair near the hairpin loop (A11·C22)) increases with decreasing pH, with

the mechanical unfolding rate relatively less dependent on pH. In contrast, mechanical unfolding rate of hairpin U29C (with an A·C pair near the terminus of the stem (A4·C29)) decreases with decreasing pH, with the mechanical folding rate essentially unchanged. Hairpin U25C (with an A·C pair at the middle of the stem (A8·C25)) shows pH dependence for both mechanical unfolding and folding reactions. Thus, proton binding facilitates the rapid formation of an A⁺·C pair and enhances the apparent mechanical stability of an RNA stem-loop structure through the increased folding rate, or decreased unfolding rate, or both, depending on the location of the A·C pair in the hairpin.

To further understand the mechanical (un)folding mechanisms and the effects of protonation, we extracted the apparent two-state mechanical unfolding and folding transition state positions from the force-dependent (un)folding studies (Equation 2.2). The $\Delta X^\ddagger$ (or $\Delta n^\ddagger$) values correlate with the slopes of the rate versus force curves, and indicate the mechanical (un)folding transition state positions. It appears that the $\Delta X^\ddagger$ values or number of nucleotides disrupted and folded at the unfolding transition state ($\Delta n^{\ddagger, \text{FU}}$) and the folding transition state ($\Delta n^{\ddagger, \text{UF}}$) are about the same for rHP, U22C and U25C, and remain unchanged with varying pH (**Figures 2.1b, 2.5 and 2.9b, Appendix Tables A1.1 and A1.2**). At the apparent unfolding and folding transition state, hairpins rHP, U22C and U25C have their bottom 9 base pairs disrupted (**Figure 2.1b**). The rapid exchange of protonation and deprotonation of an A·C pair (on the μs time)^{18,109,110} at the apparent two-state transition state allows us to observe the averaged effect of dynamic (de)protonation on mechanical (un)folding kinetics. The transition state positions revealed are consistent with the fact that U22C folding rate and U29C

unfolding rate are most sensitive to pH, respectively. For U25C, the A⁺·C pair is located adjacent to the apparent transition state positions, resulting in both unfolding and folding rates of U25C being sensitive to proton binding.

Figure 2.9. Effects of pH on mechanical (un)folding. All errors are reported as standard errors. (a) pH dependence of critical forces. The data were obtained from constant-force mode (solid line) and passive mode (dashed line). Except for rHP, the critical forces decrease with increasing pH. (b) Number of nucleotides disrupted and folded at unfolding transition states ($\Delta n^{\ddagger, FU}$) and folding transition states ($\Delta n^{\ddagger, UF}$), respectively. The data were obtained by averaging the constant-force and passive mode data. The transition state positions show no significant pH dependence. (c) pH dependence of the total number of nucleotides in folded hairpins based on the measured ΔX values (see **Table 2.2**). U29C shows decreasing number of folded nucleotides with increasing pH. (d) pH dependence

of mechanical unfolding Gibbs free energy at zero force at 22 °C (ΔG_0). Except for rHP, the ΔG_0 values decrease with increasing pH.

At pH 5.0, hairpin U29C has the bottom part of the stem nearly fully formed (**Figures 2.1b and 2.9c**), and the force-dependent kinetic studies (**Figure 2.5d**) suggest that the apparent unfolding and folding transition states (Equation 2.2) of U29C are the same as rHP. As discussed above, the bottom base pairs of hairpin U29C are not fully formed with an increased pH. However, the slopes of the rate versus force curves for hairpin U29C at pH 8.0 are largely the same as those at pH 5.0 (**Figure 2.5d**). We can conclude that the apparent folding transition state position of hairpin U29C is the same as the other three hairpins (around the A10-U23 pair, see **Figure 2.1b**), and is independent of pH. However, the apparent unfolding transition state position of hairpin U29C instead shifts up by two base pairs at a relatively high pH (7.3 and 8.0), i.e., the bottom 11 base pairs are disrupted at the unfolding transition state (**Figure 2.1b**). Consistently, unfolding rate of U29C changes significantly upon varying pH from 5.0 to 6.5, but not as much from pH 6.5 to 8.0 (**Figure 2.5d**).

The Gibbs free energy change for unzipping a hairpin at zero force ΔG_0 is obtained by subtracting the work of the mechanically unfolding hairpin by the work of stretching the single-stranded RNA from zero force to the critical force (**Table 2.2**)^{13,60}. Similar to the trend observed for critical forces, ΔG_0 of rHP is the highest at all pH's and remains at $\sim 22 \text{ kcal}\cdot\text{mol}^{-1}$, confirming that a single A-U to A-C replacement can decrease the thermodynamic stability of the hairpin. The ΔG_0 values of U22C, U25C, and U29C decrease gradually with pH

increasing from 5.0 to 8.0 (**Figure 2.9d**). It is of note that, compared to U22C and U25C, U29C shows a more significant drop in ΔG_0 , which is mainly due to the fact that several bottom base pairs in U29C fail to form when pH is above 5.0.

2.3.3 Minus-one Ribosomal Frameshifting Efficiency Positively Correlates with Critical Force and Folding Rate of the mRNA Hairpins

Our mechanical (un)folding results show that the mechanical forces needed for unfolding the studied RNA hairpin structures range from 14–18 pN, which is in the same range of the force exerted by translating ribosomes (12–20 pN).^{72,73} Thus, mRNA structures could affect the ribosome translation by serving as mechanical and kinetic barriers.¹¹¹ Hence, it is important to unravel the potential correlations between mechanical (un)folding parameters of the studied RNA hairpins and ribosomal translation. We carried out *in vitro* translation assay using a dual-luciferase reporter system to characterize how A-U to A-C substitutions in an mRNA containing the 32-mer hairpins (**Figure 2.10a**) may affect –1 ribosomal frameshifting efficiency.⁸⁷ We observed –1 frameshifting efficiencies of $2.2 \pm 0.2\%$, $1.8 \pm 0.1\%$, $1.2 \pm 0.1\%$, and $1.0 \pm 0.1\%$, respectively, for rHP, U29C, U25C, and U22C (**Figure 2.10b,c**). The frameshifting efficiency for rHP is consistent with our previously reported value.⁸⁷ We obtained the frameshifting efficiencies with relatively small errors by rigorously controlling the timing of the luminescence measurement.⁸⁷

Figure 2.10. Comparison of frameshifting efficiency at 30 °C and mechanical (un)folding parameters at 22 °C. (a) Experimental constructs for *in vitro* frameshifting assay. The slippery sequence (U UUU UUA) is shown in red and underlined. Minus-one ribosomal frameshifting results in the slippery sequence to be decoded as UUU UUU A instead of U UUU UUA, resulting in the expression of downstream FLuc at -1 frame in addition to RLuc at normal zero frame. The single-stranded spacer (GGG, brown) is located between the slippery sequence and the hairpin structure (blue). (b) Frameshifting efficiency increases with increasing critical force. The critical forces shown were averaged values based on the constant-force mode and passive mode data at pH 7.3. (c) Frameshifting efficiency increases with increasing folding rate at 15 pN. No clear

correlation is observed at forces far away from 15 pN. Errors for the critical forces are standard errors. Errors for frameshifting efficiency values are standard deviations based on three independent measurements. The mRNAs are at 0.1 μ M with the translation mixture at pH 7.2–7.6. (d) Schematics of a translating ribosome positioned at various sites of an mRNA. Multiple ribosomes may simultaneously translate one mRNA. (Top) Ribosome positioned upstream of the slippery site, with the mRNA hairpin structure unaffected. (Middle) Ribosome positioned at the slippery site, with the bottom four base pairs of the mRNA hairpin disrupted. (Bottom) Ribosome positioned downstream of the slippery site at -1 frame, with mRNA hairpin completely disrupted.

It is interesting to note that U29C and U22C show the highest and lowest frameshifting efficiencies among the three mutated hairpins. The data suggest that, for the hairpins studied, the local stability of the top part of the stem is more important than the bottom part for stimulating frameshifting. It is known that a 5–8 nucleotide single-stranded spacer between the slippery site and downstream RNA structure is optimal for stimulating frameshifting.^{74,79} Thus, with a 3-nt single-stranded spacer (**Figure 2.10a**), the bottom 2–5 base pairs of the hairpins may be pre-melted by ribosome at the slippery site. The pre-melting of the RNA structure could be due to the ribosome biasing the thermal fluctuations towards the unfolded state or the mechanical force generated during ribosome translocation.¹¹¹ Thus, we propose that the bottom four base pairs of the hairpins studied here are pre-melted at the mRNA entry site when ribosome is initially paused at the slippery site, followed by consequent unzipping of the rest of the hairpin structure (**Figure 2.10a,d**). Hence, it is reasonable that U29C shows a

minimal reduction in frameshifting efficiency compared to rHP. Consistently, our single-molecule mechanical pulling data suggest that, on average, the bottom 2–4 base pairs of U29C do not form under mechanical stretching force at pH 7.3–8.0. We observed a correlation between critical force (which is determined by both unfolding and folding rates, see **Figure 2.5**) and frameshifting efficiency (**Figure 2.10b**). Thus, for the hairpins studied, mechanical (un)folding may mimic the mRNA structure (un)folding in the presence of translating ribosomes.

U22C has a very similar frameshifting efficiency compared to U25C, which seems to be counter-intuitive, because base pairs closer to mRNA entry site of ribosome at the slippery site is expected to have a more significant effect in affecting frameshifting efficiency.^{79,80} It is likely that, in addition to local stem stability, folding rate may also affect frameshifting significantly.⁸⁵ Timely folding of the downstream mRNA structure may be critical for pausing a translating ribosome and stimulating frameshifting in consideration of the fact that multiple ribosomes translate one mRNA. U22C shows a significantly lower folding rate compared to U25C, which may result in a similar frameshifting efficiency for U22C and U25C, even though the mismatch for U22C is located further away from the slippery site. We indeed observed a correlation between frameshifting efficiency and folding rate at 15 pN (**Figure 2.10c**), with the force well within the range that can be exerted by a translating ribosome.^{72,73} No strong correlation was observed between frameshifting efficiency and folding rate at other forces. It is likely that the magnitude of such steric hindrance-induced effect is dependent on the density of ribosomes on mRNAs. It is possible that for a relatively large mRNA in the presence of multiple translating ribosomes,¹¹² the

folding of a hairpin structure may experience mechanical resistance in a relatively crowded environment of cell lysate and steric hindrance of translating ribosomes at and adjacent to the hairpin sequence (**Figure 2.10d**). We hypothesize that steric and crowding effects may be mimicked by the stretching force exerted in optical tweezers experiments.

For the hairpins studied, surprisingly, no apparent correlation is observed between frameshifting efficiency and unfolding rate (**data not shown**). The results further suggest that, folding rate may also affect frameshifting for the hairpins studied. In addition, the bottom stem of U29C is not fully formed at $\text{pH} > 5.0$, as revealed by our pulling experiments. Taken together, our data suggest that mechanical (un)folding of RNA hairpins may mimic how RNAs unfold and fold in the presence of translating ribosomes.

The thermal melting pathway of U22C is not consistent with the directional mRNA unwinding process of a ribosome. Thus, thermal melting data may not be meaningfully compared with the frameshifting data. Taken together, our results suggest that, for the hairpins studied, incorporating an A·C mismatch may reduce the mechanical stability of the stem and decrease the folding rate, hence attenuating -1 ribosomal frameshifting efficiency.

2.4 Conclusions

In summary, compared to the traditional ensemble thermal melting method, single-molecule mechanical (un)folding studies can reveal the detailed real-time folding dynamics at near physiological temperature and facilitate a deeper

understanding of how a single A-U to A·C mutation and single proton binding may affect the stabilities, (un)folding dynamics, and transition states of RNA in a position-dependent manner. For example, although our ensemble thermal melting studies show that a single A-U to A·C replacement lowers the thermal stability, the detailed dynamics and kinetic parameters and the effects of protonation, however, may not be easily obtained by pH-dependent thermal melting studies likely due to the facts that (i) relatively large hairpins may show non-two-state thermal melting transitions, (ii) the pK_a values for an A·C pair is temperature dependent, and (iii) non-native structures may form during thermal melting. Our single-molecule mechanical (un)folding studies at a constant temperature (room temperature) clearly suggest that varying proton concentration from 10 nM (pH 8.0) to 10 μ M (pH 5.0) can specifically enhance the formation of a wobble $A^+ \cdot C$ pair and in turn the mechanical stability of an $A^+ \cdot C$ pair-containing hairpin.

The mechanical unfolding forces for U22C, U25C, and U29C increase gradually with decreasing pH (**Figure 2.9a**), consistent with the fact that the protonation of the A·C mismatch is relatively fast. The exchange of protonation and deprotonation of a wobble A·C pair is on the μ s time scale.^{18,109,110} Thus, in our optical tweezers experiments with a time resolution of ms, we observed an averaged effect of dynamic (de)protonation on the stability of the apparent two-state transition state and mechanical (un)folding kinetics. The proton binding effect is similar to the salt concentration dependence as observed before.^{69,113,114} However, proton binding is a site-specific binding and is sequence and structure specific as observed for some small molecules, oligonucleotides, and

proteins.^{56,115-120} It may be expected that, with more extensive studies in an expanded pH range, the apparent pK_a values may be extracted¹⁹ from the single-molecule mechanical unfolding studies.

Our results have important implications in how a wobble A·C pair and other structures may affect RNA (un)folding dynamics and function at near physiological conditions.^{29,31,32} For example, compared to RNA thermal unfolding, RNA unfolding by mechanical stretching force may better mimic RNA (un)folding dynamics in the presence of RNA helicases (such as ribosome). Our *in vitro* translation data for the mRNAs containing the hairpins suggest that mechanical unfolding force at equilibrium (critical force) and folding rate at around 15 pN of the hairpins positively correlate with -1 ribosomal frameshifting efficiency. For the mRNA hairpins studied, mechanical unfolding rate shows no correlation with frameshifting efficiency (**data not shown**), probably because U29C hairpin is partially unfolded during pulling experiments. Indeed, with a 3-nt single-stranded spacer between the slippery site and downstream hairpin (GGG, **Figure 2.10a**), two to five base pairs of the bottom part of the stem of the hairpins are predicted to be melted by ribosome at the slippery site (**Figure 2.10d**). It is expected that the length of the single-stranded spacer is critical in determining the correlation between mechanical unfolding parameters and frameshifting efficiency. For example, the observed correlations (**Figure 2.10b,c**) are not linear correlations, which might be specific to the studied hairpins with a relatively short 3-nt single-stranded spacer.

U22C and U25C have a similar frameshifting efficiency (~1%), indicating that both unfolding and folding rates of an mRNA hairpin may affect frameshifting. It is likely that for a relatively large mRNA in the presence of multiple translating ribosomes, the folding of a hairpin structure slows down due to the steric hindrance of translating ribosomes¹¹² at and adjacent to the hairpin sequence (see **Figure 2.10d**) and mechanical resistance in a relatively crowded environment of cell lysate. Thus, RNA (un)folding under mechanical stretching force provides unique insights into how RNA structural stability and dynamics involving A·C pairs and other non-canonical pairs¹²¹ may regulate RNA-acting molecular motors such as ribosome.

References

- (1) T. R. Cech, J. A. Steitz, The Noncoding RNA Revolution-Trashing Old Rules to Forge New Ones. *Cell* **2014**, *157*, 77-94.
- (2) N. B. Leontis, A. Lescaute, E. Westhof, The building blocks and motifs of RNA architecture. *Curr Opin Struct Biol* **2006**, *16*, 279-287.
- (3) D. H. Mathews, M. D. Disney, J. L. Childs, S. J. Schroeder, M. Zuker, D. H. Turner, Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc Natl Acad Sci U S A* **2004**, *101*, 7287-7292.
- (4) I. Tinoco, Jr., C. Bustamante, How RNA folds. *J Mol Biol* **1999**, *293*, 271-281.
- (5) J. L. Chen, A. L. Dishler, S. D. Kennedy, I. Yildirim, B. Liu, D. H. Turner, M. J. Serra, Testing the Nearest Neighbor Model for Canonical RNA Base Pairs: Revision of GU Parameters. *Biochemistry* **2012**, *51*, 3508-3522.
- (6) T. Xia, J. SantaLucia, Jr., M. E. Burkard, R. Kierzek, S. J. Schroeder, X. Jiao, C. Cox, D. H. Turner, Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry* **1998**, *37*, 14719-14735.
- (7) R. Kierzek, M. E. Burkard, D. H. Turner, Thermodynamics of single mismatches in RNA duplexes. *Biochemistry* **1999**, *38*, 14214-14223.
- (8) M. Wu, J. A. Mcdowell, D. H. Turner, A Periodic-Table of Symmetrical Tandem Mismatches in RNA. *Biochemistry* **1995**, *34*, 3204-3211.
- (9) G. Chen, S. D. Kennedy, D. H. Turner, A CA(+) pair adjacent to a sheared GA or AA pair stabilizes size-symmetric RNA internal loops. *Biochemistry* **2009**, *48*, 5738-5752.
- (10) Y. V. Lerman, S. D. Kennedy, N. Shankar, M. Parisien, F. Major, D. H. Turner, NMR structure of a 4 x 4 nucleotide RNA internal loop from an R2 retrotransposon: identification of a three purine-purine sheared pair motif and comparison to MC-SYM predictions. *RNA* **2011**, *17*, 1664-1677.
- (11) C. A. Theimer, L. D. Finger, L. Trantirek, J. Feigon, Mutations linked to dyskeratosis congenita cause changes in the structural equilibrium in telomerase RNA. *Proc Natl Acad Sci U S A* **2003**, *100*, 449-454.
- (12) A. R. Davis, C. C. Kirkpatrick, B. M. Znosko, Structural characterization of naturally occurring RNA single mismatches. *Nucleic Acids Res* **2011**, *39*, 1081-1094.
- (13) Z. Zhong, L. H. Soh, M. H. Lim, G. Chen, A U·U Pair-to-U·C Pair Mutation-Induced RNA Native Structure Destabilisation and Stretching-Force-Induced RNA Misfolding. *ChemPlusChem* **2015**, *80*, 1267-1278.
- (14) A. Phan, K. Mailey, J. Saeki, X. Gu, S. J. Schroeder, Advancing viral RNA structure prediction: measuring the thermodynamics of pyrimidine-rich internal loops. *RNA* **2017**, *23*, 770-781.
- (15) M. Chawla, P. Sharma, S. Halder, D. Bhattacharyya, A. Mitra, Protonation of base pairs in RNA: context analysis and quantum chemical investigations of their geometries and stabilities. *J Phys Chem B* **2011**, *115*, 1469-1484.
- (16) N. A. Siegfried, B. O'Hare, P. C. Bevilacqua, Driving forces for nucleic acid pKa shifting in an A(+).C wobble: effects of helix position, temperature, and ionic strength. *Biochemistry* **2010**, *49*, 3225-3236.

- (17) C. L. Tang, E. Alexov, A. M. Pyle, B. Honig, Calculation of pK_as in RNA: on the structural origins and functional roles of protonated nucleotides. *J Mol Biol* **2007**, *366*, 1475-1496.
- (18) P. Legault, A. Pardi, Unusual dynamics and pK_a shift at the active site of a lead-dependent ribozyme. *J Am Chem Soc* **1997**, *119*, 6621-6628.
- (19) P. C. Bevilacqua, T. S. Brown, D. Chadalavada, J. Lecomte, E. Moody, S. I. Nakano, Linkage between proton binding and folding in RNA: implications for RNA catalysis. *Biochem Soc Trans* **2005**, *33*, 466-470.
- (20) M. L. Gleghorn, J. Zhao, D. H. Turner, L. E. Maquat, Crystal structure of a poly(rA) staggered zipper at acidic pH: evidence that adenine N1 protonation mediates parallel double helix formation. *Nucleic Acids Res* **2016**, *44*, 8417-8424.
- (21) N. Safaee, A. M. Noronha, D. Rodionov, G. Kozlov, C. J. Wilds, G. M. Sheldrick, K. Gehring, Structure of the parallel duplex of poly(A) RNA: evaluation of a 50 year-old prediction. *Angew Chem Int Ed Engl* **2013**, *52*, 10370-10373.
- (22) G. B. Goh, J. L. Knight, C. L. Brooks, 3rd, pH-dependent dynamics of complex RNA macromolecules. *J Chem Theory Comput* **2013**, *9*, 935-943.
- (23) B. E. Allred, M. Gebala, D. Herschlag, Determination of Ion Atmosphere Effects on the Nucleic Acid Electrostatic Potential and Ligand Association Using AH⁺.C Wobble Formation in Double-Stranded DNA. *J Am Chem Soc* **2017**, *139*, 7540-7548.
- (24) J. L. Wilcox, A. K. Ahluwalia, P. C. Bevilacqua, Charged Nucleobases and Their Potential for RNA Catalysis. *Acc Chem Res* **2011**, *44*, 1270-1279.
- (25) M. A. Ditzler, J. Sponer, N. G. Walter, Molecular dynamics suggest multifunctionality of an adenine imino group in acid-base catalysis of the hairpin ribozyme. *RNA* **2009**, *15*, 560-575.
- (26) N. B. Suslov, S. DasGupta, H. Huang, J. R. Fuller, D. M. J. Lilley, P. A. Rice, J. A. Piccirilli, Crystal Structure of the VS ribozyme. *Nat Chem Biol* **2015**, *11*, 840-846.
- (27) A. C. Wolter, A. K. Weickmann, A. H. Nasiri, K. Hantke, O. Ohlenschlager, C. H. Wunderlich, C. Kreutz, E. Duchardt-Ferner, J. Wohnert, A Stably Protonated Adenine Nucleotide with a Highly Shifted pK_a Value Stabilizes the Tertiary Structure of a GTP-Binding RNA Aptamer. *Angew Chem Int Ed Engl* **2017**, *56*, 401-404.
- (28) T. Tran, M. D. Disney, Molecular recognition of 6'-N-5-hexynoate kanamycin A and RNA 1x1 internal loops containing CA mismatches. *Biochemistry* **2011**, *50*, 962-969.
- (29) H. H. Bink, K. Hellendoorn, J. van der Meulen, C. W. Pleij, Protonation of non-Watson-Crick base pairs and encapsidation of turnip yellow mosaic virus RNA. *Proc Natl Acad Sci U S A* **2002**, *99*, 13465-13470.
- (30) P. Blancafort, S. V. Steinberg, B. Paquin, R. Klinck, J. K. Scott, R. Cedergren, The recognition of a noncanonical RNA base pair by a zinc finger protein. *Chem Biol* **1999**, *6*, 585-597.
- (31) K. Hellendoorn, P. J. Michiels, R. Buitenhuis, C. W. Pleij, Protonatable hairpins are conserved in the 5'-untranslated region of tymovirus RNAs. *Nucleic Acids Res* **1996**, *24*, 4910-4917.
- (32) B. Houck-Loomis, M. A. Durney, C. Salguero, N. Shankar, J. M. Nagle, S. P. Goff, V. M. D'Souza, An equilibrium-dependent retroviral mRNA switch regulates translational recoding. *Nature* **2011**, *480*, 561-564.

- (33) A. Rozov, E. Westhof, M. Yusupov, G. Yusupova, The ribosome prohibits the G*U wobble geometry at the first position of the codon-anticodon helix. *Nucleic Acids Res* **2016**, *44*, 6434-6441.
- (34) N. B. Leontis, E. Westhof, Geometric nomenclature and classification of RNA base pairs. *RNA* **2001**, *7*, 499-512.
- (35) F. H. Crick, Codon--anticodon pairing: the wobble hypothesis. *J Mol Biol* **1966**, *19*, 548-555.
- (36) G. Varani, W. H. McClain, The G x U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems. *EMBO Rep* **2000**, *1*, 18-23.
- (37) B. J. Walberer, A. C. Cheng, A. D. Frankel, Structural diversity and isomorphism of hydrogen-bonded base interactions in nucleic acids. *J Mol Biol* **2003**, *327*, 767-780.
- (38) J. D. Puglisi, J. R. Wyatt, I. Tinoco, Jr., Solution conformation of an RNA hairpin loop. *Biochemistry* **1990**, *29*, 4215-4226.
- (39) W. Saenger *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.
- (40) N. H. Woo, B. A. Roe, A. Rich, Three-dimensional structure of Escherichia coli initiator tRNA^{fMet}. *Nature* **1980**, *286*, 346-351.
- (41) E. Schmitt, M. Panvert, S. Blanquet, Y. Mechulam, Crystal structure of methionyl-tRNA^{fMet} transformylase complexed with the initiator formyl-methionyl-tRNA^{fMet}. *EMBO J* **1998**, *17*, 6819-6826.
- (42) P. C. Durant, D. R. Davis, Stabilization of the anticodon stem-loop of tRNA^{Lys,3} by an A+-C base-pair and by pseudouridine. *J Mol Biol* **1999**, *285*, 115-131.
- (43) P. Dagenais, N. Girard, E. Bonneau, P. Legault, Insights into RNA structure and dynamics from recent NMR and X-ray studies of the Neurospora Varkud satellite ribozyme. *Wiley Interdiscip Rev RNA* **2017**, *8*, e1421.
- (44) S. K. Wong, S. Sato, D. W. Lazinski, Substrate recognition by ADAR1 and ADAR2. *RNA* **2001**, *7*, 846-858.
- (45) J. M. Thomas, P. A. Beal, How do ADARs bind RNA? New protein-RNA structures illuminate substrate recognition by the RNA editing ADARs. *Bioessays* **2017**, *39*, 1600187.
- (46) D. G. Sashital, G. Cornilescu, C. J. McManus, D. A. Brow, S. E. Butcher, U2-U6 RNA folding reveals a group II intron-like domain and a four-helix junction. *Nat Struct Mol Biol* **2004**, *11*, 1237-1242.
- (47) S. L. Yean, G. Wuenschell, J. Termini, R. J. Lin, Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature* **2000**, *408*, 881-884.
- (48) S. M. Fica, N. Tuttle, T. Novak, N. S. Li, J. Lu, P. Koodathingal, Q. Dai, J. P. Staley, J. A. Piccirilli, RNA catalyses nuclear pre-mRNA splicing. *Nature* **2013**, *503*, 229-234.
- (49) X. Zhang, C. Yan, J. Hang, L. I. Finci, J. Lei, Y. Shi, An Atomic Structure of the Human Spliceosome. *Cell* **2017**, *169*, 918-929 e914.
- (50) M. Hutton, C. L. Lendon, P. Rizzu, M. Baker, S. Froelich, H. Houlden, S. Pickering-Brown, S. Chakraverty, A. Isaacs, A. Grover, J. Hackett, J. Adamson, S. Lincoln, D. Dickson, P. Davies, R. C. Petersen, M. Stevens, E. de Graaff, E. Wauters, J. van Baren, M. Hillebrand, M. Joosse, J. M. Kwon, P. Nowotny, L. K. Che, J. Norton, J. C. Morris, L. A. Reed, J. Trojanowski, H. Basun, L. Lannfelt, M. Neystat, S. Fahn, F. Dark, T. Tannenberg, P. R. Dodd, N. Hayward, J. B. Kwok, P. R. Schofield, A. Andreadis, J. Snowden, D. Craufurd, D. Neary, F. Owen, B. A.

- Oostra, J. Hardy, A. Goate, J. van Swieten, D. Mann, T. Lynch, P. Heutink, Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature* **1998**, *393*, 702-705.
- (51) A. McCarthy, R. Loneragan, D. A. Olszewska, S. O'Dowd, G. Cummins, B. Magennis, E. M. Fallon, N. Pender, E. D. Huey, S. Cosentino, K. O'Rourke, B. D. Kelly, M. O'Connell, I. Delon, M. Farrell, M. G. Spillantini, L. P. Rowland, S. Fahn, P. Craig, M. Hutton, T. Lynch, Closing the tau loop: the missing tau mutation. *Brain* **2015**, *138*, 3100-3109.
 - (52) J. L. Wilcox, P. C. Bevilacqua, pKa shifting in double-stranded RNA is highly dependent upon nearest neighbors and bulge positioning. *Biochemistry* **2013**, *52*, 7470-7476.
 - (53) J. C. Lin, C. Hyeon, D. Thirumalai, RNA under tension: Folding Landscapes, Kinetic partitioning Mechanism, and Molecular Tensegrity. *J Phys Chem Lett* **2012**, *3*, 3616-3625.
 - (54) D. Chakraborty, R. Collepardo-Guevara, D. J. Wales, Energy landscapes, folding mechanisms, and kinetics of RNA tetraloop hairpins. *J Am Chem Soc* **2014**, *136*, 18052-18061.
 - (55) C. Hyeon, D. Thirumalai, Forced-unfolding and force-quench refolding of RNA hairpins. *Biophys J* **2006**, *90*, 3410-3427.
 - (56) K. Neupane, H. Yu, D. A. Foster, F. Wang, M. T. Woodside, Single-molecule force spectroscopy of the add adenine riboswitch relates folding to regulatory mechanism. *Nucleic Acids Res* **2011**, *39*, 7677-7687.
 - (57) P. T. Li, D. Collin, S. B. Smith, C. Bustamante, I. Tinoco, Jr., Probing the mechanical folding kinetics of TAR RNA by hopping, force-jump, and force-ramp methods. *Biophys J* **2006**, *90*, 250-260.
 - (58) W. Stephenson, S. Keller, R. Santiago, J. E. Albrecht, P. N. Asare-Okai, S. A. Tenenbaum, M. Zuker, P. T. Li, Combining temperature and force to study folding of an RNA hairpin. *Phys Chem Chem Phys* **2014**, *16*, 906-917.
 - (59) M. de Messieres, J. C. Chang, A. T. Belew, A. Meskauskas, J. D. Dinman, A. La Porta, Single-molecule measurements of the CCR5 mRNA unfolding pathways. *Biophys J* **2014**, *106*, 244-252.
 - (60) J. Liphardt, B. Onoa, S. B. Smith, I. Tinoco, Jr., C. Bustamante, Reversible unfolding of single RNA molecules by mechanical force. *Science* **2001**, *292*, 733-737.
 - (61) G. Chen, J. D. Wen, I. Tinoco, Jr., Single-molecule mechanical unfolding and folding of a pseudoknot in human telomerase RNA. *RNA* **2007**, *13*, 2175-2188.
 - (62) M. T. Woodside, C. Garcia-Garcia, S. M. Block, Folding and unfolding single RNA molecules under tension. *Curr Opin Chem Biol* **2008**, *12*, 640-646.
 - (63) I. Tinoco, Jr., Force as a useful variable in reactions: unfolding RNA. *Annu Rev Biophys Biomol Struct* **2004**, *33*, 363-385.
 - (64) J. D. Wen, M. Manosas, P. T. Li, S. B. Smith, C. Bustamante, F. Ritort, I. Tinoco, Jr., Force unfolding kinetics of RNA using optical tweezers. I. Effects of experimental variables on measured results. *Biophys J* **2007**, *92*, 2996-3009.
 - (65) P. T. Li, C. Bustamante, I. Tinoco, Jr., Real-time control of the energy landscape by force directs the folding of RNA molecules. *Proc Natl Acad Sci U S A* **2007**, *104*, 7039-7044.
 - (66) G. Chen, K. Y. Chang, M. Y. Chou, C. Bustamante, I. Tinoco, Triplex structures in an RNA pseudoknot enhance mechanical stability and increase efficiency of-1 ribosomal frameshifting. *Proc Natl Acad Sci U S A* **2009**, *106*, 12706-12711.

- (67) Z. Zhong, L. Yang, H. Zhang, J. Shi, J. J. Vandana, D. T. Lam, R. C. Olsthoorn, L. Lu, G. Chen, Mechanical unfolding kinetics of the SRV-1 gag-pro mRNA pseudoknot: possible implications for -1 ribosomal frameshifting stimulation. *Sci Rep* **2016**, *6*, 39549.
- (68) Y. J. Wu, C. H. Wu, A. Y. Yeh, J. D. Wen, Folding a stable RNA pseudoknot through rearrangement of two hairpin structures. *Nucleic Acids Res* **2014**, *42*, 4505-4515.
- (69) J. Viereg, W. Cheng, C. Bustamante, I. Tinoco, Jr., Measurement of the effect of monovalent cations on RNA hairpin stability. *J Am Chem Soc* **2007**, *129*, 14966-14973.
- (70) G. Z. Yusupova, M. M. Yusupov, J. H. Cate, H. F. Noller, The path of messenger RNA through the ribosome. *Cell* **2001**, *106*, 233-241.
- (71) S. Takyar, R. P. Hickerson, H. F. Noller, mRNA helicase activity of the ribosome. *Cell* **2005**, *120*, 49-58.
- (72) T. Liu, A. Kaplan, L. Alexander, S. Yan, J. D. Wen, L. Lancaster, C. E. Wickersham, K. Fredrick, H. Noller, I. Tinoco, C. J. Bustamante, Direct measurement of the mechanical work during translocation by the ribosome. *Elife* **2014**, *3*, e03406.
- (73) J. D. Wen, L. Lancaster, C. Hodges, A. C. Zeri, S. H. Yoshimura, H. F. Noller, C. Bustamante, I. Tinoco, Following translation by single ribosomes one codon at a time. *Nature* **2008**, *452*, 598-603.
- (74) P. J. Farabaugh, Programmed translational frameshifting. *Microbiol Rev* **1996**, *60*, 103-134.
- (75) T. Jacks, H. E. Varmus, Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. *Science* **1985**, *230*, 1237-1242.
- (76) Y. G. Kim, L. Su, S. Maas, A. O'Neill, A. Rich, Specific mutations in a viral RNA pseudoknot drastically change ribosomal frameshifting efficiency. *Proc Natl Acad Sci U S A* **1999**, *96*, 14234-14239.
- (77) I. Brierley, F. J. Dos Ramos, Programmed ribosomal frameshifting in HIV-1 and the SARS-CoV. *Virus Res* **2006**, *119*, 29-42.
- (78) J. D. Dinman, Mechanisms and implications of programmed translational frameshifting. *Wiley Interdiscip Rev RNA* **2012**, *3*, 661-673.
- (79) K. D. Mouzakis, A. L. Lang, K. A. Vander Meulen, P. D. Easterday, S. E. Butcher, HIV-1 frameshift efficiency is primarily determined by the stability of base pairs positioned at the mRNA entrance channel of the ribosome. *Nucleic Acids Res* **2013**, *41*, 1901-1913.
- (80) P. Garcia-Miranda, J. T. Becker, B. E. Benner, A. Blume, N. M. Sherer, S. E. Butcher, Stability of HIV Frameshift Site RNA Correlates with Frameshift Efficiency and Decreased Virus Infectivity. *J Virol* **2016**, *90*, 6906-6917.
- (81) O. Namy, S. J. Moran, D. I. Stuart, R. J. Gilbert, I. Brierley, A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting. *Nature* **2006**, *441*, 244-247.
- (82) T. M. Hansen, S. N. S. Reihani, L. B. Oddershede, M. A. Sorensen, Correlation between mechanical strength of messenger RNA pseudoknots and ribosomal frameshifting. *Proc Natl Acad Sci U S A* **2007**, *104*, 5830-5835.
- (83) J. D. Dinman, T. Icho, R. B. Wickner, A -1 ribosomal frameshift in a double-stranded RNA virus of yeast forms a gag-pol fusion protein. *Proc Natl Acad Sci U S A* **1991**, *88*, 174-178.
- (84) C. H. Yu, M. P. Teulade-Fichou, R. C. Olsthoorn, Stimulation of ribosomal frameshifting by RNA G-quadruplex structures. *Nucleic Acids Res* **2014**, *42*, 1887-1892.

- (85) C. H. Yu, M. H. Noteborn, C. W. A. Pleij, R. C. L. Olsthoorn, Stem-loop structures can effectively substitute for an RNA pseudoknot in -1 ribosomal frameshifting. *Nucleic Acids Res* **2011**, *39*, 8952-8959.
- (86) D. B. Ritchie, D. A. Foster, M. T. Woodside, Programmed -1 frameshifting efficiency correlates with RNA pseudoknot conformational plasticity, not resistance to mechanical unfolding. *Proc Natl Acad Sci U S A* **2012**, *109*, 16167-16172.
- (87) R. Y. Puah, H. Jia, M. Maraswami, D.-F. K. Toh, R. Ero, L. Yang, K. M. Patil, A. A. L. Ong, M. S. Krishna, R. Sun, C. Tong, M. Huang, X. Chen, T. P. Loh, Y. G. Gao, D. X. Liu, G. Chen, Selective Binding to mRNA Duplex Regions by Chemically Modified Peptide Nucleic Acids Stimulates Ribosomal Frameshifting. *Biochemistry* **2018**, *57*, 149-159.
- (88) P. J. Elms, J. D. Chodera, C. J. Bustamante, S. Marqusee, Limitations of constant-force-feedback experiments. *Biophys J* **2012**, *103*, 1490-1499.
- (89) S. B. Smith, Y. Cui, C. Bustamante, Optical-trap force transducer that operates by direct measurement of light momentum. *Methods Enzymol* **2003**, *361*, 134-162.
- (90) G. Grentzmann, J. A. Ingram, P. J. Kelly, R. F. Gesteland, J. F. Atkins, A dual-luciferase reporter system for studying recoding signals. *RNA* **1998**, *4*, 479-486.
- (91) R. J. Marcheschi, K. D. Mouzakis, S. E. Butcher, Selection and Characterization of Small Molecules That Bind the HIV-1 Frameshift Site RNA. *ACS Chem Biol* **2009**, *4*, 844-854.
- (92) H. Wang, M. P. Matise, In vivo dual luciferase reporter assay with chick neural tube in ovo electroporation system. *Methods Mol Biol* **2013**, *1018*, 211-217.
- (93) J. W. Hager, J. D. Dinman, An in vivo dual-luciferase assay system for studying translational recoding in the yeast *Saccharomyces cerevisiae*. *RNA* **2003**, *9*, 1019-1024.
- (94) D. H. Mathews, Using an RNA secondary structure partition function to determine confidence in base pairs predicted by free energy minimization. *RNA* **2004**, *10*, 1178-1190.
- (95) Z. J. Lu, J. W. Gloor, D. H. Mathews, Improved RNA secondary structure prediction by maximizing expected pair accuracy. *RNA* **2009**, *15*, 1805-1813.
- (96) S. Bellaousov, D. H. Mathews, ProbKnot: fast prediction of RNA secondary structure including pseudoknots. *RNA* **2010**, *16*, 1870-1880.
- (97) N. Zhang, Y. Wang, L. An, X. Song, Q. Huang, Z. Liu, L. Yao, Entropy Drives the Formation of Salt Bridges in the Protein GB3. *Angew Chem Int Ed Engl* **2017**, *56*, 7601-7604.
- (98) J. Ubbink, T. Odijk, Polymer- and salt-induced toroids of hexagonal DNA. *Biophys J* **1995**, *68*, 54-61.
- (99) D. Bensimon, A. J. Simon, V. V. Croquette, A. Bensimon, Stretching DNA with a receding meniscus: Experiments and models. *Phys Rev Lett* **1995**, *74*, 4754-4757.
- (100) F. Vanzi, Y. Takagi, H. Shuman, B. S. Cooperman, Y. E. Goldman, Mechanical studies of single ribosome/mRNA complexes. *Biophys J* **2005**, *89*, 1909-1919.
- (101) C. Bustamante, J. F. Marko, E. D. Siggia, S. Smith, Entropic elasticity of lambda-phage DNA. *Science* **1994**, *265*, 1599-1600.
- (102) M. D. Wang, H. Yin, R. Landick, J. Gelles, S. M. Block, Stretching DNA with optical tweezers. *Biophys J* **1997**, *72*, 1335-1346.

- (103) M. T. Woodside, P. C. Anthony, W. M. Behnke-Parks, K. Larizadeh, D. Herschlag, S. M. Block, Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. *Science* **2006**, *314*, 1001-1004.
- (104) M. T. Woodside, W. M. Behnke-Parks, K. Larizadeh, K. Travers, D. Herschlag, S. M. Block, Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins. *Proc Natl Acad Sci U S A* **2006**, *103*, 6190-6195.
- (105) P. C. Bevilacqua, J. M. Bloise, Structures, kinetics, thermodynamics, and biological functions of RNA hairpins. *Annu Rev Phys Chem* **2008**, *59*, 79-103.
- (106) W. Zhang, S. J. Chen, RNA hairpin-folding kinetics. *Proc Natl Acad Sci U S A* **2002**, *99*, 1931-1936.
- (107) S. A. McKinney, C. Joo, T. Ha, Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophys J* **2006**, *91*, 1941-1951.
- (108) I. Tinoco, C. Bustamante, The effect of force on thermodynamics and kinetics of single molecule reactions. *Biophys Chem* **2002**, *101*, 513-533.
- (109) S. Ravindranathan, S. E. Butcher, J. Feigon, Adenine protonation in domain B of the hairpin ribozyme. *Biochemistry* **2000**, *39*, 16026-16032.
- (110) N. J. Reiter, H. Blad, F. Abildgaard, S. E. Butcher, Dynamics in the U6 RNA intramolecular stem-loop: a base flipping conformational change. *Biochemistry* **2004**, *43*, 13739-13747.
- (111) X. Qu, J. D. Wen, L. Lancaster, H. F. Noller, C. Bustamante, I. Tinoco, Jr., The ribosome uses two active mechanisms to unwind messenger RNA during translation. *Nature* **2011**, *475*, 118-121.
- (112) A. M. Mustoe, S. Busan, G. M. Rice, C. E. Hajdin, B. K. Peterson, V. M. Ruda, N. Kubica, R. Nutiu, J. L. Baryza, K. M. Weeks, Pervasive Regulatory Functions of mRNA Structure Revealed by High-Resolution SHAPE Probing. *Cell* **2018**, *173*, 181-195 e118.
- (113) J. M. Huguet, C. V. Bizarro, N. Forns, S. B. Smith, C. Bustamante, F. Ritort, Single-molecule derivation of salt dependent base-pair free energies in DNA. *Proc Natl Acad Sci U S A* **2010**, *107*, 15431-15436.
- (114) C. V. Bizarro, A. Alemany, F. Ritort, Non-specific binding of Na⁺ and Mg²⁺ to RNA determined by force spectroscopy methods. *Nucleic Acids Res* **2012**, *40*, 6922-6935.
- (115) J. Camunas-Soler, A. Alemany, F. Ritort, Experimental measurement of binding energy, selectivity, and allostery using fluctuation theorems. *Science* **2017**, *355*, 412-415.
- (116) M. Manosas, J. Camunas-Soler, V. Croquette, F. Ritort, Single molecule high-throughput footprinting of small and large DNA ligands. *Nat Commun* **2017**, *8*, 304.
- (117) W. J. Greenleaf, K. L. Frieda, D. A. Foster, M. T. Woodside, S. M. Block, Direct observation of hierarchical folding in single riboswitch aptamers. *Science* **2008**, *319*, 630-633.
- (118) V. K. Duesterberg, I. T. Fischer-Hwang, C. F. Perez, D. W. Hogan, S. M. Block, Observation of long-range tertiary interactions during ligand binding by the TPP riboswitch aptamer. *Elife* **2015**, *4*.
- (119) V. Chandra, Z. Hannan, H. Xu, M. Mandal, Single-molecule analysis reveals multi-state folding of a guanine riboswitch. *Nat Chem Biol* **2017**, *13*, 194-201.
- (120) M. J. McCauley, I. Rouzina, K. A. Manthei, R. J. Gorelick, K. Musier-Forsyth, M. C. Williams, Targeted binding of nucleocapsid protein transforms the folding landscape of HIV-1 TAR RNA. *Proc Natl Acad Sci U S A* **2015**, *112*, 13555-13560.

- (121) P. Shrestha, Y. Cui, J. Wei, S. Jonchhe, H. Mao, Single-Molecule Mechanochemical pH Sensing Revealing the Proximity Effect of Hydroniums Generated by an Alkaline Phosphatase. *Anal Chem* **2018**, *90*, 1718-1724.

Chapter 3

Base Triple Formation between Loop 2 and Stem 1 in SRV-1 Pseudoknot Inhibits Antisense PNA Invasion of Stem 1 and Stem 2

3.1 Introduction

RNA pseudoknot structures are important tertiary structures essential for many catalytic and regulatory functions.^{1,2} In a hairpin (H)-type RNA pseudoknot, sequences in the hairpin loop and those outside of the hairpin loop form intramolecular base pairs, resulting in the formation of two helical stems and two loops (**Figure 3.1A**). The loops in a pseudoknot could interact with adjacent stems forming tertiary structures such as major-groove and minor-groove base triples, which are often critical for various functions.³

Adenosine and other nucleotides in loop 2 interact with canonical base pairs in stem 1 to form minor-groove interactions such as the A-minor motif triples involving 2'-OH of the ribose (**Figure 3.1B**). RNA base triples are essential structural components of Group I and Group II introns,^{4,5} spliceosome,^{6,7} SAM-II riboswitch,⁸ and telomerase RNA pseudoknots.^{9,10} Moreover, RNA base triples are extensively found in ribosomal frameshift-stimulating pseudoknots in mRNAs.¹¹⁻¹⁷ For example, minor-groove base triples formed between loop 2 and stem 1 are found in various -1 ribosomal frameshifting (-1 FS) stimulatory

pseudoknot structures including beet western yellow virus (BWYV)¹³, pea enation virus type-1 (PEMV-1)¹⁴, potato leaf roll virus (PLRV)^{16,17}, and simian retrovirus type-1 (SRV-1)^{11,12}. Studies have shown that the formation of base triples could affect the thermodynamics and mechanical properties of pseudoknots and thus stimulate -1 FS.^{16,18-23} -1 FS exists in many viruses and cells for generating different ratios of functional proteins. During -1 FS, a ribosome translating a mRNA would be temporally stalled and paused over a slippery sequence (e.g., G GGA AAC) in a metastable conformational state when encountered by a downstream roadblock (located 2–8 nucleotides downstream), such as an RNA hairpin²⁴⁻²⁶ or an RNA pseudoknot.^{11,27-29} A backward slippage of the ribosome on the slippery sequence by one nucleotide results in -1 FS, after which the ribosome continues translation at the -1 frame.^{30,31}

Mutagenesis studies have shown that the disruption of the base triples in pseudoknots would drastically lower the ribosomal frameshifting efficiency.^{11,22,32} Single-molecule mechanical pulling studies showed that the disruption of the base triples in pseudoknots causes decreased mechanical stability.^{20,21,33} Correlations between the mechanical properties and -1 FS efficiency have been reported for both pseudoknots and hairpins.^{18-21,25,33} Enhanced mechanical stability of pseudoknots may provide increased resistance against a ribosome at the slippery site, resulting in the inhibition of the mRNA unwinding helicase activity of ribosome, which in turn stimulates ribosomal frameshifting.

It is crucial in understanding how base triples formed in pseudoknots would affect the structure and stability of stem 1, which is in direct contact with ribosome at the slippery site. In addition, stem 2 stability may also be critical in maintaining a pseudoknot structure for the stimulation of -1 FS. Single-molecule pulling experiments with force applied on the terminal ends of a pseudoknot are useful in revealing how base triple formation affects mechanical stability of a pseudoknot. However, the effect of base triple formation on local stability of stem 1 and stem 2 is not directly revealed by pulling experiments, as the mechanical pulling force is applied on both stem 1 and stem 2, which often results in cooperative unfolding of the pseudoknots.^{21,34,35}

Peptide nucleic acid (PNA) has enhanced binding affinity toward complementary RNAs or DNAs due to reduce charge-charge repulsion with a neutral peptide-like backbone (**Figure 3.2A–D**).³⁶ An antisense PNA can invade RNA and DNA stems by hybridizing to the complementary strand following Watson–Crick base pairing rules (**Figure 3.2A–D**), with the other strand displaced.^{37–41}

Figure 3.1. SRV-1 frameshifting RNA pseudoknots and control hairpins used in this study and PNA binding schemes. (A) Pseudoknot SF206 and its mutants SF217, SF220, and SF348 (with residues from 1 to 34), and control hairpins (with residues from 1 to 19 for H1, from 1 to 22 for H1E, and from 8 to 34 for H2 and H2S). The tertiary interactions between loop 2 and stem 1 are indicated using black dots. (B) Structures of the minor-groove base triples A21·C2·G19, A26·G4·C16, U27·C5·G15, and A28·C6·G14 formed in SF206.¹² (C) Scheme of the PNA P1 invasion through anti-parallel PNA-RNA duplex formation with the 5'-side of stem 1 in RNA pseudoknot, resulting in the disruption of stem 1 and stem-1-loop-2 tertiary interactions. (D) Structure of the complex formed between antisense PNA P1 and SF348. The self-complementary 5'-GGCC-3' segment

within the loop of H2S may form a mini duplex, resulting in a dimer structure formation. (E) Scheme of the PNA P2 invasion through anti-parallel PNA-RNA duplex formation with the 3'-side of stem 2 in RNA pseudoknot, resulting in the disruption of stem 2 and stem 1-loop 2 tertiary interactions. (F) Structure of the complex formed between antisense PNA P2 and SF348.

Studies have shown that antisense PNAs can be used for recognizing RNA component of human telomerase and inhibiting telomerase activity *in vitro*⁴² and in cell⁴³, binding to pre-mRNAs to alter splicing patterns *in vitro*,⁴⁴ inhibiting translation *in vitro* and in cell through binding to the terminus of the 5'-untranslated region of mRNA and blocking the binding of the translation machinery,⁴⁵ targeting the coding region of mRNA to arrest translation elongation,^{46,47} and invading ribosomal RNA helix structures for inhibiting ribosome activity.⁴⁸

In this study, we used an antisense PNA targeting the 5' arm of stem 1 and another antisense PNA targeting the 3' arm of stem 2 of SRV-1 frameshifting pseudoknot to investigate how minor-groove interactions affect stem 1 and stem 2 stabilities (**Figure 3.1**) by non-denaturing polyacrylamide gel electrophoresis (PAGE). We also examined the possible implications of PNA invasions for ribosomal frameshifting.

Figure 3.2. Chemical structures of PNA-RNA base pairs and PNAs P1 and P2. The letter R represents the sugar-phosphate backbone of RNA. Hydrogen bonding interactions are indicated by black dashed lines. (A–D) Watson–Crick PNA-RNA base pairs of T-A, G-C, C-G, and A-U. Chemical structures of (E) PNA P1 (NH₂-Lys-TGGCCGC-CONH₂) and (F) PNA P2 (NH₂-Lys-GCTCCA-CONH₂). The N-terminal lysine residue has an L configuration.

3.2 Materials and Methods

3.2.1 PNA Oligomer Synthesis

The PNA monomers of thymine (T), guanine (G), cytosine (C), and adenine (A) were purchased from ASM Research Chemicals. PNA oligomers P1 (NH₂-Lys-TGGCCGC-CONH₂) and P2 (NH₂-Lys-GCTCCA-CONH₂) were synthesized

manually using Boc chemistry via a Solid-Phase Peptide Synthesis protocol.⁴⁹ The PNA oligomer was purified by reverse phase high-performance liquid chromatography (RP-HPLC) and then verified through molecular weight characterization using matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF).

3.2.2 Non-denaturing Polyacrylamide Gel Electrophoresis

The RNA oligonucleotides for the SRV-1 pseudoknots (SF206: 5'-GCGGCCAGCUCCAGGCCGCCAAACAAUAUGGAGC-3'; SF217: 5'-GCGGCCAGCUCCAGGCCGCCAAACACUAUGGAGC-3'; SF220: 5'-GCGGCCAGCUCCAGGCCGCCAAACAACAUGGAGC-3'; and SF348: 5'-GCGGCCAGCUCCAGGCCGCCAAAUGGAGC-3'), control hairpins (H1: 5'-GCGGCCAGCUCCAGGCCGC-3'; H1E: 5'-GCGGCCAGCUCCAGGCCGCCAA-3'; H2: 5'-GCUCCAGGCCGCCAAACAAUAUGGAGC-3'; and H2S: 5'-GCUCCAGGCCGCCAAAUGGAGC-3'), and short RNA oligonucleotides for invading stem 1 (R1: 5'-UGGCCGC-3') were purchased from Sigma. The pseudoknots were annealed by slow cooling (heated at 95 °C for 5 min, cooled down to room temperature over 2 hours and incubated at 4 °C overnight). The control hairpins were annealed by snap cooling (heated at 95 °C for 5 min, followed by incubation on ice for 10 min). The samples were then annealed with P1/P2/R1 (slow cooling from 95 °C to room temperature and incubation at 4 °C overnight). The buffer contains 200 mM NaCl, 20 mM HEPES, 0.1 mM EDTA, pH 7.3. Prior to sample loading, 4 µL 35% glycerol was added to the sample. 20 µL of the mixture (1 µM RNA with varied P1/P2/R1 concentrations) was loaded

onto a 12% PAGE gel (with a 19:1 acrylamide to bisacrylamide ratio). The gels were run at 250 V for 6 h at 4 °C in 1× TBE (89 mM Tris, 89 mM Borate, 2 mM EDTA, pH 8.3), post-stained with ethidium bromide and imaged by a Typhoon Trio Variable Mode Imager (GE Healthcare Life Sciences, USA).

3.2.3 Data Analysis

The dissociation of the RNA-PNA bound complex is treated as a simple first-order reaction. The equilibrium binding constants can be obtained from the concentration of bound and free RNA after equilibrium has been reached. The data were fit to a quadratic equation (equation 3.2) to determine K_d from the experimental data.⁵⁰

The fraction of bound RNA can be obtained from the band intensity of the bound complex divided by the sum of the band intensities in bound and free RNA. The fraction of bound RNA (Y_{bound}) was calculated through according to

$$Y_{bound} = I_{bound}/(I_{bound}+I_{free}) \quad (3.1)$$

where I_{free} is the band intensity of free RNA with a relatively fast migration rate, I_{bound} is the band intensity of the bound RNA with a relatively slow migration rate.

The fraction of bound RNA (Y_{bound}) was fit to equation 3.2 using non-linear least squares methods.⁵¹

$$Y_{bound} = A + (B/2/R_0)(R_0 + X + K_d - [(R_0 + X + K_d)^2 - 4R_0X]^{1/2}) \quad (3.2)$$

where R_0 is the RNA pseudoknot concentration (1 μM) loaded in the gel, X is the total ligand concentration added to each RNA sample, K_d is the apparent dissociation constant, A and B are the minimum and maximum fractions of bound RNA, respectively.

If the bound RNA is not directly imaged, the fraction of free RNA was determined by the band intensity of the free RNA divided by the band intensity of free RNA with no ligand added. The fraction of free RNA was calculated from

$$Y_{free} = I_{free}/I_{free,max} \quad (3.3)$$

where I_{free} is the band intensities in the presence of ligand, $I_{free,max}$ is band intensity of free RNA in the absence of ligand.

The fraction of free RNA (Y_{free}) can be extracted from

$$Y_{free} = 1 - Y_{bound}. \quad (3.4)$$

The fraction of free RNA (Y_{free}) was then fitted to

$$Y_{free} = 1 - [A + (B/2/R_0)(R_0 + X + K_d - ((R_0 + X + K_d)^2 - 4R_0X)^{1/2})] \quad (3.5)$$

where R_0 is the total RNA pseudoknot concentration (1 μM) loaded in the gel, X is the total ligand concentration added to each RNA sample, K_d is the apparent dissociation constant, A and B are the minimum and maximum fractions of bound RNA, respectively.

3.3 Results and Discussion

3.3.1 Invasion of Stem 1 and Stem 2 in SRV-1 Pseudoknots

In SRV-1 pseudoknot SF206, there are 3 consecutive base triples A26·G4-C16, U27·C5-G15, and A28·C6-G14 together with an additionally base triple A21·C2-G18 formed between loop 2 and stem 1(**Figure 3.1**).¹² In SF217, an A26C mutation causes the disruption of the A26·G4-C16 base triple and the base stacking with adjacent bases in loop 2. In SF220, a U27C mutation may not affect the single hydrogen bond formed with C5-G15 pair. However, the U27C mutation may change the stacking interactions with A26. SF348 has a mutated loop 2 with three adenines, which may result in the disruption of all the minor-groove base triples.

We designed a short RNA oligonucleotide R1 (5'-UGGCCGC-3') with the sequence complementary to the 5' side of stem 1 for the invasion of stem 1. The PAGE data reveal that R1 indeed can bind to the pseudoknots as indicated by the appearance of a slow migrating band (**Figure 3.3**). Two slow migrating complex bands were observed for SF348 upon the addition of R1, indicating the formation of two structures between R1 and SF348. The band with a small shift may correspond to the invasion complex as expected (**Figure 3.1D**). The band with a big shift suggests the formation of a higher order complex structure. For example, invasion of stem 1 by R1 exposes the loop of stem 2 hairpin (H2S) for forming loop-loop kissing interactions, with the self-complementary 5'-GGCC-3' segment within the loop forming a mini duplex (**Figure 3.1D**). Such a higher

order structure is consistent with the relatively big shift of the second slow-moving band.

A fast migrating band was observed for all pseudoknots with increasing concentration of R1, which we assign as the self-complementary duplex formed by R1 containing four G-C pairs and two G-U pairs (5'-UGGCCGC-3'/3'-CGCCGGU-5', see **Figure 3.3F**). By fitting the PAGE data to equation 3.2, we obtained the apparent binding affinities for oligonucleotide R1 binding to the RNA pseudoknots ($K_d > 50 \mu\text{M}$ for SF206, $K_d = 37.6 \pm 7.8 \mu\text{M}$ for SF220, $K_d = 4.7 \pm 0.9 \mu\text{M}$ for SF217, and $K_d = 1.2 \pm 0.3 \mu\text{M}$ for SF348, **Figure 3.3, Table 3.1**).

Figure 3.3. Non-denaturing PAGE study of short RNA oligonucleotides R1 binding to SRV-1 RNA pseudoknots SF206, SF220, SF217, and SF348. The gels were run for 7 h at 250 V in 1× TBE. (A) Non-denaturing PAGE results of RNA pseudoknots and hairpins. The pseudoknots and hairpins were annealed by slow cooling and snap cooling, respectively. Except for hairpin H2, which shows a relatively faint band, probably due the formation of a relatively unstable structure with a relatively large loop. (B–E) Non-denaturing PAGE results of R1 binding

to RNA pseudoknots. The concentration of the RNA pseudoknots loaded was 1 μM in 20 μl . The R1 concentrations in lanes from left to right are 0, 0.01, 0.05, 0.1, 0.5, 1, 2, 5, 8, 10, 15, 20, 30, and 50 μM , respectively. (F) Non-denaturing PAGE results of R1 alone. SF206 with 0 μM R1 and 50 μM R1 were loaded at the first lane and second lane, respectively, serving as controls. The other lanes starting from the third to the very right-hand end contained only R1 with varied R1 concentrations from left to right of 0, 0.01, 0.05, 0.1, 0.5, 1, 2, 5, 8, 10, 15, 20, 30, and 50 μM , respectively. R1 alone forms dimer with increasing concentration. (G) Apparent dissociation constant (K_d) determination by non-denaturing PAGE for short RNA oligonucleotide R1 invading stem 1 of SRV-1 RNA pseudoknots. The fractions of bound RNA (both 1:1 complex and 2:2 complex) for SF206, SF220, SF217, and SF348 at varied R1 concentrations were fit to equation 3.2.

To facilitate the invasion of the pseudoknot structures, we made a PNA P1 ($\text{NH}_2\text{-Lys-TGGCCC-CONH}_2$), which has the same base sequence of RNA R1, except for replacing the U with T. The non-denaturing PAGE indeed shows that PNA P1 can in general invade SF206, SF220, SF217 and SF348 more efficiently than RNA R1, with apparent K_d values of $12.0 \pm 4.3 \mu\text{M}$, $1.6 \pm 0.4 \mu\text{M}$, $2.6 \pm 0.6 \mu\text{M}$, and $1.2 \pm 0.3 \mu\text{M}$, respectively (**Figure 3.4, Table 3.1**). The PNA-RNA complex bands are not observed (**Figure 3.4**), probably because the resulting stem-loop structure containing stem 2 and a 9-nucleotide loop is not stably formed (see **Figure 3.1C and Figure 3.3A**). Interestingly, a P1-SF348 complex band was observed (**Figure 3.4D**), which is consistent with the fact that H2S (derived from SF348, **Figure 3.1A**) is relatively more stable with a shortened loop 2 (see

Figure 3.3A). P1 invading SF206 ($K_d = 12.0 \pm 4.3 \mu\text{M}$) shows the weakest binding compared to SF217 ($K_d = 2.6 \pm 0.6 \mu\text{M}$), SF220 ($K_d = 1.6 \pm 0.4$) and SF348 ($K_d = 1.2 \pm 0.3$) (**Figure 3.4, Table 3.1**), suggesting that stem 1-loop 2 base triples can enhance stem 1 stability resulting in the resistance against P1 invasion.

Figure 3.4. Non-denaturing PAGE study of PNA P1 invading SRV-1 RNA pseudoknots of SF206, SF220, SF217, and SF348, and control hairpins. The concentration of the RNA pseudoknots and hairpins loaded was $1 \mu\text{M}$ in $20 \mu\text{l}$.

The PNA concentrations in lanes from left to right are 0, 0.01, 0.05, 0.1, 0.5, 1, 2, 5, 8, 10, 15, 20, 30, and 50 μ M, respectively. The gels were run for 7 h at 250 V in 1 $\times$ TBE. (A–D) P1 shows tighter binding to mutants SF220, SF217 and SF348 than to SF206. (E–F) P1 shows equal binding affinity to H1 and H1E. Dissociation constant (K_d) determination by non-denaturing PAGE for PNA P1 invading (G) SRV-1 RNA pseudoknots of SF206, SF220, SF217, and SF348, and (H) control hairpins. The fractions of free RNA at varied P1 concentrations were fit to equation 3.5.

Interestingly, PNA P1 invades control hairpins H1 and H1E with comparable binding affinities (K_d values of 10.3 ± 3.1 and 10.0 ± 2.3 , respectively, **Figure 3.4E,F,H, Table 3.1**) compared to the invasion of SF206, which may indicate that stem 1 within the hairpin loop of hairpins H1 and H1E is stabilized by the 7-nucleotide loop. It is likely that a potential G-C pair within the hairpin loop of H1 and H1E may form (**Figure 3.1A**). However, hairpins are known to have relatively lower efficiency in stimulating frameshifting.^{52,53} Our data suggest that a pseudoknot without tertiary stem 1-loop 2 base triple forming is less stable than a hairpin. Thus, the destabilization effect of pseudoknot formation may be compensated by stem 1-loop 2 base triple formation. Clearly, stem 1 invasion by an RNA or a PNA can reveal the energetic coupling of tertiary and secondary structures. During –FS, the 5' end of stem 1 is in direct contact with the ribosome at the slippery site, with the stability of the first few base pairs of stem 1 at the mRNA entrance tunnel critical in stimulating FS. Thus, our stem 1 invasion data indicate that tertiary interactions of loop 2 and stem 1 may help enhance the

resistance of stem 1 against the mRNA helicase activity of a translating ribosome, leading to improved efficiency in ribosomal backward slippage by one nucleotide.

Figure 3.5. Non-denaturing PAGE study of PNA P2 invading SRV-1 RNA pseudoknots of SF206, SF220, SF217, and SF348, and control hairpins. The concentration of the RNA pseudoknots and hairpins loaded was 1 μM in 20 μl . The PNA concentrations in lanes from left to right are 0, 0.01, 0.05, 0.1, 0.5, 1, 2, 5, 8, 10, 15, 20, 30, and 50 μM , respectively. The gels were run for 7 h at 250 V in 1 $\times$ TBE. (A–D) P2 shows tighter binding to mutants SF220, SF217 and SF348 than to SF206. (E–F) P2 shows no binding to H1 and H1E. Dissociation constant (K_d) determination by non-denaturing PAGE for PNA P2 invading (G) SRV-1 RNA pseudoknots of SF206, SF220, SF217, and SF348, and (H) control

hairpins. The fractions of free RNA at varied P2 concentrations were fit to equation 3.5.

Next, we designed PNA P2 (NH₂-Lys-GCTCCA-CONH₂), which is complementary to the 3' side of stem 2, to invade stem 2. Surprisingly, we observed that P2 invades SF206, SF220, SF217 and SF348 with varied K_d values of 4.0 ± 0.5 μ M, 1.1 ± 0.1 μ M, 0.5 ± 0.1 μ M, and 0.4 ± 0.1 μ M, respectively (**Figure 3.5, Table 3.1**). Stem 2 has no direct interactions with loop 2, suggesting that stem 1-loop 2 interactions can enhance stem 2 through an allosteric effect, probably through the coaxial stacking between stem 1 and stem 2 within the pseudoknot. Our non-denaturing PAGE results are consistent with our previously reported mechanical unfolding data, which also reveal energetic coupling between component stems and base triple interactions, resulting in cooperative mechanical unfolding of pseudoknots.²¹ Taken together, our data clearly suggest that the tertiary interactions formed between stem 1 and loop 2 are coupled in determining the accessibility of both stem 1 and stem 2 for PNA invasion.

Table 3.1. K_d values (in μ M) of short RNA oligonucleotides R1, PNA P1 and PNA P2 invasion obtained by non-denaturing PAGE.

RNA	K_d (μ M)		
	R1	P1	P2
SF206	>50	12.0 ± 4.3	4.0 ± 0.5
SF220	37.6 ± 7.8	1.6 ± 0.4	1.1 ± 0.1
SF217	4.7 ± 0.9	2.6 ± 0.6	0.5 ± 0.1
SF348	1.2 ± 0.3	1.2 ± 0.3	0.4 ± 0.1
H1	-	10.3 ± 3.1	>50
H1E	-	10.0 ± 2.3	>50
'-' indicates the data were not measured in this study.			

3.3.2 Correlations among Stem Invasion Efficiency, Mechanical Unfolding Force, Mechanical Unfolding Kinetics, and Frameshifting Efficiency

Our previous single-molecule mechanical unfolding studies have revealed a correlation between mechanical stability of the pseudoknots and frameshifting efficiency, with an emerging hypothesis that the mechanical resistance of both component stems 1 and 2 are stabilized by stem 1-loop 2 base triple formation.²¹ Thus, a correlation is expected between invasion efficiency of stems 1 and 2 and mechanical stability (including mechanical unfolding force and mechanical unfolding kinetics). Indeed, we observe a positive correlation between ligand invading efficiency and mechanical stabilities previously reported²¹ (**Figure 3.6A–C**). An inverse correlation is also observed between ligand invasion efficiency and mechanical unfolding rate at 0 pN of the pseudoknots (**Figure 3.6D–F**).²¹ The inverse correlation between ligand binding and mechanical unfolding rate at 25 pN is also observed. Interestingly, compared to P1 invasion of stem 1, P2 invasion efficiency of stem 2 shows a relatively better correlation with mechanical properties, which may indicate that in mechanical unfolding process, stem 2 is relatively more easily mechanically melted by stretching force applied on the terminal ends of the pseudoknots. Clearly, our stem invasion studies can reveal the component stem stabilities, which is evidenced by a positive correlation between ligand invasion and frameshifting efficiency (**Figure 3.6G–I**).¹¹ Thus, P1 invading the 5' side of stem 1 of the pseudoknot can well mimic the ribosome helicase activity of unwinding the encountered structure. Notably, stem 2 invasion efficiency by P2 also shows a correlation with frameshifting efficiency, suggesting that a pseudoknot structure is stabilized

highly by coupled secondary and tertiary interactions, which may account for the high frameshifting efficiencies generally observed for pseudoknots.

Figure 3.6. Correlation between PNA binding affinity and minus-one ribosomal frameshifting, mechanical unfolding force and unfolding rate at 0 pN. (A–C) The K_d values of R1, P1, and P2 invasion show a positive correlation with mechanical unfolding forces reported.²¹ (D–F) The K_d values of R1, P1 and P2 invasion show an inverse correlation with mechanical unfolding rates lnk at 0 pN. (G–I) The K_d values of R1, P1 and P2 invasion show a positive correlation with frameshifting efficiencies previously reported.¹¹

3.4 Conclusions

In summary, we have developed a convenient PNA invasion method for dissecting energetic coupling between tertiary and secondary interactions. The PNA invasion results show that base triples in pseudoknots are critical in determining the accessibility of both stems 1 and 2 by PNA invasion. It is remarkable to reveal the allosteric effect of stem 2 stabilization through stem 1-loop 2 interaction within a pseudoknot. The correlation between PNA invasion and frameshifting is consistent with the fact that PNA invading the 5' side of stem 1 mimics ribosomal unwinding of encountered mRNA structure.⁵⁴⁻⁵⁶ Compared to mechanical unfolding by applying a stretching force on the terminal ends, PNA invasion may be advantageous for directly characterizing the stability of a defined structural component within a pseudoknot. The PNA invasion method can be applied to many other complex RNA structures^{57,58} to reveal the correlations among structure, stability, and function. A detailed insight into the molecular mechanisms in energetic coupling and functional regulation would help the rational design of ligands for precisely targeting functional RNA structures.

References

- (1) I. Brierley, S. Pennell, R. J. Gilbert, Viral RNA pseudoknots: versatile motifs in gene expression and replication. *Nat Rev Microbiol* **2007**, *5*, 598-610.
- (2) D. P. Giedroc, P. V. Cornish, Frameshifting RNA pseudoknots: structure and mechanism. *Virus Res* **2009**, *139*, 193-208.
- (3) G. Devi, Y. Zhou, Z. S. Zhong, D. F. K. Toh, G. Chen, RNA triplexes: from structural principles to biological and biotech applications. *Wiley Interdiscip Rev RNA* **2015**, *6*, 111-128.
- (4) A. A. Szewczak, L. Ortoleva-Donnelly, S. P. Ryder, E. Moncoeur, S. A. Strobel, A minor groove RNA triple helix within the catalytic core of a group I intron. *Nat Struct Biol* **1998**, *5*, 1037-1042.
- (5) N. Toor, K. S. Keating, S. D. Taylor, A. M. Pyle, Crystal structure of a self-spliced group II intron. *Science* **2008**, *320*, 77-82.
- (6) T. A. Steitz, J. A. Steitz, A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci U S A* **1993**, *90*, 6498-6502.
- (7) S. M. Fica, M. A. Mefford, J. A. Piccirilli, J. P. Staley, Evidence for a group II intron-like catalytic triplex in the spliceosome. *Nat Struct Mol Biol* **2014**, *21*, 464-471.
- (8) S. D. Gilbert, R. P. Rambo, D. Van Tyne, R. T. Batey, Structure of the SAM-II riboswitch bound to S-adenosylmethionine. *Nat Struct Mol Biol* **2008**, *15*, 177-182.
- (9) N. B. Ulyanov, K. Shefer, T. L. James, Y. Tzfati, Pseudoknot structures with conserved base triples in telomerase RNAs of ciliates. *Nucleic Acids Res* **2007**, *35*, 6150-6160.
- (10) C. A. Theimer, C. A. Blois, J. Feigon, Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Mol Cell* **2005**, *17*, 671-682.
- (11) R. C. Olsthoorn, R. Reumerman, C. W. Hilbers, C. W. Pleij, H. A. Heus, Functional analysis of the SRV-1 RNA frameshifting pseudoknot. *Nucleic Acids Res* **2010**, *38*, 7665-7672.
- (12) P. J. Michiels, A. A. Versleijen, P. W. Verlaan, C. W. Pleij, C. W. Hilbers, H. A. Heus, Solution structure of the pseudoknot of SRV-1 RNA, involved in ribosomal frameshifting. *J Mol Biol* **2001**, *310*, 1109-1123.
- (13) L. Su, L. Q. Chen, M. Egli, J. M. Berger, A. Rich, Minor groove RNA triplex in the crystal structure of a ribosomal frameshifting viral pseudoknot. *Nat Struct Biol* **1999**, *6*, 285-292.
- (14) P. L. Nixon, A. Rangan, Y. G. Kim, A. Rich, D. W. Hoffman, M. Hennig, D. P. Giedroc, Solution structure of a luteoviral P1-P2 frameshifting mRNA pseudoknot. *J Mol Biol* **2002**, *322*, 621-633.
- (15) P. V. Cornish, M. Hennig, D. P. Giedroc, A loop 2 cytidine-stem 1 minor groove interaction as a positive determinant for pseudoknot-stimulated -1 ribosomal frameshifting. *Proc Natl Acad Sci U S A* **2005**, *102*, 12694-12699.
- (16) Y. G. Kim, S. Maas, S. C. Wang, A. Rich, Mutational study reveals that tertiary interactions are conserved in ribosomal frameshifting pseudoknots of two luteoviruses. *Rna-a Publication of the Rna Society* **2000**, *6*, 1157-1165.
- (17) P. S. Pallan, W. S. Marshall, J. Harp, F. C. Jewett, 3rd, Z. Wawrzak, B. A. Brown, 2nd, A. Rich, M. Egli, Crystal structure of a luteoviral RNA pseudoknot and

- model for a minimal ribosomal frameshifting motif. *Biochemistry* **2005**, *44*, 11315-11322.
- (18) T. M. Hansen, S. N. S. Reihani, L. B. Oddershede, M. A. Sorensen, Correlation between mechanical strength of messenger RNA pseudoknots and ribosomal frameshifting. *Proc Natl Acad Sci U S A* **2007**, *104*, 5830-5835.
 - (19) L. Green, C. H. Kim, C. Bustamante, I. Tinoco, Jr., Characterization of the mechanical unfolding of RNA pseudoknots. *J Mol Biol* **2008**, *375*, 511-528.
 - (20) G. Chen, K. Y. Chang, M. Y. Chou, C. Bustamante, I. Tinoco, Triplex structures in an RNA pseudoknot enhance mechanical stability and increase efficiency of -1 ribosomal frameshifting. *Proc Natl Acad Sci U S A* **2009**, *106*, 12706-12711.
 - (21) Z. Zhong, L. Yang, H. Zhang, J. Shi, J. J. Vandana, D. T. Lam, R. C. Olsthoorn, L. Lu, G. Chen, Mechanical unfolding kinetics of the SRV-1 gag-pro mRNA pseudoknot: possible implications for -1 ribosomal frameshifting stimulation. *Sci Rep* **2016**, *6*, 39549.
 - (22) Y. G. Kim, L. Su, S. Maas, A. O'Neill, A. Rich, Specific mutations in a viral RNA pseudoknot drastically change ribosomal frameshifting efficiency. *Proc Natl Acad Sci U S A* **1999**, *96*, 14234-14239.
 - (23) M. Y. Chou, K. Y. Chang, An intermolecular RNA triplex provides insight into structural determinants for the pseudoknot stimulator of -1 ribosomal frameshifting. *Nucleic Acids Res* **2010**, *38*, 1676-1685.
 - (24) C. H. Yu, M. H. Noteborn, C. W. A. Pleij, R. C. L. Olsthoorn, Stem-loop structures can effectively substitute for an RNA pseudoknot in -1 ribosomal frameshifting. *Nucleic Acids Res* **2011**, *39*, 8952-8959.
 - (25) L. X. Yang, Z. S. Zhong, C. L. Tong, H. Jia, Y. R. Liu, G. Chen, Single-Molecule Mechanical Folding and Unfolding of RNA Hairpins: Effects of Single A-U to A.C Pair Substitutions and Single Proton Binding and Implications for mRNA Structure-Induced -1 Ribosomal Frameshifting. *J Am Chem Soc* **2018**, *140*, 8172-8184.
 - (26) D. W. Staple, S. E. Butcher, Solution structure of the HIV-1 frameshift inducing stem-loop RNA. *Nucleic Acids Res* **2003**, *31*, 4326-4331.
 - (27) I. Brierley, F. J. Dos Ramos, Programmed ribosomal frameshifting in HIV-1 and the SARS-CoV. *Virus Res* **2006**, *119*, 29-42.
 - (28) S. Matsumoto, N. Caliskan, M. V. Rodnina, A. Murata, K. Nakatani, Small synthetic molecule-stabilized RNA pseudoknot as an activator for -1 ribosomal frameshifting. *Nucleic Acids Res* **2018**, *46*, 8079-8089.
 - (29) E. Manktelow, K. Shigemoto, I. Brierley, Characterization of the frameshift signal of Edr, a mammalian example of programmed -1 ribosomal frameshifting. *Nucleic Acids Res* **2005**, *33*, 1553-1563.
 - (30) P. J. Farabaugh, Programmed translational frameshifting. *Microbiol Rev* **1996**, *60*, 103-134.
 - (31) J. D. Dinman, Mechanisms and implications of programmed translational frameshifting. *Wiley Interdiscip Rev RNA* **2012**, *3*, 661-673.
 - (32) J. Liphardt, S. Naphthine, H. Kontos, I. Brierley, Evidence for an RNA pseudoknot loop-helix interaction essential for efficient -1 ribosomal frameshifting. *J Mol Biol* **1999**, *288*, 321-335.
 - (33) Y. T. Chen, K. C. Chang, H. T. Hu, Y. L. Chen, Y. H. Lin, C. F. Hsu, C. F. Chang, K. Y. Chang, J. D. Wen, Coordination among tertiary base pairs results in an efficient frameshift-stimulating RNA pseudoknot. *Nucleic Acids Res* **2017**, *45*, 6011-6022.

- (34) M. de Messieres, J. C. Chang, A. T. Belew, A. Meskauskas, J. D. Dinman, A. La Porta, Single-molecule measurements of the CCR5 mRNA unfolding pathways. *Biophys J* **2014**, *106*, 244-252.
- (35) G. Chen, J. D. Wen, I. Tinoco, Jr., Single-molecule mechanical unfolding and folding of a pseudoknot in human telomerase RNA. *RNA* **2007**, *13*, 2175-2188.
- (36) P. E. Nielsen, M. Egholm, An introduction to peptide nucleic acid. *Curr Issues Mol Biol* **1999**, *1*, 89-104.
- (37) M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden, P. E. Nielsen, PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules. *Nature* **1993**, *365*, 566-568.
- (38) S. C. Brown, S. A. Thomson, J. M. Veal, D. G. Davis, NMR solution structure of a peptide nucleic acid complexed with RNA. *Science* **1994**, *265*, 777-780.
- (39) J. Lohse, O. Dahl, P. E. Nielsen, Double duplex invasion by peptide nucleic acid: A general principle for sequence-specific targeting of double-stranded DNA. *Proc Natl Acad Sci U S A* **1999**, *96*, 11804-11808.
- (40) V. V. Demidov, E. Protozanova, K. I. Izvolsky, C. Price, P. E. Nielsen, M. D. Frank-Kamenetskii, Kinetics and mechanism of the DNA double helix invasion by pseudocomplementary peptide nucleic acids. *Proc Natl Acad Sci U S A* **2002**, *99*, 5953-5958.
- (41) I. V. Smolina, V. V. Demidov, V. A. Soldatenkov, S. G. Chasovskikh, M. D. Frank-Kamenetskii, End invasion of peptide nucleic acids (PNAs) with mixed-base composition into linear DNA duplexes. *Nucleic Acids Res* **2005**, *33*, e146.
- (42) J. C. Norton, M. A. Piatyszek, W. E. Wright, J. W. Shay, D. R. Corey, Inhibition of human telomerase activity by peptide nucleic acids. *Nat Biotechnol* **1996**, *14*, 615-619.
- (43) R. Villa, M. Folini, S. Lualdi, S. Veronese, M. G. Daidone, N. Zaffaroni, Inhibition of telomerase activity by a cell-penetrating peptide nucleic acid construct in human melanoma cells. *FEBS Lett* **2000**, *473*, 241-248.
- (44) J. G. Karras, M. A. Maier, T. Lu, A. Watt, M. Manoharan, Peptide nucleic acids are potent modulators of endogenous pre-mRNA splicing of the murine interleukin-5 receptor-alpha chain. *Biochemistry* **2001**, *40*, 7853-7859.
- (45) D. F. Doyle, D. A. Braasch, C. G. Simmons, B. A. Janowski, D. R. Corey, Inhibition of gene expression inside cells by peptide nucleic acids: effect of mRNA target sequence, mismatched bases, and PNA length. *Biochemistry* **2001**, *40*, 53-64.
- (46) N. Dias, S. Dheur, P. E. Nielsen, S. Gryaznov, A. Van Aerschot, P. Herdewijn, C. Helene, T. E. Saison-Behmoaras, Antisense PNA tridecamers targeted to the coding region of Ha-ras mRNA arrest polypeptide chain elongation. *J Mol Biol* **1999**, *294*, 403-416.
- (47) N. Dias, C. Senamaud-Beaufort, E. Forestier El, C. Auvin, C. Helene, T. Ester Saison-Behmoaras, RNA hairpin invasion and ribosome elongation arrest by mixed base PNA oligomer. *J Mol Biol* **2002**, *320*, 489-501.
- (48) M. Kulik, A. Markowska-Zagrajek, M. Wojciechowska, R. Grzela, T. Witula, J. Trylska, Helix 69 of Escherichia coli 23S ribosomal RNA as a peptide nucleic acid target. *Biochimie* **2017**, *138*, 32-42.
- (49) D. F. K. Toh, K. M. Patil, G. Chen, Sequence-specific and Selective Recognition of Double-stranded RNAs over Single-stranded RNAs by Chemically Modified Peptide Nucleic Acids. *J Vis Exp* **2017**.
- (50) P. Thordarson, Determining association constants from titration experiments in supramolecular chemistry. *Chem Soc Rev* **2011**, *40*, 1305-1323.

- (51) S. Y. Lin, A. D. Riggs, Lac repressor binding to non-operator DNA: detailed studies and a comparison of equilibrium and rate competition methods. *J Mol Biol* **1972**, 72, 671-690.
- (52) P. Somogyi, A. J. Jenner, I. Brierley, S. C. Inglis, Ribosomal pausing during translation of an RNA pseudoknot. *Mol Cell Biol* **1993**, 13, 6931-6940.
- (53) D. P. Giedroc, C. A. Theimer, P. L. Nixon, Structure, stability and function of RNA pseudoknots involved in stimulating ribosomal frameshifting. *J Mol Biol* **2000**, 298, 167-185.
- (54) S. Takyar, R. P. Hickerson, H. F. Noller, mRNA helicase activity of the ribosome. *Cell* **2005**, 120, 49-58.
- (55) H. Amiri, H. F. Noller, Structural Evidence for Product Stabilization by the Ribosomal mRNA Helicase. *RNA* **2018**, *In press*.
- (56) Y. Zhang, S. Hong, A. Ruangprasert, G. Skinotis, C. M. Dunham, Alternative Mode of E-Site tRNA Binding in the Presence of a Downstream mRNA Stem Loop at the Entrance Channel. *Structure* **2018**, 26, 437-445 e433.
- (57) C. D. Torgerson, D. A. Hiller, S. Stav, S. A. Strobel, Gene regulation by a glycine riboswitch singlet uses a finely tuned energetic landscape for helical switching. *RNA* **2018**, 24, 1813-1827.
- (58) B. Liu, N. Shankar, D. H. Turner, Fluorescence competition assay measurements of free energy changes for RNA pseudoknots. *Biochemistry* **2010**, 49, 623-634.

Chapter 4

Single-Molecule Mechanical Folding and Unfolding of Disease-causing Tau pre-mRNA Exon 10 Splice Site Hairpins: Characterization of Secondary Structure Rearrangement and the Implications for Alternative Splicing

4.1 Introduction

Frontotemporal dementia with Parkinsonism linked to chromosome-17 (FTDP-17) is an autosomal dominant neurological disorder in adult human brain that causes behavioural disturbances, aphasia, cognitive impairment, and Parkinsonism.¹⁻⁵ It has been found that microtubule-associated protein tau (*MAPT*) in neurons, which associates the microtubule assembly, stability, and function, is linked to the cause of FTDP-17.²⁻⁸ Tau is abundantly expressed in neurons of the central nervous system. Abnormal aggregations of tau proteins are known hallmarks of various neurodegenerative conditions.^{2,9,10}

Alternative splicing of exons 2, 3, and 10 of the *MAPT* pre-mRNA transcript from the tau gene can generate six tau isoforms. The alternative splicing of exon 10 produces tau isoforms with four microtubule-binding repeats (4R) or three repeats (3R) in the case of exon 10 inclusion or skipping, respectively (**Figure 4.1A–C**). In healthy neurons, the ratio of 4R:3R is maintained at roughly 1:1.^{1,2,6,11} Excessive 4R isoforms generated can cause tau proteins aggregation, while overabundant 3R isoform production is speculated to trigger apoptosis

pathways.^{2,12,13} Mutations in the *MAPT* gene cause altered pre-mRNA splicing and subsequent imbalance of tau proteins, and lead to neurodegenerative diseases.^{2,4,9,10,14} Studies revealed a stem-loop structure composed of 18–25 nucleotides at the exon 10 5' splice site (5'ss, **Figure 4.1B,C**) was involved in the regulation of exon 10 alternative splicing.^{3,7,15–18} A 25-nucleotide (nt) hairpin (from residues –6 to +19) is composed of two stems (top stem and bottom stem) flanking an A bulge (**Figure 4.1B**). An 18-nt hairpin (from residues –2 to +16) contains only the top stem. The hairpin formation prevents recognition of the 5'ss by the U1 small nuclear RNA (U1 snRNA) (**Figure 4.1A**).¹⁹ Recognition of 5'ss by base pairing with U1 snRNA is a key initial step in splicing, by using multiple registers and noncanonical base pairs, through the formation of a maximum of 11 possible base pairs.^{20,21}

Figure 4.1. *MAPT* exon 10 5' splice site. (A) U1 snRNA recognition of *MAPT* exon 10 5'ss sequence. U1 snRNA can form 7 base pairs with the 5'ss

(highlighted in gray). (B) Proposed structures for 25-nt and 30-nt wild-type and +19G *MAPT* exon 10 5'ss regulatory hairpins. The structures were predicted by the RNAstructure program with the free energy values (1M NaCl, 37 °C) shown. (C) Schematic of *MAPT* pre-mRNA around exon 9 (gray box), exon 10 (white box), and exon 11 (gray box). Exons are represented by the numbered boxes, introns being the intervening lines. The regulatory hairpin is located at the junction of exon 10 and intron 10. 4R and 3R isoforms are generated based on the inclusion or skipping of *MAPT* exon 10. (D) Construction of the UMV *MAPT* exon 10 hybrid minigene for splicing assay. We cloned the *MAPT* exon 10 (white box) and flanking intronic sequences (dashed lines) into the UMV minigene. The flanking exons are constitutively spliced.

Alternative splicing of tau pre-mRNA may also be regulated by *trans*-acting factors. For example, DEAD-Box Helicase 5 (DDX5, also known as RNA helicase p68)²²⁻²⁴ may facilitate the unwinding of the hairpin and/or the duplex formed between 5'ss and U1 snRNA, thereby increasing exon 10 inclusion. Splicing Factor Proline- and Glutamine-Rich (SFPQ), also known as polypyrimidine tract binding protein associated splicing factor,²⁴⁻²⁷ is also involved in regulating tau pre-mRNA exon 10 splicing. Other *trans* factors may also regulate alternative splicing of tau.^{4,28-31}

Several FTDP-17-associated point mutations occur within the hairpin region (**Figure 4.1B,C**).^{3,7,15-18,30} The mutations at the top stem typically result in the disruption a base pair, and tend to increase the 4R:3R ratio. However, a single C-to-G mutation in the bottom stem at the 19th nucleotide downstream of the 5'ss

(+19G) causes a significant decrease in 4R:3R ratio.¹² This is surprising because the +19G mutation is expected to disrupt a G-C base pair, (G-6)-(C+19), in the bottom stem of the hairpin (**Figure 4.1B**).^{12,31}

During the alternative splicing of exon 10 in *MAPT* pre-mRNA, the U1 snRNP binding to the 5' splice-site sequence of exon 10 would require the unfolding of the downstream RNA hairpin structure.³¹ The structural information and mechanical properties of this downstream RNA were not well studied. Single-molecule unfolding of the tau splice site mRNA structure and its disease-causing mutants using optical tweezers can help probe the hairpin structures, their mechanical stability and (un)folding kinetics and how these mechanical properties are related to the alternative splicing efficiency of exon 10.

Here, we investigated how the +19G mutation in the structure of the tau pre-mRNA exon 10 5'ss hairpin decreases exon 10 inclusion by mini-gene splicing assay in cell culture, and biophysical assays including oligonucleotide thermal melting and single-molecule mechanical unfolding using high-resolution optical tweezers. Our results indicate that, for the +19G mutation, an expanded hairpin of 30 nt (+19G-30nt) instead of 25 nt (+19G-25nt) (**Figure 4.1B**) needs to be considered for correlating splicing activity and hairpin structure and stability. The biophysical data of the +19G-30nt and other mutants suggest that +19G mutation induces a conformational change that leads to the formation of a 27-nt hairpin with the formation of a new bottom stem with enhanced stability. The stabilized bottom stem may further enhance the local stability of the splice site at the top stem through coaxial stacking, resulting in the inhibition of exon 10

inclusion. Furthermore, we characterized how the stabilizing +19G mutation in the bottom stem and various destabilizing mutations in the top stem of the hairpin coordinate in regulating splicing through the regulation of hairpin structure and stability in a position dependent manner. Mechanical unfolding directly reveals the structural changes upon single-nucleotide mutations. Our single-molecule force spectroscopy studies reveal an excellent correlation between splicing activity and mechanical unfolding rate at 10 pN, mimicking the unfolding of the splice site hairpins aided by helicases.

4.2 Materials and Methods

4.2.1 RNA Secondary Structure Prediction

We used the online version of the RNAstructure software^{32,33} to predict the structures and free energies of various *MAPT* exon 10 5'ss hairpins. We inputted the sequences of the 25nt and 30nt WT hairpins: GGCAGU|GUGAGUACCUUCACACGUC and GGCGGCAGU|GUGAGUACCUUCACACGUCCC respectively; the position of the 5'ss is marked by the vertical bar. We produced predictions for each mutant hairpin by making the appropriate sequence alteration(s) corresponding to their respective mutation(s). We kept all other parameters at default settings (1 M NaCl, 37 °C).

4.2.2 *MAPT* Exon 10 Hybrid Minigenes Construction and Splicing Assay in Cells

MAPT exon 10 hybrid minigenes in UMV²¹ were constructed by our lab members in order to test the change in splicing caused by the various point mutants. Briefly, the exon 10 of *MAPT* flanked by intronic sequences extending 39nt upstream and 63nt downstream was amplified using plasmids containing the relevant sequences³¹ as templates and proper PCR primers (sequences available upon request). This fragment was then cloned into UMV to form the final hybrid wild-type (WT) minigene.²¹ Mutations in the minigene were introduced through PCR mutagenesis as described previously,²¹ using appropriate pairs of primers, and the wild-type minigene as the template. The *MAPT* hybrid minigenes were then mixed with control plasmids at a 1:11 ratio to create the DNA mix, with which ~50% confluent HEK293T cells were transfected. The transfected cells were harvested, extracted and reverse transcribed the RNA. UMV-expressed cDNAs were amplified using proper primer pair. Semi-quantitative PCR were performed, and the PCR products were separated by 10% native PAGE at 10 V/cm for 6 h 30 min in 1x TBE buffer. The gels were imaged with Typhoon Trio Variable Mode Imager (GE Healthcare Life Sciences, USA). The gel bands were quantified to obtain the exon inclusion percentages from the experimental triplicates.

4.2.3 Thermal Melting Experiments

We carried out the UV-absorbance detected thermal melting experiments using a Shimadzu 2550 spectrometer. The concentration of all RNA samples used was fixed at 2 μ M. The RNA samples were snap-cooled prior to the UV melting

experiments. During the experiments, the temperature was raised from 15 °C or 20 °C to 95 °C with a ramp rate of 0.5 °C/min with interval of 0.5 °C, and the absorbance at 260 nm (A_{260}) were recorded. We obtained the melting temperatures from the first derivatives of the UV absorbance melting (heating) curves. The melting experiments were conducted in two buffer conditions: buffer 1 (200mM NaCl, 10mM HEPES, 0.1mM EDTA, pH 7.3), without MgCl₂ and spermine, and buffer 2 (5 mM MgCl₂, 100 μ M spermine, 200 mM NaCl, 10 mM HEPES, 0.1 mM EDTA, pH7.3). The measurements for each individual sample were performed thrice.

4.2.4 Non-denaturing Polyacrylamide Gel Electrophoresis

RNA oligonucleotides for the 25-mer hairpins, 30-mer hairpins, and short RNA oligonucleotides U1R1 (U1R1: 5'-AUACUUACCUG-3') for non-denaturing polyacrylamide gel electrophoresis (PAGE) studies were purchased from Sigma. The hairpins were folded in the presence of U1R1 oligonucleotide by slow cooling from 90 °C to room temperature and incubation at 4 °C overnight. Samples were incubated with 200 mM NaCl, 10 mM HEPES, 0.1 mM EDTA, pH 7.3. Prior to the sample loading, 35% glycerol (20% of the total volume) was added to the sample mixtures. 20 μ L of each sample (1 μ M RNA with varied U1R1 concentrations) was loaded onto a 12% non-denaturing polyacrylamide gel (with a 19:1 acrylamide to bisacrylamide ratio). The gels were run at 250 V for 5 h at 4 °C in 1 $\times$ TBE (89 mM Tris, 89 mM Borate, 2 mM EDTA, pH 8.3), and post-stained with ethidium bromide. We imaged the gels by a Typhoon Trio Variable Mode Imager.

4.2.5 Single-molecule Optical Tweezers Experiments

The *puc19* plasmids containing the target sequences of WT (5'-GGCGGCAGUGUGAGUACCUUCACACGUCCC-3'), +19G (5'-GGCGGCAGUGUGAGUACCUUCACACGUGCC-3'), +14U (5'-GGCGGCAGUGUGAGUACCUUCAUACGUCCC-3') and +14U+19G (5'-GGCGGCAGUGUGAGUACCUUCAUACGUGCC-3') at the exon 10 5' splice site were constructed previously.³¹ They were used as templates for PCR amplification using the following primers: The forward primer PUC19-F with the sequence 5'-AATTAACCCTCACTAAAGCTTGTCTGTAAGCGGATGCCG-3' (T3 promoter underlined) and the reverse primer PUC19-R with the sequence 5'-CAGGGTCGGAACAGGAGAGC-3'. *In vitro* transcription by T3 RNA polymerase was performed to generate the RNA molecules used in the single-molecule mechanical (un)folding studies. As described previously,³⁴⁻³⁷ the RNA molecules were then annealed with two PCR-amplified supplementary DNA sequences to form the RNA/DNA hybrid handles. The 5' terminal of the handle B was labelled with digoxigenin, which was introduced by using digoxigenin-labelled reverse primer during PCR. The 3' terminal of handle A was labelled with biotin through biotin modified dUTP. During the experiments, handle B would attach to an anti-digoxigenin-coated polystyrene bead (about 3 μm in diameter) and be trapped by a dual-laser optical trap. In the same time, handle A would attach to a streptavidin-coated polystyrene bead (about 1.7 μm in diameter) and be held onto a micropipette by suction. Experiments were done in three modes, force-ramp mode, constant-force mode and passive mode. In a force-ramp mode, the force is loaded and unloaded at a rate of 4 pN/s. In a constant-force mode, the force is maintained at certain values

with an electronic force feedback of 5 ms. In a passive mode, the positions of the pipette bead and the trap are fixed, with the force showing a sudden drop/increase when the RNA structure unfolds/folds. The data acquisition rate was 1000Hz. The experiments were carried out at 22 ± 1 °C in buffer 2: 5 mM MgCl₂, 100 μM spermine, 200 mM NaCl, 10 mM HEPES, 0.1 mM EDTA, pH 7.3.

4.3 Results and Discussion

4.3.1 Thermal Melting of Expanded 30-mer Splice Site Hairpin Constructs Reveals a Stabilization Effect of +19G Mutation

We used the RNAstructure program (<https://rna.urmc.rochester.edu/RNAstructure.html>)^{32,33} to predict the structure and folding free energy (1 M NaCl, 37 °C) of the +19G mutant hairpin and the wild-type (WT) with varied sequence length. Interestingly, the 30-mer +19G mutant hairpin is predicted to have a structural rearrangement forming an extended bottom stem below the (G-2)-(C+16) base pair (**Figure 4.1B**). Upon the structural rearrangement, the bulge residue A-3 becomes base paired with the residue U+18, with G+17 instead forming a bulge (**Figure 4.1B**). Thus, instead of weakening the hairpin due to the disruption of the (G-6)-(C+19) pair, the +19G mutation is predicted to result in the formation of an extended 4-base pair (bp) bottom stem. Such a local secondary structural rearrangement is only possible for a 30-nt hairpin sequence^{4,18} but not for the 25-nt one often studied previously (**Figure 4.1B**).^{15,31,38-42}

Figure 4.2. Normalized thermal melting curves monitored at 260 nm in buffer 2 (100 μ M spermine, 5 mM MgCl_2 , 200 mM NaCl, 10 mM HEPES, 0.1 mM EDTA, pH7.3). (A) Representative melting curves for 25-mer constructs: WT-25nt, +14U-25nt and +19G-25nt. (B,C) Representative melting curves for 30-mer constructs of (B) WT-30nt and disease-causing mutants (+14U, +19G, +3A, and +11C), and (C) 30-mer artificial mutants (17Uins-30nt, +19G+17U-30nt, +10C-30nt, +17A-30nt, +18C-30nt and +14U+19G-30nt). Only the heating trajectories are shown. The heating and cooling curves are not superimposable.

To experimentally detect the potential structural rearrangement, we carried out UV absorbance-detected thermal melting experiment for a series of RNA oligonucleotide hairpins (**Figure 4.2, Table 4.1**). The UV thermal melting data show that the T_m values of the 25-mer +19G-25nt and WT-25nt hairpins are almost the same, consistent with previously reported results.³¹ The bottom two base pairs (G-U and G-C) below the A bulge in the 25-mer WT hairpin are not stably formed as evidenced by previous NMR results.¹⁶ Consistently, we observed no significant change of T_m in the 25-mer constructs (76.6 $^{\circ}\text{C}$ for +19G-25nt versus 76.0 $^{\circ}\text{C}$ for 25-mer WT-25nt, in buffer 2, **Figure 4.2**), similar to previously reported thermal melting data.³¹ We tested two buffers for the thermal melting experiments, and see the same trend for the T_m values. We did not obtain equilibrium thermodynamic parameters due to the hysteresis observed.

Table 4.1. Oligonucleotide thermal melting results.

	T_m (°C)	
	Buffer 1 ^a	Buffer 2 ^a
WT-25nt	72.0±0.3	76.0±0.3
+19G-25nt	72.4±0.3	76.6±0.3
+14U-25nt	55.5±0.3	59.4±0.4
WT-30nt	70.3±0.4	75.5±0.4
+17Uins-30nt ^b	62.2±0.8, 80.9±0.4	65.1±0.6, 85.8±0.3
+18C-30nt	78.4±0.2	82.9±0.4
+19G-30nt ^b	54.8±0.2, 74.4±0.2	59.6±1.1, 79.9±0.1
+17A-30nt	70±0.3	74.4±0.4
+11C-30nt	66.1±0.5	71.9±0.3
+14U+19G-30nt	61.9±0.3	67.7±0.2
+14U-30nt	57.2±0.3	61.9±0.3
+3A-30nt	56.2±0.4	62.2±0.3
^a Buffer 1: 200mM NaCl, 10mM HEPES, 0.1mM EDTA, pH 7.3. Buffer 2: 200mM NaCl, 5mM MgCl ₂ , 100μM spermine, 10mM HEPES, 0.1mM EDTA, pH 7.3.		
^b An extra small low-temperature melting was observed.		

The 30-mer hairpin WT-30nt shows a T_m value similar to WT-25nt (**Figure 4.2**), suggesting that extending the bottom stem with a C bulge and two G-C pairs does not enhance the stability of the top stem. Significantly, the 30-mer +19G-30nt hairpin shows a T_m value 4 °C higher than that of the 30-mer WT-30nt hairpin (79.9 versus 75.5 °C). The thermal melting results clearly suggest the formation of a stabilized bottom stem in +19G-30nt, which cooperatively melts with the top stem. The coaxial stacking of the bottom stem with the top stem causes the stabilization of splice site-containing top stem, and may thus disfavor exon 10 inclusion. A small broad low-temperature melting transition is observed for +19G-30nt and +17U+19G (**Figure 4.2B,C**), which might be due to the melting of the bottom stem alone for a small fraction of the ensemble in solution. The non-denaturing PAGE data suggest that all the hairpin constructs fold properly (**Figure 4.3**). Taken together, the direct comparison of thermal melting data for

the 25-mer and 30-mer hairpin constructs shows a secondary structure rearrangement of the 30-mer +19G splice site hairpin, demonstrating the importance of the hairpin structure in regulating alternative splicing of tau pre-mRNA.

Figure 4.3. Non-denaturing PAGE results for RNA oligonucleotides. The same constructs were used for the thermal melting studies. (A) 25-mer RNA oligonucleotides consisting of residues from -6 to +19 with RNA concentrations of 1 μM (on the left) and 2 μM (on the right). (B) RNA oligonucleotides consisting of residues from -9 to 21 with RNA concentrations of 1 μM (on the left) and 2 μM (on the right). There're no dimers or duplexes forming at RNA concentrations of 2 μM (the same RNA concentration used in melting study), ruling out dimers or duplexes contributing to the two melting transitions of the RNA oligonucleotides.

4.3.2 Cell Culture pre-mRNA Splicing Assay Shows that a Stabilizing Mutation in the Bottom Stem Results in Inhibition of the 5'ss Recognition on the Top Stem

We carried out cell culture splicing minigene mutational study (**Figure 4.4A**) to correlate the structural rearrangement with alternative splicing activity. We constructed UMV hybrid minigenes containing *MAPT* exon 10 (93 nt) with its flanking intronic sequences (39 nt upstream and 63 nt downstream, see Materials and Methods, and **Figure 4.1D**). We then transfected HEK293T cells with the minigenes and characterized the pre-mRNA splicing patterns via semi-quantitative RT-PCR.

Figure 4.4. +19G rearrangement affects splicing in cells. (A) Proposed structures for the 30-nt wild-type and various mutated exon 10 5'ss regulatory hairpins. The structures were predicted by the RNAstructure program with the free energy values (1M NaCl, 37 °C) shown. (B,C) Representative RT-PCR data of the cell culture splicing assays. The exon 10 inclusion percentage values were derived

from three experimental replicates (samples from independent transfections). The standard deviations are less than 10%.

The +19G minigene shows a close to 0% exon 10 inclusion, which is significantly lower than that of the WT (56%) (**Figure 4.4B**), consistent with the low exon 10 inclusion levels previously reported.^{12,31} Deletion of the G+17 residue in the +19G minigene (+19G+17Gdel, **Figure 4.4A**) results in no change in splicing (**Figure 4.4B**). The result is consistent with the fact that, in +19G hairpin, G+17 residue is forming a bulge, without forming significantly stabilizing interactions. We made a -4G mutant, which is predicted to have a very similar structural rearrangement as that of +19G (**Figure 4.4A**). Indeed, we observed that the -4G mutant minigene shares a very similar splicing pattern as that of +19G (**Figure 4.4C**, Lane 4). Furthermore, the two double mutants -4G+17A and +17A+19G, which are predicted to form an A+17 bulge and a 4-bp bottom stem, also show close to 0% exon 10 inclusion (**Figure 4.4C**). Clearly, a stabilized bottom stem away from the splice site can significantly inhibit splicing activity. The +17A point mutation is predicted to weaken the bottom stem forming a relatively unstable 2×2 internal loop (**Figure 4.4A**). We did observe a slight decrease in the T_m of the +17A-30nt hairpin compared to WT-30nt. +17A mutant minigene shows a relatively large increase in exon 10 inclusion (70%, **Figure 4.4B**), probably due to the formation of a relatively flexible 2×2 internal loop.

In the +19G-30nt construct, the residue G+19 is predicted to pair with residue C-4 in the bottom stem. It is likely that other interactions in the +19G mutant may also form resulting in the inhibition of splicing. For example, it was

proposed that the residue G+19 might not form a base pair in the bottom stem, but may instead pair with residue C-40, strengthening the stability of a large RNA secondary structure.³¹ Indeed, we observed that the -40A point mutant gene shows a significantly increased exon inclusion level (93%, **Figure 4.4B**, lane 8), which is consistent with a destabilizing effect expected for the disruption of a potential pair, (G+19)-(C-40).³¹ However, the observed increase in exon 10 inclusion for the -40A point mutant gene could also be due to the disruption of a known exonic splicing silencer element and/or other indirect effect.^{29-31,43,44} We generated a double mutant -40A+19G, which disrupts the proposed (C-40)-(G+19) pair. The double mutant -40A+19G shows an exon inclusion level (1%), similar to +19G (**Figure 4.4B**, lane 9). Thus, +19G mutation can have a dominant effect in stabilizing the splice site hairpin through the formation of a stabilized bottom stem, accounting for the observed significant decrease in splicing of *MAPT* exon 10.

A bottom stem-stabilizing mutant +18C (a U-to-C substitution at position +18) shows a comparable splicing activity as +19G (**Figure 4.4A,C**), further suggesting that a bottom stem could be important in stabilizing the top stem, resulting the inhibition of exon inclusion. The +17Uins mutation (**Figure 4.4A**, also known as the WTI-17T mutation^{15,31}) involves the insertion of an additional U nucleotide between the positions +16 and +17, forming a pair with residue A-3, resulting in the removal of the bulge structure. Consistent with previously reported results,^{15,31} the +17Uins mutant shows a reduced exon 10 inclusion level (8%, **Figure 4.4C**). It is interesting to note that both the 30-mer +19G and +18C hairpins are predicted to be less stable than that of +17Uins, as confirmed by the

experimental thermal melting T_m values (**Figure 4.2**), although +19G and +18C show slightly lower exon 10 inclusion than that of +17Uins (**Figure 4.4C**). Additionally, the exon 10 inclusion level of the double mutant +17Uins+18C (3%) is similar to that of +18C (2%) (**Figure 4.4C**). Similarly, +19G and +17Gdel+19G have the same exon 10 inclusion level (1%) (**Figure 4.4C**). Thus, extending the splice site-containing top stem with an A-U pair by removing the bulge residue below (G-2)-(C+16) pair may enhance the inhibition of splicing; stabilizing the bottom stem by increasing the content of G-C pairs is, however, clearly more effective in the inhibition of splicing.

We further generated the +17Ains+18Uins and the +17Uins+18Ains minigenes (**Figure 4.4A**) by inserting dinucleotide sequences 5'AU3' and 5'UA3', respectively, between the positions +16 and +17. Interestingly, +17Ains+18Uins has a significantly higher exon 10 inclusion level than +17Uins+18Ains (20% versus 8%), although the hairpins are predicted to have the same folding free energy (**Figure 4.4A,C**). The data further suggest that extending the splice site-containing top stem with an A-U pair immediately below (G-2)-(C+16) pair (+17Uins+18Ains) can indeed stabilize the top stem and thus inhibit the recognition of the 5'ss, with comparable exon 10 inclusion level as +19G (8% versus 1%). Extending the bottom stem with an A-U pair (+17Ains+18Uins), however, is less effective in blocking the recognition of the 5'ss, compared to +19G (20% versus 1%). The data also further suggest that increasing the G-C pair content in the bottom stem (e.g., +18C and +19G) can significantly stabilize the top stem and thus inhibit splicing.

4.3.3 Disease-causing Destabilizing Mutations on the Top Stem Rescue the Splicing of +19G Mutant

The majority of disease-causing mutations at the top stem (**Figure 4.1B**) destabilize the hairpin structure resulting in an increased tau exon 10 inclusion. We hypothesized that the bottom stem-stabilizing disease-causing +19G mutation may counteract the destabilizing mutations on the top stem (**Figure 4.5A**), resulting in the rescue of aberrant splicing. We generated minigenes by combining the +19G mutation with various disease-causing mutations known to increase exon inclusion, namely -2A, +3A, +11C, +12U, +13G, +14U (also known as DDPAC), +15C, and +16U, respectively.^{7,15,31} The bottom stem structural rearrangement (+19G rearrangement) is predicted to occur in all the double mutant hairpins, resulting in a stabilization of the hairpin structure. Consistently, the oligonucleotide thermal melting reveals T_m values of 79.9, 75.5, 67.7, and 61.9 °C, respectively, for +19G-30nt, WT-30nt, +14U+19G-30nt, +14U-30nt (**Figure 4.2**).

Figure 4.5. Position-dependent counteracting effect of the top stem mutations on +19G mutation. (A) Mutants studied. The region showing a dominant effect is boxed in dashed line. (B) Representative RT-PCR data of the cell culture splicing assays. The exon 10 inclusion percentage values were derived from three

experimental replicates (samples from independent transfections). The standard deviations are less than 10%.

The +11C and the +16U mutants show high exon 10 inclusion levels of 96% and 100%, respectively (**Figure 4.5B**). Upon the +19G substitution on the two mutants, we observe a significant decrease in exon 10 inclusion level (24% and 31%, respectively, for +11C+19G and +16U+19G, **Figure 4.5B**). Clearly, a stabilizing +19G mutation on the bottom stem of the hairpin can compensate the destabilizing effect of the +11C and +16U.

Remarkably, the mutants with the disease-causing mutations from positions +12 to +15 (**Figure 4.5B**, lanes 9–14, 17–18), as well as positions –2 and +3 (**Figure 4.5B**, lanes 3–6), are resistant to the influence of the +19G mutation. Only a small decrease (< 5%) is observed in exon inclusion induced by the +19G mutation. Furthermore, the double mutant +14U+18C also shows a relatively high exon inclusion (88%, **Figure 4.5B**, lane 15), in comparison to +14U (99%, lane 13), +18C (2%, lane 16), +14U+19G (95%, lane 14), and WT (56%, lane 1). Clearly, the destabilizing effect of +14U and other mutations located near the 5'ss in the hairpin is not efficiently rescued by the stabilizing mutation +19G on the bottom stem (**Figure 4.5A**). It is probable that a mutation adjacent to the splice site (**Figure 4.5A**, red dashed box) is critical in determining the splicing activity. For example, the +14U mutation results in the weakening of the base pair, (G+1)-(C+14), which is immediately next to the 5'ss-containing hairpin. Similarly, stabilizing mutations (e.g., +11C and insertion of a U residue between C+16 and G+17) have shown inefficient inhibition of splicing on the +14U mutant.¹⁵ Note

that the -2A and +3A mutations can strengthen the tau exon 10 5'ss, by forming a more stable duplex with U1 snRNA (see **Figure 4.1A**), in addition to destabilizing the 5'ss-containing hairpin. Indeed, +3A-30nt hairpin shows a relatively low T_m of 62.2 °C (**Figure 4.2**). Taken together, our data show that the splicing activity is critically determined by the local stability of the stem structure surrounding 5'ss, and the overall hairpin stability may not always correlate with splicing activity.

4.3.4 Correlations Between Splicing Activity and T_m and Predicted Folding Free Energy

We obtained the thermal melting data for WT, +14U, and +19G as both 25-mer and 30-mer hairpin constructs. A direct comparison of the 25-mer and 30-mer constructs clearly suggest that the T_m values of the 30-mer constructs have an improved correlation with splicing activity (**Figure 4.6A**). A poor correlation is also observed between splicing activity and predicted free energy for the 25-mer constructs (**Figure 4.6B**). We next focused on the analysis on the 30-mer constructs. The exon inclusion levels correlate with the T_m values of the 30-mer hairpins (**Figure 4.6C**), which is similar to the previous reported result for 25-mer hairpins without the +19G mutant.¹⁵ It is probably because the thermal melting transitions largely reflect the stability of the top stem, because the bottom stem without the structural rearrangement as observed for +19G mutant is relatively unstable as evidenced by the NMR.¹⁶

Figure 4.6. Potential correlations between exon 10 inclusion percentage (4R%) and experimental thermal melting T_m value and predicted folding free energy. The structures and folding free energy values (1M NaCl, 37 °C) were predicted by the RNAstructure program. (A) 4R% shows a better correlation with the T_m (in buffer 2) for the 30-mer than the 25-mer hairpins. (B) Comparison of 4R% and predicted free energies for all hairpins containing 25 nucleotides (consisting of residues from -6 to +19). (C) Comparison of 4R% and T_m (in buffer 2) for all the 30-mer hairpins. The T_m values of the major transitions were used, if there are two melting transitions. (D) Comparison of 4R% and predicted free energies for all hairpins containing 30 nucleotides. Two classes of minigenes (shown in red open squares) have a poor correlation between 4R% and predicted folding free energy. The first class of minigenes include the double mutants +13G+19G, +14U+18C, +14U+19G, and +12U+19G (ordered with increasingly favorable predicted free energy), which have mutations at the positions within three base

pairs from the splice site. The mutant +15C+19G is not excluded for analysis because a significant structural rearrangement in the top stem occurs upon +15C mutation. The mutants -2A+19G and +3A+19G are not excluded for analysis because -2A and +3A mutations result in strengthening of the splice site (see **Figure 4.1A**). The second class of minigenes include +17Uins+18C, +17Gdel+19G and +17Uins (ordered with increasingly favorable predicted free energy), which have the bulge residue removed. Removing the two classes of outliers results in a significant improvement of the correlation between exon 10 inclusion level and folding free energy (adjusted R^2 values of 0.60 versus 0.83).

Due to the stabilization of the bottom stem upon structural rearrangement for the mutants (e.g., +19G and -4G), which in turn stabilizes the top stem by coaxial stacking, it becomes valid to observe correlation between predicted free energies of the 30-mer constructs and splicing activities ($R^2 = 0.60$, **Figure 4.6D**). Two classes of minigenes have a poor correlation between exon 10 inclusion level and folding free energy. The first class of minigenes include the double mutants +13G+19G, +14U+18C, +14U+19G, and +12U+19G (ordered with increasingly favorable predicted free energy), which have mutations at the positions within three base pairs from the splice site. The mutant +15C+19G is not excluded for analysis because a significant structural rearrangement in the top stem occurs upon +15C mutation. The mutants -2A+19G and +3A+19G are not excluded for analysis because -2A and +3A mutations result in strengthening of the splice site (see **Figure 4.1A**). The second class of minigenes include +17Uins+18C, +17Gdel+19G and +17Uins (ordered with increasingly favorable predicted free energy), which have the bulge residue removed. Removing the two classes of

outliers results in a significant improvement of the correlation between exon 10 inclusion level and folding free energy (adjusted R^2 values of 0.60 versus 0.83, **Figure 4.6D**).

We further tested if U1 snRNA can directly bind to the splice site hairpins. An 11-mer unmodified RNA oligonucleotide U1R1 (5'-AUACUUACCUG-3', see **Figure 4.1A**) with the same sequence as that of the 5' end of U1 snRNA was used to mimic U1 snRNA. We did not observe appreciable binding by PAGE (**Figure 4.7**). Thus, *trans* factors such as helicases are needed to unwind the hairpin structures to facilitate U1 snRNA binding.

Figure 4.7. Non-denaturing PAGE results for RNA oligonucleotides targeted by U1R1 (5'-AUACUUACCUG-3'), mimicking the 5' end of U1 snRNA. No binding is observed between the U1R1 oligonucleotide and any of the (A–C) 25-mer or (D–F) 30-mer RNA hairpins. The RNA hairpin concentrations are 1 μ M. The U1R1 concentrations are varied and from left to right are 0, 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 50 μ M.

4.3.5 Single-molecule RNA Hairpin Force Measurements Directly Reveal +19G Hairpin Structural Rearrangement

We used force-measuring optical tweezers to mimic the unwinding of the splice site hairpins by helicases and spliceosome complex (**Figure 4.8A**). We carried out single-molecule studies for WT, +19G, +14U, and +14U+19G hairpins (**Figure 4.8B**). The +14U mutant has been studied extensively.^{15,22,25,38} The WT, +19G, +14U, and +14U+19G hairpins show a one-step mechanical unfolding transition (**Figure 4.8C,D, Figures 4.9 and 4.10**). In a constant-force mode, the extension trajectories directly reveal the real-time folding/unfolding dynamics of a hairpin, with large and small extensions corresponding to unfolded and folded states, respectively (**Figure 4.9**). In a passive mode, the force trajectories directly reveal the real-time folding/unfolding dynamics of a hairpin, with larger and smaller forces corresponding to folded and unfolded states, respectively (**Figure 4.10**). The end-to-end extension changes at the equilibrium forces for the studied WT, +19G, +14U, and +14U+19G are 6.5 ± 1.4 , 9.4 ± 1.3 , 6.2 ± 1.7 , and 8.3 ± 1.6 nm, respectively (**Table 4.2, Appendix Tables A2.1 and A2.2**), which correspond to folded hairpins consisting of 20.2 ± 3.3 , 27.1 ± 3.1 , 21.2 ± 4.4 , and 25.3 ± 3.9 nucleotides, as estimated by an extensible wormlike chain (EWLC) model.⁴⁵⁻⁴⁹ Thus, the single-molecule pulling experiment directly reveal that +19G hairpin has more nucleotides folded compared to the WT (27 nt versus 20 nt).

Figure 4.8. Single-molecule pulling experiments directly reveal the +19G structural rearrangement. (A) Optical tweezers setup scheme. (B) Schematic of the RNA constructs used in the single-molecule mechanical pulling experiments. The ssRNA linkers are shown in gray. The secondary structures are shown based on the measured number of nucleotides folded in the mechanical pulling experiments. WT and +14U only have 21 nucleotides folded during the pulling. +19G has 27 nucleotides folded during the pulling, consistent with the structure rearrangement hypothesis. +14U+19G is proposed to adopt two conformations similar to that of +19G (type 1) and WT (type 2), respectively. The unfolding and folding transition state positions are indicated with open arrow and solid arrow, respectively. (C) Representative force-ramp curves. Force-loading and force-unloading curves are shown in colors and black, respectively. The curves are smoothed to 200 Hz. (D) Representative constant-force mode traces for WT and +19G at 14.1 pN. Compared to WT, +19G shows a larger extension change upon

(un)folding and with a slower unfolding rate, and a more populated hairpin structure. The data are smoothed to 200 Hz.

Figure 4.9. Representative constant-force traces. (A) WT at 13.4 pN; (B) +19G at 14.1 pN; (C) +14U at 9.9 pN; and (D) +14U+19G at 11.6 pN. The curves are smoothed to 200 Hz. The force distributions are fit to one Gaussian distributions (binned to 0.05 pN). The extension distributions are fitted to two Gaussian distributions (binned to 1 nm). The extension curves were fitted to hidden Markov model.⁵⁰ The fitting trajectories are shown in black. A one-step mechanical unfolding transition is observed. The dashed cyan and green lines correspond to the unfolded states and folded states, respectively.

Figure 4.10. Representative passive-mode traces of (A) WT, (B) +19G, (C) +14U and (D) +14U+19G at around critical forces. The data acquisition rate was 1000 Hz. The force distributions (binned to 0.05 pN) and extension distributions (binned to 1 nm) are fitted to two Gaussian distributions. The extension curves were fitted to hidden Markov model.⁵⁰ The fitting trajectories are shown in black. The dashed cyan and green lines correspond to the unfolded states and folded states, respectively.

The WT and +14U hairpins contain about 20 nucleotides folded during the pulling process, even though the constructs consist of more than 30 nucleotides. Thus, the bottom 9 bases of the WT-30nt and +14U-30nt hairpins are not stably paired (**Figure 4.2B**), which is in agreement with the previous NMR results for a 25-mer WT-25nt.¹⁶ The fact that +19G hairpin has 27 nucleotides folded during the pulling process further confirms that +19G adopts a rearrangement in secondary structure (**Figure 4.2B**). The +14U+19G hairpin adopts a similar secondary structure with a slightly decreased number of nucleotides folded

compared to +19G (25 nt versus 27 nt), suggesting a similar structural rearrangement present in +14U+19G hairpin.

Figure 4.11. Example of (un)folding rates extraction from constant-force data.

(A) Force trajectory and the corresponding extension trajectory (blue) of WT at 13.7 pN from constant-force mode. The curves are smoothed to 200 Hz. The extension curves were fitted to hidden Markov model.⁵⁰ The fitting trajectories are shown in black. (B) Histogram of lifetime distributions of folded state and unfolded state fitted to single exponential equation based on one-step reaction model for extracting unfolding and folding rates, respectively. The mean lifetime of the folded state (τ_{folded}) and the unfolded state (τ_{unfolded}) obtained are 0.059 ± 0.001 s ($R^2 = 0.998$) and 0.083 ± 0.003 s ($R^2 = 0.996$), respectively. The uncertainties are reported as 95% confidence interval for the single-exponential fitting.

The extension and force trajectories were fitted using a hidden Markov model (**Figure 4.11**).⁵⁰ The histograms of lifetime distributions of the folded states and unfolded states were then fitted to single exponential functions to extract the one-step unfolding and folding kinetic rates of the hairpins (**Figure 4.11B**, **Appendix Tables A2.1 and A2.3**). The force-dependent kinetics were then fitted to the Bell

model.^{51,52} Critical forces were obtained from the crossing points of the unfolding rate versus force and folding rate versus force curves (**Figure 4.12A,B**). The mechanical unfolding and folding transition state positions can be obtained from the slopes of the unfolding and folding rates versus force curves. The critical forces of the WT, +19G, +14U, and +14U+19G hairpins are 13.5 ± 0.1 , 14.1 ± 0.1 , 10.0 ± 0.2 , and 11.9 ± 0.2 pN, respectively (**Table 4.2**). The hairpins +19G (14.1 pN) and +14U (10.0 pN) have the highest and lowest critical forces, respectively. The critical force of +14U+19G (11.9 pN) is lower than WT (13.5 pN) and 19G (14.1 pN), but higher than +14U (10.0 pN), suggesting that mechanical stability of +14U+19G is primarily determined by the destabilizing +14U mutation, largely through increasing the mechanical unfolding rate (**Figure 4.12B**). Thus, mechanical unfolding can help directly capture the structural rearrangement and the resulting conformational stability change.

Table 4.2. Comparison of the critical forces, extension changes, and the unfolding free energies at the corresponding critical force and zero force of WT, +19G, +14U and +14U+19G in 100 uM spermine, 5 mM MgCl₂, 200 mM NaCl, 10 mM HEPES, 0.1 mM EDTA, pH7.3, from constant-force experiments.

RNA molecule	ΔX_c (nm)	F_c (pN) ^a	$-\Delta G_{F_c}$ (kcal·mol ⁻¹) ^b	$-\Delta G_0$ (kcal·mol ⁻¹) ^c
WT	6.5±1.4	13.5 ± 0.1	12.6 ± 2.9	7.7 ± 2.9
+19G	9.4±1.3	14.1 ± 0.1	19.1 ± 2.6	12.5 ± 2.6
+14U	6.2±1.7	10.0 ± 0.2	8.9 ± 2.5	5.1 ± 2.5
+14U+19G	8.3±1.6	11.9 ± 0.2	13.9 ± 2.7	8.6 ± 2.7
^a F_c is the critical force at which the unfolding and folding rates equal each other. ^b ΔG_{F_c} is the free energy for unfolding reaction at critical force ($\Delta G_{F_c} = N_A F_c \Delta X_c$). N_A is Avogadro constant. ^c ΔG_0 is the Gibbs free energy change in the unfolding process with no external force and can be calculated from $\Delta G_0 = \Delta G_{F_c} - W_{\text{stretching}}$. $W_{\text{stretching}}$ is the work from the stretching of single-stranded RNA from 0 force to F_c , obtained through the integration of the EWLC curve. All uncertainties reported here are standard errors.

Figure 4.12. Comparison between exon 10 inclusion level and mechanical (un)folding parameters. (A) Force-dependent unfolding (open markers) and folding (filled markers) rates of WT and +19G obtained through constant-force mode. (B) Force-dependent unfolding (open markers) and folding (filled markers) rates of WT, +19G, +14U and +14U+19G obtained through passive mode. Results of WT and +19G drawn from constant-force mode and passive mode are consistent. The R^2 values of the linear fitting curves range from 0.973 to 0.999. (C) The inverse correlation between the 4R% inclusion and the critical force. (D) 4R% inclusion vs the experimentally determined free energy changes

at zero force. (E) A positive correlation observed between the 4R% inclusion and the unfolding rate at 10 pN. (F) No correlation observed between the 4R% inclusion and the folding rate at 10 pN.

The slopes of the folding rates versus force are almost the same, suggesting that the folding transition state positions of the four hairpins are identical (between base pairs (U+2)-(A+13) and (G+1)-(C+14), **Figure 4.8B**), presumably because the hairpins have the same the loop and closing base pairs, consistent with previously reported results for other hairpins.^{37,53-55} The mutations affect more significantly the mechanical unfolding rate than the folding rate. The +19G and +14U hairpins show the slowest and fastest unfolding rates, respectively (**Figure 4.12A,B**). We did not obtain the kinetics parameters for +14U and +14U+19G due to the relatively fast (un)folding rates for the two hairpins. The critical force for +14U and +14U+19G were obtained from the (un)folding population versus force plots (**Appendix Table A2.2**). The mechanical unfolding rates of +14U+19G is faster than WT but slower than +14U (**Figure 4.12B**).

We observed that exon 10 inclusion level decreased with increasing critical force (**Figure 4.12C**). The folding free energy experimentally can be obtained from the mechanical unfolding studies (**Table 4.2**). The exon 10 inclusion level generally decreases with increasingly favourable folding free energy (**Figure 4.12D**). Significantly, we observe that the exon 10 inclusion level correlates positively with the unfolding rate in the force range of 0 to 15 pN (**Figure 4.12E**, **Figure 4.13**). An excellent correlation is observed at 10 pN (**Figure 4.12E**), which is around the mechanical critical force, consistent with the fact that the

unfolding of the splice site hairpins is aided by helicases. A weak correlation was observed between exon 10 inclusion level and mechanical folding rate (**Figure 4.12F, Figure 4.13**). Taken together, our single-molecule studies show that mechanical unfolding can mimic how the splice site hairpins are unfolded in the cell for the recognition by U1 snRNA.

The unfolding rate versus force curve of +19G has a steeper slope compared to those of the other three hairpins (**Figure 4.12A,B**), suggesting that, more base pairs of +19G need to be unwound to reach the unfolding transition state position, consistent with the fact that +19G has an extended and stabilized bottom stem (**Figure 4.8B**). The unfolding transition state of +14U+19G is positioned lower than that of +19G, probably because a +14U mutation destabilizes the top stem as well as the bottom stem, resulting in the lowering of the unfolding barrier (+14U+19G-type 1, **Figure 4.8B**). It is also possible that +14U+19G may occasionally fold into an alternative conformation similar to WT and +14U (+14U+19G-type 2, **Figure 4.8B**). The number of folded nucleotides (as calculated based on the extension changes) of +14U+19G (25 nt) is in between WT and +14U (21 nt) and +19G (27 nt), consistent with the hairpin +14U+19G folding into two conformations, which are not experimentally resolved.

Figure 4.13. 4R% inclusion vs the unfolding kinetics and folding kinetics at 0, 10, 15, and 20 pN. (A–D) 4R% virus $\ln k_{\text{unfolding}}$ at 0 pN, 10 pN, 15 pN and 20 pN. The best correlation at a stretching force of around 10 pN is revealed. (E–H) 4R% virus $\ln k_{\text{folding}}$ at 0 pN, 10 pN, 15 pN and 20 pN. No correlation is revealed.

4.4 Conclusions

In summary, we have shown that the disease-causing +19G mutation in the tau pre-mRNA exon 10 5'ss hairpin causes a significant structural rearrangement in the bottom stem. This change in hairpin structure results in enhanced hairpin stability, which leads to decreased exon 10 inclusion. The structural rearrangement model explains the apparently aberrant phenotype of the +19G mutation, which was initially thought to weaken the hairpin structure by disrupting a single base pair in the bottom stem of the 25-mer splice site hairpin.¹² The structural rearrangement model takes an extended 30-mer hairpin sequence into consideration, instead of a 25-mer sequence typically employed by the majority of prior studies.^{15,31,39,56} The structural change is evidenced by our biophysical characterizations including oligonucleotide thermal melting experiments and single-molecule mechanical measurements. Local secondary structures and potential structural rearrangements may be generally utilized for regulating other splicing events in cells.^{17,57-65}

Clearly, a single mutation could either destabilize or stabilize the splice site hairpin structure^{3,16,18,66,67} leading to increased or decreased inclusion of tau exon 10, thus disturbing the 4R:3R tau protein isoform ratio and causing FTDP-17. By designing minigenes containing the stabilizing +19G mutation on the bottom stem and other disease-causing destabilizing mutations on the top stem, we determined that the structured region within 3 base pairs of the splice site (ranging from positions -2 and +3 and +12 to +15) is critical for splicing regulation. Destabilizing mutations in this region in top stem can not be compensated by stabilizing mutations in the bottom such as +19G and -4G.

Numerous *trans*-acting factors (other than the spliceosomal components) are present in the cell, and may play important roles in determining *MAPT* exon 10 splicing. Mutations in pre-mRNA may affect *trans*-factor binding and thus exon 10 splicing. For example, it has been reported that mutations at the +19 position (+19A or +19U) also result in a small but non-significant decrease in exon inclusion, even though they do not seem to increase hairpin stability,¹² indicating that mutations may affect the binding of *trans*-acting protein factors including DDX5 and SFPQ.^{4,22-27,29-31,68}

Protein *trans* factors such as helicases may help unwind the hairpin structure to allow the U1 snRNA binding to the splice site as U1 snRNA alone may not be able to invade the hairpin structure. It is interesting to note that mechanical unfolding rate at around 10 pN has an excellent correlation with splicing activity, consistent with the fact that the splice site hairpins need to be unwound by helicases to allow the 5'ss recognition by U1 snRNA. Mechanical unfolding studies of RNA structures may be generally useful in providing insights into RNA structure-mediated processes. A similar correlation has been observed between the ribosomal frameshifting efficiency and the mechanical unfolding rate within a force range of the frameshift-stimulating mRNA pseudoknot, suggesting the possible active role of ribosome helicase activity in unwinding RNA structures.³⁶ Clearly, single-molecule mechanical unfolding studies can reveal directly the structural changes upon single-nucleotide mutations and can mimic how the splice site hairpins are unfolded in the cell for the recognition by U1 snRNA.

References

- (1) Z. K. Wszolek, Y. Tsuboi, B. Ghetti, S. Pickering-Brown, Y. Baba, W. P. Cheshire, Frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP-17). *Orphanet J Rare Dis* **2006**, *1*, 30-30.
- (2) C. Li, J. Gotz, Tau-based therapies in neurodegeneration: opportunities and challenges. *Nat Rev Drug Discov* **2017**, *16*, 863-883.
- (3) M. Hutton, C. L. Lendon, P. Rizzu, M. Baker, S. Froelich, H. Houlden, S. Pickering-Brown, S. Chakraverty, A. Isaacs, A. Grover, J. Hackett, J. Adamson, S. Lincoln, D. Dickson, P. Davies, R. C. Petersen, M. Stevens, E. de Graaff, E. Wauters, J. van Baren, M. Hillebrand, M. Joosse, J. M. Kwon, P. Nowotny, L. K. Che, J. Norton, J. C. Morris, L. A. Reed, J. Trojanowski, H. Basun, L. Lannfelt, M. Neystat, S. Fahn, F. Dark, T. Tannenberg, P. R. Dodd, N. Hayward, J. B. Kwok, P. R. Schofield, A. Andreadis, J. Snowden, D. Craufurd, D. Neary, F. Owen, B. A. Oostra, J. Hardy, A. Goate, J. van Swieten, D. Mann, T. Lynch, P. Heutink, Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature* **1998**, *393*, 702-705.
- (4) I. D'Souza, P. Poorkaj, M. Hong, D. Nochlin, V. M. Lee, T. D. Bird, G. D. Schellenberg, Missense and silent tau gene mutations cause frontotemporal dementia with parkinsonism-chromosome 17 type, by affecting multiple alternative RNA splicing regulatory elements. *Proc Natl Acad Sci U S A* **1999**, *96*, 5598-5603.
- (5) M. G. Spillantini, J. R. Murrell, M. Goedert, M. R. Farlow, A. Klug, B. Ghetti, Mutation in the tau gene in familial multiple system tauopathy with presenile dementia. *Proc Natl Acad Sci U S A* **1998**, *95*, 7737-7741.
- (6) B. Ghetti, A. L. Oblak, B. F. Boeve, K. A. Johnson, B. C. Dickerson, M. Goedert, Invited review: Frontotemporal dementia caused by microtubule-associated protein tau gene (MAPT) mutations: a chameleon for neuropathology and neuroimaging. *Neuropathol Appl Neurobiol* **2015**, *41*, 24-46.
- (7) A. McCarthy, R. Loneran, D. A. Olszewska, S. O'Dowd, G. Cummins, B. Magennis, E. M. Fallon, N. Pender, E. D. Huey, S. Cosentino, K. O'Rourke, B. D. Kelly, M. O'Connell, I. Delon, M. Farrell, M. G. Spillantini, L. P. Rowland, S. Fahn, P. Craig, M. Hutton, T. Lynch, Closing the tau loop: the missing tau mutation. *Brain* **2015**, *138*, 3100-3109.
- (8) I. Grundke-Iqbal, K. Iqbal, Y. C. Tung, M. Quinlan, H. M. Wisniewski, L. I. Binder, Abnormal phosphorylation of the microtubule-associated protein tau (tau) in Alzheimer cytoskeletal pathology. *Proc Natl Acad Sci U S A* **1986**, *83*, 4913-4917.
- (9) A. W. P. Fitzpatrick, B. Falcon, S. He, A. G. Murzin, G. Murshudov, H. J. Garringer, R. A. Crowther, B. Ghetti, M. Goedert, S. H. W. Scheres, Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* **2017**, *547*, 185-190.
- (10) B. Falcon, W. Zhang, A. G. Murzin, G. Murshudov, H. J. Garringer, R. Vidal, R. A. Crowther, B. Ghetti, S. H. W. Scheres, M. Goedert, Structures of filaments from Pick's disease reveal a novel tau protein fold. *Nature* **2018**, *561*, 137-140.
- (11) J. J. Young, M. Lavakumar, D. Tampi, S. Balachandran, R. R. Tampi, Frontotemporal dementia: latest evidence and clinical implications. *Ther Adv Psychopharmacol* **2018**, *8*, 33-48.

- (12) P. M. Stanford, C. E. Shepherd, G. M. Halliday, W. S. Brooks, P. W. Schofield, H. Brodaty, R. N. Martins, J. B. J. Kwok, P. R. Schofield, Mutations in the tau gene that cause an increase in three repeat tau and frontotemporal dementia. *Brain* **2003**, *126*, 814-826.
- (13) K. M. Schoch, S. L. DeVos, R. L. Miller, S. J. Chun, M. Norrbom, D. F. Wozniak, H. N. Dawson, C. F. Bennett, F. Rigo, T. M. Miller, Increased 4R-Tau Induces Pathological Changes in a Human-Tau Mouse Model. *Neuron* **2016**, *90*, 941-947.
- (14) M. Hong, V. Zhukareva, V. Vogelsberg-Ragaglia, Z. Wszolek, L. Reed, B. I. Miller, D. H. Geschwind, T. D. Bird, D. McKeel, A. Goate, J. C. Morris, K. C. Wilhelmsen, G. D. Schellenberg, J. Q. Trojanowski, V. M. Lee, Mutation-specific functional impairments in distinct tau isoforms of hereditary FTDP-17. *Science* **1998**, *282*, 1914-1917.
- (15) C. P. Donahue, C. Muratore, J. Y. Wu, K. S. Kosik, M. S. Wolfe, Stabilization of the tau exon 10 stem loop alters pre-mRNA splicing. *J Biol Chem* **2006**, *281*, 23302-23306.
- (16) L. Varani, M. Hasegawa, M. G. Spillantini, M. J. Smith, J. R. Murrell, B. Ghetti, A. Klug, M. Goedert, G. Varani, Structure of tau exon 10 splicing regulatory element RNA and destabilization by mutations of frontotemporal dementia and parkinsonism linked to chromosome 17. *Proc Natl Acad Sci U S A* **1999**, *96*, 8229-8234.
- (17) E. Buratti, F. E. Baralle, Influence of RNA secondary structure on the pre-mRNA splicing process. *Mol Cell Biol* **2004**, *24*, 10505-10514.
- (18) A. Grover, H. Houlden, M. Baker, J. Adamson, J. Lewis, G. Prihar, S. Pickering-Brown, K. Duff, M. Hutton, 5' splice site mutations in tau associated with the inherited dementia FTDP-17 affect a stem-loop structure that regulates alternative splicing of exon 10. *J Biol Chem* **1999**, *274*, 15134-15143.
- (19) X. Roca, A. R. Krainer, I. C. Eperon, Pick one, but be quick: 5' splice sites and the problems of too many choices. *Genes Dev* **2013**, *27*, 129-144.
- (20) X. Roca, M. Akerman, H. Gaus, A. Berdeja, C. F. Bennett, A. R. Krainer, Widespread recognition of 5' splice sites by noncanonical base-pairing to U1 snRNA involving bulged nucleotides. *Genes Dev* **2012**, *26*, 1098-1109.
- (21) J. Tan, Jia Xin J. Ho, Z. Zhong, S. Luo, G. Chen, X. Roca, Noncanonical registers and base pairs in human 5' splice-site selection. *Nucleic Acids Res* **2016**, *44*, 3908-3921.
- (22) A. Kar, K. Fushimi, X. Zhou, P. Ray, C. Shi, X. Chen, Z. Liu, S. Chen, J. Y. Wu, RNA helicase p68 (DDX5) regulates tau exon 10 splicing by modulating a stem-loop structure at the 5' splice site. *Mol Cell Biol* **2011**, *31*, 1812-1821.
- (23) C. Lin, L. Yang, J. J. Yang, Y. Huang, Z. R. Liu, ATPase/helicase activities of p68 RNA helicase are required for pre-mRNA splicing but not for assembly of the spliceosome. *Mol Cell Biol* **2005**, *25*, 7484-7493.
- (24) S. A. Park, S. I. Ahn, J. M. Gallo, Tau mis-splicing in the pathogenesis of neurodegenerative disorders. *BMB Rep* **2016**, *49*, 405-413.
- (25) P. Ray, A. Kar, K. Fushimi, N. Havlioglu, X. Chen, J. Y. Wu, PSF Suppresses Tau Exon 10 Inclusion by Interacting with a Stem-Loop Structure Downstream of Exon 10. *J Mol Neurosci* **2011**, *45*, 453-466.
- (26) S. Ishigaki, Y. Fujioka, Y. Okada, Y. Riku, T. Udagawa, D. Honda, S. Yokoi, K. Endo, K. Ikenaka, S. Takagi, Y. Iguchi, N. Sahara, A. Takashima, H. Okano, M. Yoshida, H. Warita, M. Aoki, H. Watanabe, H. Okado, M. Katsuno, G. Sobue, Altered Tau Isoform Ratio Caused by Loss of FUS and SFPQ Function Leads to FTLD-like Phenotypes. *Cell Rep* **2017**, *18*, 1118-1131.

- (27) Y. D. Ke, J. Dramiga, U. Schutz, J. J. Kril, L. M. Ittner, H. Schroder, J. Gotz, Tau-mediated nuclear depletion and cytoplasmic accumulation of SFPQ in Alzheimer's and Pick's disease. *PloS one* **2012**, *7*, e35678.
- (28) I. D'Souza, G. D. Schellenberg, Determinants of 4-repeat tau expression. Coordination between enhancing and inhibitory splicing sequences for exon 10 inclusion. *J Biol Chem* **2000**, *275*, 17700-17709.
- (29) I. D'Souza, G. D. Schellenberg, tau Exon 10 expression involves a bipartite intron 10 regulatory sequence and weak 5' and 3' splice sites. *J Biol Chem* **2002**, *277*, 26587-26599.
- (30) J. Zhou, Q. Yu, T. Zou, Alternative splicing of exon 10 in the tau gene as a target for treatment of tauopathies. *BMC Neurosci* **2008**, *9 Suppl 2*, S10.
- (31) J. Lisowiec, D. Magner, E. Kierzek, E. Lenartowicz, R. Kierzek, Structural determinants for alternative splicing regulation of the MAPT pre-mRNA. *RNA Biol* **2015**, *12*, 330-342.
- (32) J. S. Reuter, D. H. Mathews, RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics* **2010**, *11*, 129-129.
- (33) D. H. Mathews, M. D. Disney, J. L. Childs, S. J. Schroeder, M. Zuker, D. H. Turner, Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc Natl Acad Sci U S A* **2004**, *101*, 7287-7292.
- (34) J. Liphardt, B. Onoa, S. B. Smith, I. Tinoco, Jr., C. Bustamante, Reversible unfolding of single RNA molecules by mechanical force. *Science* **2001**, *292*, 733-737.
- (35) S. B. Smith, Y. Cui, C. Bustamante, Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science* **1996**, *271*, 795-799.
- (36) Z. Zhong, L. Yang, H. Zhang, J. Shi, J. J. Vandana, D. T. Lam, R. C. Olsthoorn, L. Lu, G. Chen, Mechanical unfolding kinetics of the SRV-1 gag-pro mRNA pseudoknot: possible implications for -1 ribosomal frameshifting stimulation. *Sci Rep* **2016**, *6*, 39549.
- (37) L. X. Yang, Z. S. Zhong, C. L. Tong, H. Jia, Y. R. Liu, G. Chen, Single-Molecule Mechanical Folding and Unfolding of RNA Hairpins: Effects of Single A-U to A.C Pair Substitutions and Single Proton Binding and Implications for mRNA Structure-Induced -1 Ribosomal Frameshifting. *J Am Chem Soc* **2018**, *140*, 8172-8184.
- (38) Y. Luo, M. D. Disney, Bottom-up design of small molecules that stimulate exon 10 skipping in mutant MAPT pre-mRNA. *Chembiochem* **2014**, *15*, 2041-2044.
- (39) C. P. Donahue, J. Ni, E. Rozners, M. A. Glicksman, M. S. Wolfe, Identification of tau stem loop RNA stabilizers. *J Biomol Screen* **2007**, *12*, 789-799.
- (40) Y. Liu, E. Peacey, J. Dickson, C. P. Donahue, S. Zheng, G. Varani, M. S. Wolfe, Mitoxantrone analogues as ligands for a stem-loop structure of tau pre-mRNA. *J Med Chem* **2009**, *52*, 6523-6526.
- (41) S. Zheng, Y. Chen, C. P. Donahue, M. S. Wolfe, G. Varani, Structural basis for stabilization of the tau pre-mRNA splicing regulatory element by novantrone (mitoxantrone). *Chem Biol* **2009**, *16*, 557-566.
- (42) G. Artigas, P. Lopez-Senin, C. Gonzalez, N. Escaja, V. Marchan, Ametantrone-based compounds as potential regulators of Tau pre-mRNA alternative splicing. *Org Biomol Chem* **2015**, *13*, 452-464.
- (43) J. Wang, Q.-S. Gao, Y. Wang, R. Lafyatis, S. Stamm, A. Andreadis, Tau exon 10, whose missplicing causes frontotemporal dementia, is regulated by an

- intricate interplay of cis elements and trans factors. *J Neurochem* **2004**, *88*, 1078-1090.
- (44) R. Sud, E. T. Geller, G. D. Schellenberg, Antisense-mediated Exon Skipping Decreases Tau Protein Expression: A Potential Therapy For Tauopathies. *Mol Ther Nucleic Acids* **2014**, *3*, e180.
 - (45) J. Ubbink, T. Odijk, Polymer- and salt-induced toroids of hexagonal DNA. *Biophys J* **1995**, *68*, 54-61.
 - (46) D. Bensimon, A. J. Simon, V. V. Croquette, A. Bensimon, Stretching DNA with a receding meniscus: Experiments and models. *Phys Rev Lett* **1995**, *74*, 4754-4757.
 - (47) F. Vanzi, Y. Takagi, H. Shuman, B. S. Cooperman, Y. E. Goldman, Mechanical studies of single ribosome/mRNA complexes. *Biophys J* **2005**, *89*, 1909-1919.
 - (48) C. Bustamante, J. F. Marko, E. D. Siggia, S. Smith, Entropic elasticity of lambda-phage DNA. *Science* **1994**, *265*, 1599-1600.
 - (49) M. D. Wang, H. Yin, R. Landick, J. Gelles, S. M. Block, Stretching DNA with optical tweezers. *Biophys J* **1997**, *72*, 1335-1346.
 - (50) S. A. McKinney, C. Joo, T. Ha, Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophys J* **2006**, *91*, 1941-1951.
 - (51) J. D. Wen, M. Manosas, P. T. Li, S. B. Smith, C. Bustamante, F. Ritort, I. Tinoco, Jr., Force unfolding kinetics of RNA using optical tweezers. I. Effects of experimental variables on measured results. *Biophys J* **2007**, *92*, 2996-3009.
 - (52) I. Tinoco, C. Bustamante, The effect of force on thermodynamics and kinetics of single molecule reactions. *Biophys Chem* **2002**, *101*, 513-533.
 - (53) M. T. Woodside, P. C. Anthony, W. M. Behnke-Parks, K. Larizadeh, D. Herschlag, S. M. Block, Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. *Science* **2006**, *314*, 1001-1004.
 - (54) M. T. Woodside, W. M. Behnke-Parks, K. Larizadeh, K. Travers, D. Herschlag, S. M. Block, Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins. *Proc Natl Acad Sci U S A* **2006**, *103*, 6190-6195.
 - (55) Z. Zhong, L. H. Soh, M. H. Lim, G. Chen, A U·U Pair-to-U·C Pair Mutation-Induced RNA Native Structure Destabilisation and Stretching-Force-Induced RNA Misfolding. *ChemPlusChem* **2015**, *80*, 1267-1278.
 - (56) L. Varani, M. G. Spillantini, M. Goedert, G. Varani, Structural basis for recognition of the RNA major groove in the tau exon 10 splicing regulatory element by aminoglycoside antibiotics. *Nucleic Acids Res* **2000**, *28*, 710-719.
 - (57) E. W. J. Wallace, J. D. Beggs, Extremely fast and incredibly close: cotranscriptional splicing in budding yeast. *RNA* **2017**, *23*, 601-610.
 - (58) E. C. Merkhofer, P. Hu, T. L. Johnson, Introduction to Cotranscriptional RNA Splicing. *Methods Mol Biol* **2014**, *1126*, 83-96.
 - (59) C. K. Damgaard, T. O. Tange, J. Kjems, hnRNP A1 controls HIV-1 mRNA splicing through cooperative binding to intron and exon splicing silencers in the context of a conserved secondary structure. *RNA* **2002**, *8*, 1401-1415.
 - (60) N. R. Zearfoss, E. S. Johnson, S. P. Ryder, hnRNP A1 and secondary structure coordinate alternative splicing of Mag. *RNA* **2013**, *19*, 948-957.
 - (61) N. N. Singh, R. N. Singh, E. J. Androphy, Modulating role of RNA structure in alternative splicing of a critical exon in the spinal muscular atrophy genes. *Nucleic Acids Res* **2007**, *35*, 371-389.
 - (62) J. Wang, P. G. Schultz, K. A. Johnson, Mechanistic studies of a small-molecule modulator of SMN2 splicing. *Proc Natl Acad Sci U S A* **2018**, *115*, E4604-E4612.

- (63) J. Palacino, S. E. Swalley, C. Song, A. K. Cheung, L. Shu, X. Zhang, M. Van Hoosear, Y. Shin, D. N. Chin, C. G. Keller, M. Beibel, N. A. Renaud, T. M. Smith, M. Salcius, X. Shi, M. Hild, R. Servais, M. Jain, L. Deng, C. Bullock, M. McLellan, S. Schuierer, L. Murphy, M. J. Blommers, C. Blaustein, F. Berenshteyn, A. Lacoste, J. R. Thomas, G. Roma, G. A. Michaud, B. S. Tseng, J. A. Porter, V. E. Myer, J. A. Tallarico, L. G. Hamann, D. Curtis, M. C. Fishman, W. F. Dietrich, N. A. Dales, R. Sivasankaran, SMN2 splice modulators enhance U1-pre-mRNA association and rescue SMA mice. *Nat Chem Biol* **2015**, *11*, 511-517.
- (64) M. Sivaramakrishnan, K. D. McCarthy, S. Campagne, S. Huber, S. Meier, A. Augustin, T. Heckel, H. Meistermann, M. N. Hug, P. Birrer, A. Moursy, S. Khawaja, R. Schmucki, N. Berntenis, N. Giroud, S. Golling, M. Tzouros, B. Banfai, G. Duran-Pacheco, J. Lamerz, Y. Hsiu Liu, T. Luebbbers, H. Ratni, M. Ebeling, A. Clery, S. Paushkin, A. R. Krainer, F. H. Allain, F. Metzger, Binding to SMN2 pre-mRNA-protein complex elicits specificity for small molecule splicing modifiers. *Nat Commun* **2017**, *8*, 1476.
- (65) A. Garcia-Lopez, F. Tessaro, H. R. A. Jonker, A. Wacker, C. Richter, A. Comte, N. Berntenis, R. Schmucki, K. Hatje, O. Petermann, G. Chiriano, R. Perozzo, D. Sciarra, P. Konieczny, I. Faustino, G. Fournet, M. Orozco, R. Artero, F. Metzger, M. Ebeling, P. Goekjian, B. Joseph, H. Schwalbe, L. Scapozza, Targeting RNA structure in SMN2 reverses spinal muscular atrophy molecular phenotypes. *Nat Commun* **2018**, *9*, 2032.
- (66) C. Dumanchin, A. Camuzat, D. Campion, P. Verpillat, D. Hannequin, B. Dubois, P. Saugier-Veber, C. Martin, C. Penet, F. Charbonnier, Y. Agid, T. Frebourg, A. Brice, Segregation of a missense mutation in the microtubule-associated protein tau gene with familial frontotemporal dementia and parkinsonism. *Hum Mol Genet* **1998**, *7*, 1825-1829.
- (67) M. G. Spillantini, M. Goedert, Tau protein pathology in neurodegenerative diseases. *Trends Neurosci* **1998**, *21*, 428-433.
- (68) N. A. Faustino, T. A. Cooper, Pre-mRNA splicing and human disease. *Genes Dev* **2003**, *17*, 419-437.

Chapter 5

Conclusion and Future Work

Force is involved in a wide range of important biological activities such as cargo particles delivery by motor proteins, RNA polymerase moving along the DNA template during transcription,^{1,2} ribosome unfolding mRNA structure into single strand during ribosomal translation,³⁻⁶ and pre-mRNA splicing by spliceosome.^{7,8} Optical tweezers employ gradient force generated by highly focused laser beams with Gaussian distribution to form a stable potential well or optical trap so that it can steadily capture and manipulate dielectric particles and apply external forces on the single molecules of interest attached to the particles.^{9,10} With the help of this method, direct measurement of conformation and mechanochemical transformations under mechanical force of single molecule is achieved. In recent decades, optical tweezers have appeared to be a very useful and practical tool to achieve single-molecule manipulation and probe the mechanical properties of RNA structures.¹¹⁻¹⁸ RNAs are important in regulating gene expression and cellular functions such as serving as catalysis, generating multiple functional proteins through ribosomal frameshifting during translation, regulating pre-mRNA alternative splicing, etc. The RNA structures, stabilities and kinetics need to be investigated to provide better understanding of the related biological functions.

In summary, in this thesis, we employed single-molecule manipulation using high-resolution optical tweezers, bulk thermal melting experiments, *in vitro*

frameshifting assays, cell culture splicing assays, non-denaturing PAGE studies, and PNA binding to study the structures, thermal and mechanical stabilities, and unfolding and folding kinetics of RNA structures, and how they regulate the RNA structure-induced minus-one ribosomal frameshifting and pre-mRNA alternative splicing.

The studies have clearly revealed that the mechanical stability as well as the kinetics of the RNA structures play vital roles in stimulating minus-one ribosomal frameshifting and regulating *MAPT* exon 10 splicing. We found that the higher mechanical stability of the mRNA structure at the frameshift site can provide more resistance against the unwinding of the elongating ribosome, thus stimulating minus-one ribosomal frameshifting. We observed an excellent correlation between exon 10 inclusion level and mechanical unfolding rate at 10 pN, consistent with the fact that the unfolding of the splice site hairpins is aided by helicases. The mechanical unfolding directly reveals the structural changes and mechanical properties, and can mimic the ribosome activity during frameshift and how the splice site hairpins are unfolded in the cell for the recognition by U1 snRNA. The single-molecule studies using optical tweezers are very informative in unveiling the structural information and mechanical properties of RNA. Thus, the mechanical manipulation using high-resolution optical tweezers serves as a very good tool to study the mechanical properties of RNA and the possible implications towards RNA-regulated biological activities.

We also employed PNAs to interact with target RNA sequences and structures with high selectivity and binding affinity. PNA binding has been extensively used

in RNA recognitions, diagnostics, and anti-viral treatments. It is shown by our study that PNA can specifically bind to target RNA, reveal the relative stability of the target native and mutated structures. PNAs could be used to probe the complex ribosomal translation processes and uncover the roles of *cis*-acting mRNA structures. The RNA-binding PNAs can be applied in future to target viral RNA structures to modulate the -1 programmed ribosomal frameshifting *in vivo*.

In future, it would be much valuable to combine the single-molecule mechanical study using optical tweezers with PNA binding studies. We can flow in PNA to the sample chamber of the optical tweezers machine, and directly study the interaction of the PNA with target RNA structure at the single-molecule level. Such single-molecule RNA and PNA interactions can potentially reveal the detailed effects of ligand binding, the ligand binding modes, and the binding kinetics, etc. We can also directly monitor and characterize the helicase activity in unwinding RNA structure using optical tweezers. In combination with bulk experiment, we could gain a deeper understanding of the detailed mechanisms of RNA-based biological activities.

References

- (1) E. A. Abbondanzieri, W. J. Greenleaf, J. W. Shaevitz, R. Landick, S. M. Block, Direct observation of base-pair stepping by RNA polymerase. *Nature* **2005**, *438*, 460-465.
- (2) M. H. Larson, R. Landick, S. M. Block, Single-molecule studies of RNA polymerase: one singular sensation, every little step it takes. *Mol Cell* **2011**, *41*, 249-262.
- (3) G. Z. Yusupova, M. M. Yusupov, J. H. Cate, H. F. Noller, The path of messenger RNA through the ribosome. *Cell* **2001**, *106*, 233-241.
- (4) S. Takyar, R. P. Hickerson, H. F. Noller, mRNA helicase activity of the ribosome. *Cell* **2005**, *120*, 49-58.
- (5) T. Liu, A. Kaplan, L. Alexander, S. Yan, J. D. Wen, L. Lancaster, C. E. Wickersham, K. Fredrick, H. Noller, I. Tinoco, C. J. Bustamante, Direct measurement of the mechanical work during translocation by the ribosome. *Elife* **2014**, *3*, e03406.
- (6) J. D. Wen, L. Lancaster, C. Hodges, A. C. Zeri, S. H. Yoshimura, H. F. Noller, C. Bustamante, I. Tinoco, Following translation by single ribosomes one codon at a time. *Nature* **2008**, *452*, 598-603.
- (7) N. A. Faustino, T. A. Cooper, Pre-mRNA splicing and human disease. *Genes Dev* **2003**, *17*, 419-437.
- (8) J. P. Staley, C. Guthrie, Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **1998**, *92*, 315-326.
- (9) A. Ashkin, J. M. Dziedzic, J. E. Bjorkholm, S. Chu, Observation of a single-beam gradient force optical trap for dielectric particles. *Opt Lett* **1986**, *11*, 288.
- (10) A. Ashkin, Acceleration and Trapping of Particles by Radiation Pressure. *Phys Rev Lett* **1970**, *24*, 156-8.
- (11) J. Liphardt, B. Onoa, S. B. Smith, I. Tinoco, Jr., C. Bustamante, Reversible unfolding of single RNA molecules by mechanical force. *Science* **2001**, *292*, 733-737.
- (12) G. Chen, J. D. Wen, I. Tinoco, Jr., Single-molecule mechanical unfolding and folding of a pseudoknot in human telomerase RNA. *RNA* **2007**, *13*, 2175-2188.
- (13) M. T. Woodside, C. Garcia-Garcia, S. M. Block, Folding and unfolding single RNA molecules under tension. *Curr Opin Chem Biol* **2008**, *12*, 640-646.
- (14) G. Chen, K. Y. Chang, M. Y. Chou, C. Bustamante, I. Tinoco, Triplex structures in an RNA pseudoknot enhance mechanical stability and increase efficiency of -1 ribosomal frameshifting. *Proc Natl Acad Sci U S A* **2009**, *106*, 12706-12711.
- (15) Z. Zhong, L. Yang, H. Zhang, J. Shi, J. J. Vandana, D. T. Lam, R. C. Olsthoorn, L. Lu, G. Chen, Mechanical unfolding kinetics of the SRV-1 gag-pro mRNA pseudoknot: possible implications for -1 ribosomal frameshifting stimulation. *Sci Rep* **2016**, *6*, 39549.
- (16) J. T. Finer, R. M. Simmons, J. A. Spudich, Single myosin molecule mechanics: piconewton forces and nanometre steps. *Nature* **1994**, *368*, 113-119.
- (17) T. J. Purcell, C. Morris, J. A. Spudich, H. L. Sweeney, Role of the lever arm in the processive stepping of myosin V. *Proc Natl Acad Sci U S A* **2002**, *99*, 14159-14164.
- (18) L. X. Yang, Z. S. Zhong, C. L. Tong, H. Jia, Y. R. Liu, G. Chen, Single-Molecule Mechanical Folding and Unfolding of RNA Hairpins: Effects of Single A-U to A.C Pair Substitutions and Single Proton Binding and Implications for mRNA Structure-Induced -1 Ribosomal Frameshifting. *J Am Chem Soc* **2018**, *140*, 8172-8184.

Appendixes

Appendix 1. The parameters obtained from mechanical pulling experiments for rHP, U22C, U25C, and U29C at varied pH.

Table A1.1. Parameters obtained from constant-force experiments at 200 mM NaCl with varied pH.

pH	Force (pN)	Measured ΔX (nm)	Measured nt ^a	τ_{folded} (s) ^b	τ_{unfolded} (s)	F_c (pN)	$\Delta X^{\ddagger, \text{FU}}$ (nm)	$\Delta n^{\ddagger, \text{FU}}$	$\Delta X^{\ddagger, \text{UF}}$ (nm)	$\Delta n^{\ddagger, \text{FU}}$
rHP										
5.0	16.3	12.4±1.1	32.8±2.5	0.645±0.023	0.263±0.008	16.7±0.1	7.7±0.6	17±1	4.0±0.6	14±1
	16.5	12.4±1.1	32.7±2.5	0.467±0.015	0.317±0.016					
	16.7	12.5±1.4	32.8±3.2	0.371±0.014	0.331±0.011					
	16.9	12.4±1.3	32.5±2.9	0.216±0.016	0.497±0.019					
	17.1	12.3±1.2	32.2±2.7	0.144±0.007	0.565±0.028					
	$\ln(k_{\text{unfolding}}) = (1.88 \pm 0.15) F - (30.35 \pm 2.49)$, $R^2 = 0.991$ $\ln(k_{\text{folding}}) = (-0.99 \pm 0.14) F + (17.50 \pm 2.41)$, $R^2 = 0.970$									
6.5	16.7	12.9±1.1	33.8±2.5	0.733±0.058	0.171±0.005	17.1±0.1	7.3±0.4	16±1	5.7±1.0	17±2
	16.9	13.0±1.1	33.9±2.5	0.450±0.021	0.209±0.004					
	17.1	12.8±1.1	33.4±2.5	0.342±0.010	0.226±0.004					
	17.3	12.9±1.1	33.5±2.5	0.237±0.003	0.421±0.008					
	17.5	12.8±1.1	33.2±2.5	0.169±0.005	0.494±0.011					
	$\ln(k_{\text{unfolding}}) = (1.79 \pm 0.09) F - (29.49 \pm 1.49)$, $R^2 = 0.996$ $\ln(k_{\text{folding}}) = (-1.41 \pm 0.24) F + (25.41 \pm 4.03)$, $R^2 = 0.961$									
7.3	16.1	12.7±1.1	33.6±2.5	0.658±0.020	0.192±0.010	16.5±0.1	8.5±0.4	19±1	5.0±0.7	16±2
	16.4	12.7±1.1	33.4±2.5	0.394±0.010	0.272±0.007					
	16.7	12.7±1.1	33.3±2.5	0.192±0.006	0.341±0.004					
	17.0	12.6±1.1	33.0±2.5	0.102±0.003	0.610±0.027					
	$\ln(k_{\text{unfolding}}) = (2.10 \pm 0.09) F - (33.50 \pm 1.55)$, $R^2 = 0.998$ $\ln(k_{\text{folding}}) = (-1.23 \pm 0.17) F + (21.51 \pm 2.76)$, $R^2 = 0.982$									
8.0	16.4	12.9±1.2	33.9±2.7	0.707±0.022	0.218±0.010	16.8±0.1	8.8±0.7	20±2	3.7±0.8	13±2
	16.7	12.7±1.1	33.3±2.5	0.303±0.009	0.239±0.009					
	17.0	12.7±1.1	33.2±2.5	0.175±0.004	0.312±0.008					
	17.3	12.6±1.1	32.8±2.5	0.098±0.003	0.498±0.014					
	$\ln(k_{\text{unfolding}}) = (2.16 \pm 0.16) F - (34.98 \pm 2.71)$, $R^2 = 0.994$ $\ln(k_{\text{folding}}) = (-0.91 \pm 0.20) F + (16.62 \pm 3.34)$, $R^2 = 0.956$									
U22C										
5.0 Tether1	14.9	12.5±1.2	33.7±2.8	0.694±0.034	0.178±0.006	15.4±0.1	6.2±0.2	14±0	4.9±0.4	16±1
	15.2	12.6±1.2	33.7±2.8	0.470±0.011	0.288±0.005					
	15.5	12.5±1.2	33.4±2.8	0.298±0.004	0.439±0.015					
	15.8	12.4±1.1	33.0±2.5	0.184±0.006	0.589±0.014					
	16.1	12.4±1.1	32.9±2.5	0.114±0.006	0.759±0.023					
	$\ln(k_{\text{unfolding}}) = (1.52 \pm 0.04) F - (22.27 \pm 0.56)$, $R^2 = 0.999$ $\ln(k_{\text{folding}}) = (-1.21 \pm 0.10) F + (19.60 \pm 1.48)$, $R^2 = 0.991$									
5.0 Tether2	15.0	12.6±1.2	33.9±2.8	0.677±0.014	0.234±0.003	15.4±0.1	7.1±0.1	16±0	4.6±0.4	15±1
	15.3	12.2±1.2	32.8±2.8	0.423±0.020	0.275±0.012					
	15.6	12.8±1.2	34.0±2.8	0.239±0.005	0.466±0.009					
	15.9	12.6±1.2	33.4±2.7	0.148±0.004	0.653±0.021					
	16.2	12.3±1.2	32.6±2.7	0.083±0.005	0.820±0.020					
	$\ln(k_{\text{unfolding}}) = (1.75 \pm 0.03) F - (25.87 \pm 0.54)$, $R^2 = 0.999$ $\ln(k_{\text{folding}}) = (-1.12 \pm 0.10) F + (18.36 \pm 1.61)$, $R^2 = 0.988$									

Tether1																
&2	$\ln(k_{\text{unfolding}}) = (1.63 \pm 0.05) F - (24.10 \pm 0.72), R^2 = 0.997$															
pooled ^d	$\ln(k_{\text{folding}}) = (-1.16 \pm 0.06) F + (18.94 \pm 0.98), R^2 = 0.989$				15.4±0.1	6.6±0.2	15±0	4.7±0.2	15±0							
6.5	14.9	12.7±1.1	34.1±2.6	1.212±0.058	0.464±0.010	15.2±0.1	6.7±1.4	15±3	4.7±0.9	15±2						
	15.1	12.7±1.0	34.0±2.3	0.897±0.032	0.661±0.039											
	15.3	12.4±1.0	33.2±2.3	0.510±0.042	0.862±0.029											
	15.5	12.8±1.0	34.1±2.3	0.489±0.028	0.928±0.034											
	$\ln(k_{\text{unfolding}}) = (1.64 \pm 0.34) F - (24.66 \pm 5.25), R^2 = 0.959$															
Tether1	$\ln(k_{\text{folding}}) = (-1.17 \pm 0.22) F + (18.17 \pm 3.41), R^2 = 0.965$				15.2±0.2	7.3±0.5	17±1	5.1±0.5	16±1							
	14.8	12.5±1.4	33.7±3.3	1.577±0.075						0.507±0.023						
	15.1	12.4±1.4	33.3±3.2	0.946±0.044						0.629±0.019						
	15.4	12.6±1.4	33.7±3.2	0.484±0.032						0.962±0.027						
	15.7	12.4±1.4	33.1±3.2	0.336±0.009						1.514±0.013						
6.5	$\ln(k_{\text{unfolding}}) = (1.77 \pm 0.13) F - (26.63 \pm 2.04), R^2 = 0.994$				14.8±0.1	6.5±0.5	15±1	4.5±0.5	15±1							
	$\ln(k_{\text{folding}}) = (-1.24 \pm 0.13) F + (19.04 \pm 2.02), R^2 = 0.989$															
	Tether1															
	&2	$\ln(k_{\text{unfolding}}) = (1.72 \pm 0.14) F - (25.88 \pm 2.07), R^2 = 0.982$														
	pooled	$\ln(k_{\text{folding}}) = (-1.23 \pm 0.12) F + (19.02 \pm 1.82), R^2 = 0.973$								15.2±0.1	7.3±0.6	17±1	5.1±0.5	16±1		
7.3	14.1	12.5±1.4	34.1±3.3	2.366±0.141	0.461±0.015	14.6±0.1	7.7±0.6	18±1	5.7±0.8	18±2						
	14.4	12.8±1.4	34.6±3.3	1.559±0.081	0.644±0.027											
	14.7	12.8±1.1	34.5±2.6	0.744±0.043	0.871±0.017											
	15.0	12.7±1.1	34.1±2.6	0.456±0.035	1.674±0.079											
	$\ln(k_{\text{unfolding}}) = (1.89 \pm 0.14) F - (27.60 \pm 2.04), R^2 = 0.995$															
Tether1	$\ln(k_{\text{folding}}) = (-1.39 \pm 0.19) F + (20.44 \pm 2.78), R^2 = 0.982$				14.8±0.1	6.5±0.5	15±1	4.5±0.5	15±1							
	14.2	12.6±0.9	34.3±2.1	1.790±0.075						0.423±0.020						
	14.5	12.6±0.9	34.1±2.1	1.305±0.101						0.571±0.025						
	14.8	12.5±1.1	33.7±2.6	0.752±0.072						0.721±0.021						
	15.1	12.5±1.4	33.6±3.2	0.438±0.031						1.186±0.071						
7.3	$\ln(k_{\text{unfolding}}) = (1.59 \pm 0.13) F - (23.25 \pm 1.93), R^2 = 0.993$				14.5±0.2	6.8±1.3	15±3	5.8±0.4	18±1							
	$\ln(k_{\text{folding}}) = (-1.11 \pm 0.13) F + (16.64 \pm 1.91), R^2 = 0.986$															
	Tether1															
	&2	$\ln(k_{\text{unfolding}}) = (1.73 \pm 0.10) F - (25.27 \pm 1.51), R^2 = 0.989$														
	pooled	$\ln(k_{\text{folding}}) = (-1.18 \pm 0.21) F + (17.56 \pm 3.13), R^2 = 0.953$								14.7±0.2	7.0±0.4	16±1	4.8±0.9	16±2		
8.0	14.1	12.7±1.2	34.6±2.8	1.650±0.081	0.553±0.014	14.5±0.2	6.8±1.3	15±3	5.8±0.4	18±1						
	14.4	12.8±1.1	34.6±2.6	1.224±0.072	0.830±0.027											
	14.7	12.7±1.1	34.2±2.6	0.831±0.032	1.424±0.041											
	15.0	12.7±1.2	34.1±2.8	0.349±0.025	1.905±0.090											
	$\ln(k_{\text{unfolding}}) = (1.68 \pm 0.31) F - (24.35 \pm 4.57), R^2 = 0.967$															
Tether1	$\ln(k_{\text{folding}}) = (-1.42 \pm 0.11) F + (20.56 \pm 1.58), R^2 = 0.994$				14.5±0.1	7.2±0.5	17±1	5.5±0.7	17±2							
	14.0	12.7±1.5	34.6±3.5	2.656±0.117						0.540±0.024						
	14.3	12.7±1.5	34.5±3.5	1.581±0.08						0.698±0.020						
	14.6	12.6±1.5	34.1±3.5	1.017±0.095						0.988±0.038						
	14.9	12.7±1.5	34.1±3.5	0.522±0.058						1.849±0.107						
8.0	$\ln(k_{\text{unfolding}}) = (1.77 \pm 0.11) F - (25.83 \pm 1.53), R^2 = 0.996$				14.5±0.1	7.2±0.5	17±1	5.5±0.7	17±2							
	$\ln(k_{\text{folding}}) = (-1.35 \pm 0.20) F + (19.55 \pm 2.89), R^2 = 0.979$															
	Tether1															
	&2	$\ln(k_{\text{unfolding}}) = (1.76 \pm 0.17) F - (25.60 \pm 2.40), R^2 = 0.975$														
	pooled	$\ln(k_{\text{folding}}) = (-1.38 \pm 0.09) F + (20.09 \pm 1.35), R^2 = 0.987$								14.5±0.1	7.2±0.7	17±2	5.6±0.4	18±1		
U25C																
5.0	15.9	12.4±0.9	33.0±2.1	0.680±0.016	0.127±0.004	16.4±0.2	8.4±1.1	19±3	4.6±0.9	15±2						
	16.2	12.2±1.0	32.4±2.3	0.261±0.005	0.137±0.004											
	16.5	12.4±1.0	32.7±2.3	0.191±0.003	0.252±0.008											
	16.8	12.3±0.9	32.4±2.0	0.095±0.004	0.317±0.019											
	$\ln(k_{\text{unfolding}}) = (2.07 \pm 0.28) F - (32.45 \pm 4.56), R^2 = 0.982$															
6.5	$\ln(k_{\text{folding}}) = (-1.12 \pm 0.23) F + (19.92 \pm 3.76), R^2 = 0.960$				15.4±0.1	7.7±0.2	18±0	5.0±0.8	16±2							
	14.9	12.5±1.1	33.9±2.6	0.433±0.014						0.078±0.002						
	15.2	12.6±1.1	33.7±2.5	0.236±0.005						0.089±0.002						
	15.5	12.6±1.1	33.6±2.5	0.131±0.004						0.153±0.003						
	15.8	12.5±1.1	33.2±2.5	0.079±0.003						0.222±0.005						
	$\ln(k_{\text{unfolding}}) = (1.90 \pm 0.06) F - (27.41 \pm 0.85), R^2 = 0.999$															
	$\ln(k_{\text{folding}}) = (-1.23 \pm 0.19) F + (20.92 \pm 2.86), R^2 = 0.978$															
	14.3	12.2±1.0	33.3±2.3	0.208±0.008	0.056±0.001											
	14.6	12.5±1.0	33.8±2.3	0.117±0.003	0.083±0.003											

7.3	14.9	12.4±1.0	33.4±2.3	0.061±0.002	0.135±0.005	14.7±0.1	7.9±0.3	18±1	6.6±0.4	20±1
	15.2	12.2±1.0	32.8±2.3	0.037±0.001	0.242±0.003					
	$\ln(k_{unfolding}) = (1.94 \pm 0.07) F - (26.22 \pm 0.98), R^2 = 0.999$ $\ln(k_{folding}) = (-1.63 \pm 0.10) F + (26.18 \pm 1.48), R^2 = 0.996$									
8.0	13.7	12.8±1.0	35.0±2.4	0.361±0.007	0.042±0.001	14.4±0.1	7.6±0.2	18±0	5.3±0.5	17±1
	14.0	12.8±1.0	34.9±2.4	0.215±0.005	0.064±0.002					
	14.3	12.7±1.0	34.5±2.3	0.123±0.004	0.074±0.001					
	14.6	12.7±1.0	34.3±2.3	0.071±0.002	0.144±0.003					
	14.9	12.6±0.9	33.9±2.1	0.038±0.001	0.195±0.004					
	$\ln(k_{unfolding}) = (1.87 \pm 0.04) F - (24.63 \pm 0.51), R^2 = 0.999$ $\ln(k_{folding}) = (-1.29 \pm 0.13) F + (20.92 \pm 1.88), R^2 = 0.985$									
U29C										
5.0 Tether1	15.7	12.2±1.2	32.6±2.8	0.404±0.010	0.109±0.004	16.2±0.1	7.4±0.4	17±1	4.0±0.5	14±1
	16.0	12.2±1.2	32.5±2.7	0.251±0.003	0.143±0.002					
	16.3	12.1±1.2	32.1±2.7	0.149±0.002	0.175±0.003					
	16.6	12.2±1.2	32.2±2.7	0.078±0.001	0.275±0.005					
	$\ln(k_{unfolding}) = (1.82 \pm 0.09) F - (27.68 \pm 1.49), R^2 = 0.997$ $\ln(k_{folding}) = (-0.99 \pm 0.12) F + (17.83 \pm 1.96), R^2 = 0.985$									
5.0 Tether2	15.9	12.1±1.0	32.3±2.3	0.249±0.003	0.122±0.001	16.2±0.1	6.9±0.7	16±2	4.9±0.3	16±1
	16.1	12.2±0.9	32.4±2.1	0.232±0.006	0.152±0.002					
	16.3	12.1±0.9	32.1±2.0	0.144±0.003	0.214±0.008					
	16.5	12.1±0.9	32.0±2.0	0.100±0.002	0.255±0.03					
	16.7	12.1±0.9	31.9±2.0	0.069±0.001	0.316±0.05					
Tether1 &2 pooled	$\ln(k_{unfolding}) = (1.70 \pm 0.18) F - (25.82 \pm 2.93), R^2 = 0.984$ $\ln(k_{folding}) = (-1.21 \pm 0.07) F + (21.34 \pm 1.19), R^2 = 0.994$					16.2±0.1	7.2±0.4	16±1	4.5±0.3	15±1
	$\ln(k_{unfolding}) = (1.77 \pm 0.09) F - (26.83 \pm 1.44), R^2 = 0.991$									
	$\ln(k_{folding}) = (-1.10 \pm 0.07) F + (19.62 \pm 1.14), R^2 = 0.986$									
6.5 Tether1	14.6	11.6±1.4	31.7±3.3	0.219±0.006	0.048±0.001	15.2±0.1	7.2±0.2	17±0	3.7±0.4	13±1
	14.9	11.3±1.4	30.9±3.3	0.143±0.002	0.057±0.001					
	15.2	11.1±1.3	30.3±3.0	0.081±0.002	0.070±0.001					
	15.5	11.2±1.2	30.4±2.8	0.047±0.001	0.096±0.003					
	15.8	10.6±1.2	28.9±2.8	0.028±0.001	0.144±0.003					
6.5 Tether2	$\ln(k_{unfolding}) = (1.74 \pm 0.04) F - (23.96 \pm 0.66), R^2 = 0.999$ $\ln(k_{folding}) = (-0.91 \pm 0.09) F + (16.34 \pm 1.44), R^2 = 0.984$					15.3±0.1	7.1±0.3	16±1	3.5±0.2	13±0
	14.7	10.0±1.3	28.0±3.0	0.166±0.007	0.035±0.002					
	14.9	10.1±1.2	28.1±2.8	0.108±0.001	0.040±0.001					
	15.2	10.1±1.2	28.0±2.8	0.074±0.000	0.049±0.001					
	15.5	10.0±1.2	27.6±2.8	0.042±0.001	0.066±0.001					
Tether1 &2 pooled	15.8	9.8±1.3	27.0±3.0	0.023±0.001	0.091±0.001	15.3±0.2	7.1±0.3	16±1	3.6±0.7	13±2
	$\ln(k_{unfolding}) = (1.75 \pm 0.08) F - (23.99 \pm 1.17), R^2 = 0.992$									
	$\ln(k_{folding}) = (-0.88 \pm 0.17) F + (16.08 \pm 2.66), R^2 = 0.871$									
7.3 Tether1	14.3	10.6±1.2	29.5±2.8	0.155±0.006	0.034±0.002	14.8±0.1	7.9±0.2	18±0	4.3±0.9	15±2
	14.5	10.4±1.2	29.0±2.8	0.102±0.002	0.048±0.001					
	14.7	10.4±1.2	28.9±2.8	0.069±0.001	0.059±0.001					
	14.9	10.3±1.2	28.6±2.8	0.048±0.001	0.064±0.001					
	$\ln(k_{unfolding}) = (1.95 \pm 0.04) F - (26.06 \pm 0.64), R^2 = 0.999$ $\ln(k_{folding}) = (-1.05 \pm 0.21) F + (18.36 \pm 3.04), R^2 = 0.963$									
7.3 Tether2	14.4	10.5±1.2	29.2±2.8	0.200±0.003	0.031±0.001	15.0±0.1	7.0±0.1	16±0	4.1±0.2	14±0
	14.7	10.3±1.2	28.6±2.8	0.112±0.001	0.041±0.001					
	15.0	10.0±1.2	27.8±2.8	0.067±0.001	0.055±0.001					
	15.3	10.1±1.2	27.9±2.8	0.041±0.001	0.072±0.001					
	15.6	9.9±1.1	27.4±2.5	0.025±0.000	0.106±0.001					
Tether1 &2 pooled	$\ln(k_{unfolding}) = (1.72 \pm 0.03) F - (23.09 \pm 0.48), R^2 = 0.999$ $\ln(k_{folding}) = (-1.01 \pm 0.04) F + (18.00 \pm 0.59), R^2 = 0.998$					15.0±0.2	5.9±0.8	14±2	3.4±0.5	13±1
	$\ln(k_{unfolding}) = (1.46 \pm 0.20) F - (19.10 \pm 3.01), R^2 = 0.939$									
	$\ln(k_{folding}) = (-0.83 \pm 0.13) F + (15.31 \pm 1.91), R^2 = 0.926$									
	14.0	9.9±1.1	28.0±2.6	0.558±0.018	0.027±0.001					
	14.5	10.0±1.1	28.0±2.6	0.237±0.003	0.039±0.000					

	15.0	9.7±1.0	27.1±2.3	0.105±0.001	0.056±0.001					
8.0	15.5	9.5±1.0	26.5±2.3	0.043±0.001	0.093±0.001					
Tether1	16.0	9.7±0.8	26.7±1.8	0.023±0.000	0.174±0.003	15.2±0.1	6.6±0.2	15±0	3.7±0.3	13±1
		ln($k_{\text{unfolding}}$) = (1.62 ± 0.05) F – (22.01 ± 0.75), R ² = 0.999								
		ln(k_{folding}) = (–0.92 ± 0.07) F + (16.56 ± 1.00), R ² = 0.992								
	14.6	10.2±1.2	28.5±2.8	0.152±0.002	0.042±0.001					
	14.9	10.2±1.1	28.3±2.6	0.089±0.001	0.054±0.001					
	15.2	10.2±1.2	28.2±2.8	0.056±0.002	0.076±0.002					
8.0	15.5	10.0±1.1	27.6±2.5	0.037±0.001	0.094±0.001					
Tether2	15.8	9.6±1.1	26.6±2.5	0.020±0.001	0.135±0.001	15.1±0.1	6.7±0.2	16±0	3.9±0.2	14±0
		ln($k_{\text{unfolding}}$) = (1.64 ± 0.06) F – (22.05 ± 0.88), R ² = 0.998								
		ln(k_{folding}) = (–0.96 ± 0.04) F + (17.25 ± 0.62), R ² = 0.997								
Tether1										
&2	ln($k_{\text{unfolding}}$) = (1.67 ± 0.10) F – (22.63 ± 1.49), R ² = 0.986									
pooled	ln(k_{folding}) = (–0.93 ± 0.04) F + (16.70 ± 0.57), R ² = 0.993									
						15.1±0.1	6.8±0.4	16±1	3.8±0.2	13±0

^a The measured nucleotides (nt) are calculated using EWLC model (Equation 3.1) based on the extension change measured at each force with the diameter of the hairpin stem (2 nm) taken into account.

^b The uncertainties for τ are reported as 95% confidence interval for the single-exponential fit. All other experimental data are reported as mean ± standard error.

^c The number of nucleotides disrupted or folded at the unfolding transition state are calculated using EWLC model (Equation 3.1).

^d Experimental data of tether 1 and tether 2 were pooled together and fit to the Bell’s model (Equation 3.2) to extract the kinetic parameters.

Table A1.2. Parameters obtained by passive mode at 200 mM NaCl with varied pH.

pH	Force (pN)	τ_{folded} (s) ^a	Force (pN)	τ_{unfolded} (s)	F_c (pN)	$\Delta X^{\ddagger, \text{FU}}$ (nm)	$\Delta n^{\ddagger, \text{FU}}$	$\Delta X^{\ddagger, \text{UF}}$ (nm)	$\Delta n^{\ddagger, \text{FU}}$
rHP									
5.0		rHP unfolding		rHP folding					
	16.66	0.658±0.019	15.87	0.137±0.004					
	17.30	0.240±0.008	16.50	0.271±0.012					
	17.46	0.150±0.005	16.68	0.250±0.010					
	17.59	0.149±0.010	16.83	0.295±0.018					
	17.83	0.078±0.006	17.07	0.382±0.008					
	18.01	0.057±0.003	17.26	0.433±0.032	17.0±0.1	7.4±0.4	17±1	3.3±0.3	12±1
	$\ln(k_{\text{unfolding}}) = (1.82 \pm 0.10) F - (29.94 \pm 1.80)$, $R^2 = 0.994$ $\ln(k_{\text{folding}}) = (-0.81 \pm 0.08) F + (14.79 \pm 1.33)$, $R^2 = 0.981$								
6.5		rHP unfolding		rHP folding					
	16.63	0.540±0.023	15.79	0.059±0.003					
	16.88	0.425±0.015	16.05	0.088±0.003					
	17.32	0.151±0.004	16.50	0.147±0.003					
	17.38	0.153±0.003	16.54	0.154±0.008					
	17.50	0.108±0.006	16.67	0.159±0.013					
	17.64	0.095±0.005	16.82	0.258±0.010					
	17.82	0.075±0.004	16.97	0.231±0.009					
	17.97	0.049±0.001	17.13	0.321±0.009					
	18.14	0.039±0.002	17.30	0.363±0.006	17.0±0.1	7.3±0.3	16±1	4.9±0.3	16±1
	$\ln(k_{\text{unfolding}}) = (1.80 \pm 0.07) F - (29.32 \pm 1.18)$, $R^2 = 0.995$ $\ln(k_{\text{folding}}) = (-1.20 \pm 0.07) F + (21.76 \pm 1.17)$, $R^2 = 0.988$								
7.3		rHP unfolding		rHP folding					
	16.18	1.069±0.054	15.36	0.043±0.003					
	16.89	0.281±0.007	16.08	0.101±0.003					
	16.94	0.291±0.008	16.13	0.115±0.003					
	17.38	0.111±0.003	16.56	0.168±0.004					

	17.41	0.111±0.002	16.59	0.172±0.008					
	17.76	0.052±0.001	16.95	0.288±0.006					
	17.91	0.047±0.002	17.10	0.339±0.009	16.9±0.1	7.6±0.2	17±0	4.8±0.2	15±0
		$\ln(k_{\text{unfolding}}) = (1.86 \pm 0.05) F - (30.24 \pm 0.91), R^2 = 0.998$ $\ln(k_{\text{folding}}) = (-1.17 \pm 0.04) F + (21.11 \pm 0.58), R^2 = 0.998$							
8.0		rHP unfolding		rHP folding					
	16.74	0.297±0.007	16.14	0.097±0.006					
	16.84	0.261±0.011	16.25	0.125±0.011					
	16.96	0.220±0.008	16.37	0.156±0.005					
	17.14	0.198±0.005	16.53	0.171±0.010					
	17.23	0.116±0.005	16.62	0.188±0.009					
	17.42	0.102±0.005	16.81	0.314±0.019					
	17.71	0.058±0.002	17.02	0.337±0.007	16.8±0.1	7.0±0.6	16±1	5.8±0.5	18±1
		$\ln(k_{\text{unfolding}}) = (1.71 \pm 0.15) F - (27.42 \pm 2.62), R^2 = 0.981$ $\ln(k_{\text{folding}}) = (-1.43 \pm 0.13) F + (25.35 \pm 2.21), R^2 = 0.979$							
U22C									
5.0		U22C unfolding		U22C folding					
	14.72	1.012±0.042	13.92	0.050±0.002					
	14.94	0.807±0.044	14.15	0.070±0.004					
	15.06	0.596±0.024	14.24	0.090±0.004					
	15.42	0.348±0.009	14.61	0.127±0.002					
	15.95	0.128±0.006	15.14	0.317±0.014					
	16.04	0.096±0.002	15.23	0.296±0.004					
	16.08	0.101±0.005	15.26	0.271±0.006					
	16.13	0.112±0.003	15.31	0.306±0.010					
	16.64	0.040±0.002	15.82	0.776±0.024					
	16.75	0.034±0.002	15.94	0.559±0.029	15.4±0.1	7.0±0.2	16±0	5.2±0.3	17±1
	$\ln(k_{\text{unfolding}}) = (1.72 \pm 0.04) F - (25.33 \pm 0.68), R^2 = 0.997$ $\ln(k_{\text{folding}}) = (-1.28 \pm 0.07) F + (20.73 \pm 1.10), R^2 = 0.987$								
6.5	15.68	0.233±0.017	14.91	0.592±0.020					
	15.54	0.314±0.012	14.75	0.557±0.015					
	15.45	0.369±0.011	14.66	0.542±0.017					
	15.37	0.427±0.028	14.62	0.485±0.020					
	15.17	0.605±0.04	14.38	0.312±0.008					
	14.98	0.682±0.023	14.22	0.292±0.012					
	14.76	0.829±0.034	13.99	0.209±0.006					
	14.69	0.891±0.023	13.92	0.234±0.010					
	14.49	1.572±0.153	13.73	0.211±0.013					
	14.40	1.909±0.089	13.68	0.183±0.010	15.0±0.1	6.0±0.4	14±1	4.1±0.3	14±1
	$\ln(k_{\text{unfolding}}) = (1.48 \pm 0.09) F - (21.73 \pm 1.35), R^2 = 0.985$ $\ln(k_{\text{folding}}) = (-1.01 \pm 0.08) F + (15.59 \pm 1.12), R^2 = 0.977$								
7.3	14.35	1.274±0.058	13.60	0.262±0.015					
	14.52	0.933±0.032	13.77	0.347±0.018					
	15.03	0.388±0.020	14.29	0.701±0.019					
	15.17	0.251±0.016	14.41	0.675±0.019					
	15.48	0.243±0.009	14.72	1.087±0.040	14.7±0.1	6.5±0.9	15±2	5.0±0.4	16±1
		$\ln(k_{\text{unfolding}}) = (1.59 \pm 0.21) F - (23.00 \pm 3.11), R^2 = 0.975$ $\ln(k_{\text{folding}}) = (-1.23 \pm 0.09) F + (18.07 \pm 1.24), R^2 = 0.992$							
8.0	14.78	1.544±0.104	14.10	0.301±0.014					
	14.97	1.149±0.042	14.26	0.477±0.014					
	15.16	0.877±0.037	14.47	0.553±0.025					
	15.18	0.895±0.051	14.49	0.630±0.030					
	16.15	0.182±0.007	15.44	2.955±0.200	14.9±0.1	6.4±0.2	15±0	6.7±0.4	20±1
		$\ln(k_{\text{unfolding}}) = (1.57 \pm 0.04) F - (23.63 \pm 0.57), R^2 = 0.999$ $\ln(k_{\text{folding}}) = (-1.65 \pm 0.09) F + (24.42 \pm 1.32), R^2 = 0.996$							
U25C									
5.0		U25C unfolding		U25C folding					
	15.76	0.358±0.009	15.08	0.059±0.005					
	15.91	0.301±0.014	15.20	0.097±0.006					
	16.10	0.228±0.007	15.40	0.107±0.006					
	16.14	0.234±0.005	15.46	0.111±0.006					

	16.23	0.204±0.010	15.54	0.142±0.004					
	16.33	0.184±0.008	15.63	0.165±0.010					
	16.51	0.133±0.006	15.81	0.208±0.014					
	16.69	0.094±0.002	15.99	0.228±0.005					
	16.80	0.065±0.003	16.11	0.254±0.007					
	16.90	0.049±0.003	16.21	0.315±0.014					
	17.01	0.038±0.002	16.31	0.385±0.012	16.0±0.1	7.3±0.4	17±1	5.5±0.3	17±1
		$\ln(k_{\text{unfolding}}) = (1.80 \pm 0.11) F - (27.46 \pm 1.81), R^2 = 0.983$							
		$\ln(k_{\text{folding}}) = (-1.34 \pm 0.08) F + (22.90 \pm 1.29), R^2 = 0.984$							
6.5	15.68	0.101±0.003	14.86	0.066±0.001					
	15.73	0.087±0.001	14.92	0.074±0.001					
	15.77	0.077±0.002	14.95	0.074±0.001					
	15.88	0.070±0.001	15.07	0.090±0.001					
	16.48	0.023±0.001	15.66	0.194±0.009	15.4±0.1	7.3±0.3	17±1	5.5±0.1	17±0
		$\ln(k_{\text{unfolding}}) = (1.79 \pm 0.08) F - (25.77 \pm 1.24), R^2 = 0.997$							
		$\ln(k_{\text{folding}}) = (-1.34 \pm 0.03) F + (22.54 \pm 0.47), R^2 = 0.999$							
7.3	14.41	0.109±0.002	13.75	0.015±0.001					
	14.83	0.054±0.001	14.18	0.029±0.001					
	15.01	0.042±0.001	14.36	0.035±0.001					
	15.21	0.033±0.001	14.53	0.042±0.001					
	15.32	0.026±0.000	14.63	0.060±0.001					
	15.49	0.019±0.001	14.83	0.086±0.002					
	15.95	0.009±0.000	15.32	0.139±0.003	14.8±0.1	6.6±0.2	15±0	5.9±0.4	18±1
		$\ln(k_{\text{unfolding}}) = (1.61 \pm 0.04) F - (20.99 \pm 0.55), R^2 = 0.999$							
		$\ln(k_{\text{folding}}) = (-1.46 \pm 0.09) F + (24.33 \pm 1.27), R^2 = 0.991$							
8.0	14.24	0.139±0.003	13.43	0.025±0.000					
	14.35	0.116±0.004	13.54	0.034±0.001					
	14.43	0.089±0.002	13.62	0.033±0.001					
	14.49	0.084±0.002	13.67	0.040±0.001					
	14.66	0.055±0.001	13.85	0.051±0.001					
	14.81	0.049±0.001	14.00	0.063±0.002					
	14.94	0.037±0.000	14.13	0.071±0.001	14.4±0.1	7.7±0.4	18±1	5.9±0.4	17±1
		$\ln(k_{\text{unfolding}}) = (1.88 \pm 0.11) F - (24.80 \pm 1.59), R^2 = 0.992$							
		$\ln(k_{\text{folding}}) = (-1.46 \pm 0.11) F + (23.28 \pm 1.46), R^2 = 0.987$							
U29C									
		U29C							
		unfolding		U29C folding					
5.0 Tether1	15.78	0.390±0.012	14.99	0.052±0.001					
	15.93	0.304±0.013	15.14	0.057±0.004					
	15.99	0.254±0.006	15.21	0.064±0.003					
	16.22	0.158±0.002	15.43	0.077±0.002					
	16.47	0.123±0.002	15.69	0.102±0.001					
	16.62	0.097±0.002	15.85	0.097±0.002					
	16.66	0.072±0.001	15.89	0.108±0.003					
	16.73	0.074±0.002	15.95	0.117±0.003					
	16.84	0.056±0.001	16.08	0.131±0.003					
	17.03	0.036±0.001	16.25	0.188±0.005					
	17.08	0.035±0.002	16.30	0.178±0.005					
	17.15	0.032±0.001	16.38	0.206±0.004	16.3±0.1	7.5±0.2	17±0	3.9±0.2	14±0
		$\ln(k_{\text{unfolding}}) = (1.83 \pm 0.05) F - (28.01 \pm 0.88), R^2 = 0.996$							
		$\ln(k_{\text{folding}}) = (-0.96 \pm 0.05) F + (17.39 \pm 0.85), R^2 = 0.985$							
5.0 Tether2	15.71	0.391±0.014	14.93	0.052±0.001					
	15.78	0.436±0.017	15.00	0.052±0.002					
	15.94	0.265±0.008	15.16	0.073±0.003					
	16.26	0.156±0.005	15.47	0.085±0.003					
	16.34	0.128±0.003	15.55	0.088±0.002					
	16.35	0.124±0.003	15.56	0.089±0.002					
	16.82	0.055±0.002	16.04	0.153±0.004					
	16.87	0.053±0.003	16.09	0.155±0.003					
	16.88	0.049±0.002	16.10	0.164±0.004					
	16.92	0.043±0.001	16.14	0.140±0.003					
	17.24	0.023±0.000	16.44	0.199±0.008					
	17.26	0.029±0.001	16.48	0.211±0.006	16.2±0.1	7.5±0.2	17±0	3.7±0.2	14±0

$\ln(k_{\text{unfolding}}) = (1.83 \pm 0.05) F - (27.87 \pm 0.84), R^2 = 0.996$ $\ln(k_{\text{folding}}) = (-0.90 \pm 0.04) F + (16.42 \pm 0.62), R^2 = 0.991$ $\ln(k_{\text{unfolding}}) = (1.83 \pm 0.04) F - (27.92 \pm 0.61), R^2 = 0.996$ $\ln(k_{\text{folding}}) = (-0.93 \pm 0.03) F + (16.82 \pm 0.55), R^2 = 0.985$									
Tether1& 2 pooled ^c					16.2±0.1	7.5±0.2	17±0	3.8±0.1	13±0
6.5 Tether1	15.23	0.082±0.002	14.61	0.040±0.001					
	15.31	0.071±0.002	14.70	0.049±0.002					
	15.33	0.070±0.002	14.72	0.037±0.003					
	15.40	0.063±0.001	14.78	0.041±0.001					
	15.81	0.029±0.001	15.23	0.074±0.001					
	15.89	0.027±0.001	15.31	0.090±0.002					
	16.14	0.018±0.001	15.57	0.106±0.005					
	16.21	0.018±0.000	15.63	0.112±0.003					
	16.33	0.013±0.000	15.78	0.129±0.005					
	16.62	0.009±0.000	16.08	0.189±0.010	15.3±0.1	6.6±0.2	15±0	4.4±0.2	15±0
$\ln(k_{\text{unfolding}}) = (1.61 \pm 0.04) F - (22.03 \pm 0.57), R^2 = 0.998$ $\ln(k_{\text{folding}}) = (-1.08 \pm 0.06) F + (19.02 \pm 0.94), R^2 = 0.987$									
6.5 Tether2	14.6	0.176±0.004	13.94	0.017±0.001					
	14.76	0.116±0.004	14.09	0.021±0.001					
	14.8	0.107±0.003	14.14	0.019±0.000					
	15.02	0.077±0.001	14.36	0.022±0.000					
	15.46	0.040±0.001	14.81	0.037±0.001					
	15.48	0.038±0.001	14.82	0.035±0.001					
	15.84	0.025±0.000	15.18	0.047±0.001					
	15.86	0.022±0.000	15.20	0.047±0.000					
	16.31	0.013±0.000	15.67	0.079±0.001	15.3±0.1	5.9±0.2	14±0	3.5±0.1	13±0
$\ln(k_{\text{unfolding}}) = (1.45 \pm 0.04) F - (19.28 \pm 0.60), R^2 = 0.997$ $\ln(k_{\text{folding}}) = (-0.87 \pm 0.03) F + (16.29 \pm 0.42), R^2 = 0.995$ $\ln(k_{\text{unfolding}}) = (1.44 \pm 0.06) F - (19.25 \pm 0.92), R^2 = 0.986$									
Tether1& 2 pooled					15.2±0.1	5.9±0.2	14±0	4.4±0.3	15±1
7.3 Tether1	14.71	0.075±0.001	14.16	0.030±0.000					
	14.77	0.071±0.001	14.22	0.035±0.000					
	14.89	0.063±0.001	14.34	0.038±0.001					
	14.94	0.053±0.001	14.39	0.037±0.001					
	15.14	0.037±0.000	14.60	0.051±0.001					
	15.66	0.017±0.000	15.12	0.083±0.001	14.8±0.1	6.6±0.3	15±1	4.2±0.2	14±0
$\ln(k_{\text{unfolding}}) = (1.61 \pm 0.07) F - (21.12 \pm 0.98), R^2 = 0.997$ $\ln(k_{\text{folding}}) = (-1.03 \pm 0.06) F + (18.08 \pm 0.94), R^2 = 0.992$									
7.3 Tether2	14.42	0.130±0.002	13.85	0.018±0.001					
	14.66	0.092±0.001	14.09	0.024±0.001					
	14.69	0.080±0.001	14.13	0.021±0.000					
	15.03	0.057±0.001	14.47	0.032±0.000					
	15.13	0.049±0.000	14.57	0.037±0.000					
	15.19	0.043±0.001	14.63	0.036±0.000					
	15.57	0.024±0.000	15.01	0.056±0.001					
	15.61	0.022±0.000	15.06	0.054±0.001					
	16.25	0.009±0.000	15.72	0.111±0.002	15.0±0.1	5.9±0.1	14±0	4.0±0.2	14±0
$\ln(k_{\text{unfolding}}) = (1.46 \pm 0.03) F - (18.95 \pm 0.48), R^2 = 0.998$ $\ln(k_{\text{folding}}) = (-0.97 \pm 0.04) F + (17.51 \pm 0.52), R^2 = 0.995$ $\ln(k_{\text{unfolding}}) = (1.46 \pm 0.06) F - (18.98 \pm 0.88), R^2 = 0.989$									
Tether1& 2 pooled					15.0±0.1	5.9±0.2	14±0	3.8±0.4	13±1
8.0 Tether1	13.86	0.831±0.044	13.16	0.010±0.000					
	14.28	0.427±0.016	13.59	0.013±0.000					
	14.34	0.436±0.019	13.65	0.012±0.001					
	14.46	0.395±0.013	13.78	0.019±0.001					
	14.94	0.140±0.007	14.25	0.028±0.000					
	15.02	0.098±0.002	14.35	0.029±0.001					
	15.13	0.097±0.002	14.46	0.035±0.000					

	15.27	0.087±0.002	14.60	0.039±0.001					
	15.65	0.037±0.001	15.00	0.054±0.002					
	15.72	0.045±0.002	15.05	0.054±0.001					
	15.79	0.033±0.00	15.12	0.058±0.001					
	16.20	0.019±0.00	15.59	0.095±0.001					
	16.50	0.011±0.00	15.90	0.129±0.003					
	16.70	0.010±0.00	16.00	0.137±0.003					
	17.10	0.006±0.000	16.47	0.271±0.007	15.3±0.1	6.5±0.2	14±0	4.0±0.1	14±0
	$\ln(k_{\text{unfolding}}) = (1.59 \pm 0.05) F - (21.76 \pm 0.68), R^2 = 0.995$								
	$\ln(k_{\text{folding}}) = (-0.97 \pm 0.02) F + (17.42 \pm 0.32), R^2 = 0.997$								
8.0	14.66	0.102±0.001	14.03	0.021±0.001					
Tether2	14.71	0.108±0.002	14.07	0.023±0.001					
	14.95	0.068±0.002	14.32	0.031±0.000					
	15.02	0.066±0.001	14.39	0.033±0.001					
	15.26	0.043±0.001	14.63	0.040±0.001					
	15.53	0.032±0.000	14.90	0.065±0.001					
	15.57	0.032±0.001	14.94	0.056±0.001					
	15.76	0.022±0.000	15.14	0.073±0.002					
	15.90	0.019±0.001	15.29	0.074±0.001					
	15.94	0.017±0.000	15.33	0.083±0.001					
	15.95	0.016±0.000	15.34	0.095±0.001					
	16.10	0.015±0.000	15.49	0.101±0.002					
	16.18	0.014±0.000	15.56	0.109±0.001					
	16.35	0.011±0.000	15.76	0.115±0.002	15.1±0.1	5.6±0.1	13±0	4.1±0.2	14±0
	$\ln(k_{\text{unfolding}}) = (1.37 \pm 0.03) F - (17.86 \pm 0.50), R^2 = 0.997$								
	$\ln(k_{\text{folding}}) = (-1.01 \pm 0.04) F + (18.02 \pm 0.59), R^2 = 0.991$								
	$\ln(k_{\text{unfolding}}) = (1.58 \pm 0.08) F - (21.33 \pm 1.24), R^2 = 0.967$								
Tether1&	$\ln(k_{\text{folding}}) = (-0.99 \pm 0.02) F + (17.65 \pm 0.31),$								
2 pooled	$R^2 = 0.994$				15.1±0.1	6.4±0.3	15±1	4.0±0.1	14±0
^a The uncertainties for τ are reported as 95% confidence interval for the single-exponential fit. All other experimental data are reported as mean $\pm$ standard error. ^b The number of nucleotides disrupted or folded at the unfolding transition state is calculated using EWLC model (Equation 3.1). ^c Force-dependent kinetics of tether 1 and tether 2 are pooled together and fit to the Bell's model (Equation 3.2) to extract the kinetic parameters.									

Appendix 2. The parameters obtained from mechanical pulling experiments for tau pre-mRNA splice site RNA hairpins of WT, +14U, +19G and +14U+19G.

Table A2.1. Parameters of WT and +19G from constant-force experiments in 100 μ M spermine, 5 mM MgCl₂, 200 mM NaCl, 10 mM HEPES, 0.1 mM EDTA, pH7.3.

	Force (pN)	Measured ΔX (nm)	Measured nt ^a	τ_{folded} (s) ^b	τ_{unfolded} (s)	F_c (pN)	$\Delta X^{\ddagger, \text{FU}}$ (nm)	$\Delta n^{\ddagger, \text{FU}}$	$\Delta X^{\ddagger, \text{UF}}$ (nm)	$\Delta n^{\ddagger, \text{FU}}$						
WT-1	13.1	6.6±1.3	20.6±3.1	0.090±0.002	0.053±0.003	13.5±0.1	3.2±0.2	8±1	2.9±0.2	12±1						
	13.4	6.4±1.3	20.0±3.1	0.077±0.001	0.061±0.003											
	13.7	6.4±1.2	19.9±2.8	0.059±0.001	0.083±0.003											
	14.0	6.4±1.2	19.8±2.8	0.044±0.001	0.104±0.002											
	14.3	6.3±1.3	19.5±3.0	0.037±0.001	0.119±0.002											
	$\ln(k_{\text{unfolding}}) = (0.78 \pm 0.05) F - (7.83 \pm 0.63)$, $R^2 = 0.995$															
	$\ln(k_{\text{folding}}) = (-0.72 \pm 0.05) F + (12.35 \pm 0.75)$, $R^2 = 0.991$															
WT-2	12.0	6.7±1.3	21.3±3.2	0.329±0.009	0.026±0.003	13.6±0.1	3.7±0.2	9±1	2.5±0.2	11±1						
	12.5	6.5±1.3	20.6±3.1	0.184±0.003	0.037±0.002											
	12.7	6.5±1.3	20.5±3.1	0.159±0.002	0.040±0.001											
	13.0	6.6±1.3	20.6±3.1	0.123±0.003	0.050±0.004											
	13.1	6.5±1.3	20.3±3.1	0.110±0.004	0.051±0.004											
	13.2	6.4±1.3	20.1±3.1	0.108±0.003	0.053±0.003											
	13.5	6.4±1.2	20.0±2.9	0.078±0.002	0.065±0.003											
	13.7	6.5±1.2	20.1±2.8	0.062±0.002	0.084±0.002											
	$\ln(k_{\text{unfolding}}) = (0.91 \pm 0.05) F - (9.69 \pm 0.66)$, $R^2 = 0.989$															
$\ln(k_{\text{folding}}) = (-0.62 \pm 0.04) F + (11.12 \pm 0.51)$, $R^2 = 0.986$																
WT-3	12.8	6.6±1.5	20.7±3.6	0.123±0.002	0.044±0.002	13.5±0.1	3.6±0.1	9±0	3.0±0.2	12±1						
	13.1	6.5±1.4	20.3±3.4	0.096±0.001	0.051±0.003											
	13.4	6.5±1.4	20.2±3.3	0.071±0.002	0.063±0.002											
	13.7	6.6±1.5	20.4±3.6	0.054±0.001	0.083±0.003											
	14.0	6.7±1.4	20.5±3.3	0.043±0.001	0.105±0.003											
	$\ln(k_{\text{unfolding}}) = (0.89 \pm 0.02) F - (9.33 \pm 0.30)$, $R^2 = 0.999$															
	$\ln(k_{\text{folding}}) = (-0.74 \pm 0.04) F + (12.67 \pm 0.60)$, $R^2 = 0.995$															
+19G -1	13.3	9.3±1.3	26.9±3.1	0.343±0.013	0.067±0.003	14.0±0.1	6.0±0.1	14±0	3.6±0.2	13±1						
	13.6	9.4±1.2	27.0±2.8	0.224±0.010	0.092±0.001											
	13.9	9.4±1.2	26.9±2.8	0.135±0.007	0.118±0.001											
	14.2	9.5±1.2	27.0±2.8	0.087±0.002	0.141±0.003											
	14.5	9.3±1.2	26.4±2.8	0.057±0.002	0.184±0.003											
	14.8	9.2±1.1	26.0±2.6	0.038±0.001	0.268±0.008											
	$\ln(k_{\text{unfolding}}) = (1.48 \pm 0.02) F - (18.61 \pm 0.35)$, $R^2 = 0.999$															
	$\ln(k_{\text{folding}}) = (-0.88 \pm 0.04) F + (14.33 \pm 0.62)$, $R^2 = 0.995$															
	+19G -2	13.5	9.0±1.1	26.1±2.6	0.425±0.015						0.080±0.003	14.2±0.1	6.3±0.4	15±1	3.3±0.4	12±1
		13.8	9.3±1.2	26.7±2.8	0.302±0.006						0.099±0.002					
		14.1	9.3±1.2	26.6±2.8	0.144±0.006						0.170±0.004					
		14.4	9.3±1.2	26.4±2.8	0.098±0.004						0.167±0.002					
		14.7	9.2±1.2	26.1±2.8	0.068±0.001						0.214±0.003					
		15.0	9.1±1.1	25.7±2.6	0.044±0.002						0.281±0.004					
		$\ln(k_{\text{unfolding}}) = (1.54 \pm 0.09) F - (19.96 \pm 1.24)$, $R^2 = 0.994$														
$\ln(k_{\text{folding}}) = (-0.82 \pm 0.10) F + (13.51 \pm 1.39)$, $R^2 = 0.973$																
+19G -3	13.6	9.2±1.2	26.6±2.8	0.284±0.016	0.069±0.001	14.2±0.1	6.1±0.5	14±1	2.8±0.2	11±1						
	13.9	9.5±1.3	27.1±3.1	0.139±0.007	0.085±0.003											
	14.2	9.2±1.3	26.3±3.1	0.109±0.006	0.110±0.002											
	14.5	9.0±1.2	25.7±2.8	0.062±0.003	0.121±0.003											
	14.8	9.2±1.2	26.0±2.8	0.045±0.003	0.163±0.002											
	$\ln(k_{\text{unfolding}}) = (1.50 \pm 0.13) F - (19.00 \pm 1.88)$, $R^2 = 0.989$															
$\ln(k_{\text{folding}}) = (-0.69 \pm 0.05) F + (12.06 \pm 0.73)$, $R^2 = 0.992$																

^aThe measured nucleotides (nt) and the number of nucleotides disrupted or folded at the unfolding transition state are calculated using EWLC model based on the extension change measured at each force with the diameter of the hairpin stem (2 nm) taken into account.

^bThe uncertainties for τ are reported as 95% confidence interval for the single-exponential fitting. All other experimental data are reported as mean $\pm$ standard error.

Table A2.2. Parameters of +14U and +14U+19G from constant-force experiments in 100 μ M spermine, 5 mM MgCl₂, 200 mM NaCl, 10 mM HEPES, 0.1 mM EDTA, pH7.3.

RNA	Force (pN)	Measured ΔX (nm)	Measured nt ^a	p_{folded} ^b	p_{unfolded} ^b	F_c (pN)	
+14U-1	8.8	6.3±1.4	22.3±3.8	0.872	0.128	10.0±0.2	
	9.3	6.5±1.4	22.4±3.7	0.761	0.239		
	9.8	6.4±1.5	21.8±3.9	0.541	0.459		
	10.3	6.3±1.6	21.2±4.1	0.346	0.654		
+14U-2	9.3	6.3±1.5	21.9±4.0	0.735	0.265	10.1±0.2	
	9.6	6.2±1.6	21.4±4.2	0.645	0.355		
	9.9	6.2±1.7	21.2±4.4	0.563	0.437		
	10.2	6.0±1.6	20.5±4.1	0.426	0.574		
	10.5	5.9±1.5	20.1±3.8	0.373	0.627		
	10.8	6.0±1.5	20.2±3.8	0.273	0.727		
+14U+19G-1	11.0	7.8±1.3	24.6±3.3	0.807	0.193	11.6±0.2	
	11.3	7.9±1.3	24.6±3.2	0.695	0.305		
	11.6	8.1±1.4	24.9±3.5	0.511	0.489		
	11.9	8.2±1.4	25.0±3.4	0.345	0.655		
	12.2	8.0±1.4	24.4±3.4	0.210	0.790		
+14U+19G-2	11.2	9.0±1.6	27.4±4.0	0.873	0.127	12.1±0.2	
	11.5	8.4±1.6	25.7±4.0	0.761	0.239		
	11.8	8.3±1.6	25.3±3.9	0.651	0.349		
	12.1	8.1±1.6	24.7±3.9	0.541	0.459		
	12.4	8.3±1.6	25.0±3.9	0.429	0.571		
	+14U+19G-2	12.7	8.1±1.6	24.9±3.9	0.284		0.716

^a The measured nucleotides (nt) are calculated using EWLC model based on the extension change measured at each force with the diameter of the hairpin stem (2 nm) taken into account.

^b The probability of the folded state (p_{folded}) or the unfolded sate (p_{unfolded}) are obtained from the extension distribution. All experimental data are reported as mean ± standard error.

Table A2.3. Parameters of WT, +19G, +14U and +14U+19G from passive mode in 100 μ M spermine, 5 mM MgCl₂, 200 mM NaCl, 10 mM HEPES, 0.1 mM EDTA, pH7.3.

RNA	Force (pN)	τ_{folded} (s) ^a	Force (pN)	τ_{unfolded} (s) ^a	F_c (pN)	$\Delta X^{\ddagger, \text{FU}}$ (nm)	$\Delta n^{\ddagger, \text{FU}}$	$\Delta X^{\ddagger, \text{UF}}$ (nm)	$\Delta n^{\ddagger, \text{UF}}$
WT	13.03	0.052 $\pm$ 0.003	12.62	0.024 $\pm$ 0.001					
	13.17	0.044 $\pm$ 0.001	12.77	0.028 $\pm$ 0.001					
	13.31	0.041 $\pm$ 0.002	12.92	0.030 $\pm$ 0.001					

	13.66	0.031±0.001	13.26	0.038±0.001					
	13.94	0.026±0.001	13.53	0.055±0.001					
	14.04	0.023±0.001	13.64	0.056±0.001					
	14.07	0.020±0.001	13.68	0.060±0.001					
	14.25	0.020±0.001	13.85	0.061±0.001					
	14.27	0.017±0.001	13.87	0.067±0.001	13.3±0.1	3.4±0.2	8±1	3.3±0.2	13±1
	$\ln(k_{\text{unfolding}}) = (0.83 \pm 0.05) F - (7.83 \pm 0.62), R^2 = 0.990$ $\ln(k_{\text{folding}}) = (-0.82 \pm 0.04) F + (14.03 \pm 0.54), R^2 = 0.992$								
+19G	13.50	0.368±0.028	12.99	0.051±0.003					
	14.09	0.184±0.010	13.54	0.074±0.003					
	14.16	0.159±0.003	13.60	0.069±0.002					
	14.31	0.149±0.004	13.76	0.083±0.003					
	14.76	0.071±0.002	14.20	0.130±0.002					
	15.08	0.032±0.002	14.56	0.178±0.004					
	15.16	0.030±0.001	14.65	0.178±0.004					
	15.27	0.028±0.001	14.84	0.204±0.005	14.3±0.1	6.4±0.4	15±1	3.3±0.2	12±1
	$\ln(k_{\text{unfolding}}) = (1.56 \pm 0.09) F - (20.18 \pm 1.25), R^2 = 0.991$ $\ln(k_{\text{folding}}) = (-0.80 \pm 0.04) F + (13.48 \pm 0.56), R^2 = 0.993$								
+14U	9.14	0.029±0.001	8.77	0.008±0.001					
	9.31	0.025±0.001	8.97	0.009±0.001					
	9.56	0.024±0.001	9.18	0.010±0.001					
	9.77	0.022±0.001	9.39	0.013±0.001					
	9.94	0.018±0.001	9.55	0.015±0.001					
	10.07	0.016±0.001	9.69	0.017±0.001					
	10.35	0.014±0.001	9.97	0.022±0.001					
	10.59	0.012±0.001	10.19	0.026±0.001					
	10.81	0.009±0.001	10.42	0.032±0.001					
	10.99	0.008±0.001	10.60	0.035±0.001	9.9±0.1	2.8±0.2	7±1	3.5±0.1	14±0
	$\ln(k_{\text{unfolding}}) = (0.69 \pm 0.04) F - (2.84 \pm 0.38), R^2 = 0.988$ $\ln(k_{\text{folding}}) = (-0.85 \pm 0.02) F + (12.36 \pm 0.21), R^2 = 0.998$								
+14U+	11.70	0.029±0.001	11.31	0.012±0.001					
19G	11.87	0.026±0.001	11.48	0.013±0.001					
	12.18	0.022±0.001	11.80	0.020±0.002					
	12.33	0.020±0.001	11.95	0.021±0.002					
	12.39	0.019±0.001	12.02	0.024±0.003					
	13.01	0.010±0.001	12.64	0.042±0.001					
	13.08	0.010±0.001	12.69	0.046±0.001	12.0±0.1	3.3±0.2	8±1	4.0±0.2	15±1
	$\ln(k_{\text{unfolding}}) = (0.81 \pm 0.05) F - (6.801 \pm 0.66), R^2 = 0.989$ $\ln(k_{\text{folding}}) = (-0.98 \pm 0.04) F + (15.54 \pm 0.43), R^2 = 0.997$								
^a The uncertainties for τ are reported as 95% confidence interval for the single-exponential fitting. All other experimental data are reported as mean $\pm$ standard error.									
^b The number of nucleotides disrupted or folded at the unfolding transition state is calculated using EWLC model.									

List of Publications

1. **L. X. Yang**, Z. S. Zhong, C. L. Tong, H. Jia, Y. R. Liu, G. Chen. (2018) Single-Molecule Mechanical Folding and Unfolding of RNA Hairpins: Effects of Single A-U to A-C Pair Substitutions and Single Proton Binding and Implications for mRNA Structure-Induced -1 Ribosomal Frameshifting. [*J Am Chem Soc*](#), 140, 8172-8184.
2. **L. X. Yang**, D. F. K. Toh, Z. S. Zhong, G. Chen. (2019) Base Triple Formation between Loop 2 and Stem 1 in SRV-1 Pseudoknot Inhibits Antisense PNA Invasion of Stem 1 and Stem 2. Submitted.
3. J. Z. Tan, = **L. X. Yang**, = A. A. L. Ong, J. H. Shi, Z. S. Zhong, M. L. Lye, S. Y. Liu, J. Lisowiec-Wachnicka, R. Kierzek, X. Roca, G. Chen. (2019) A Disease-Causing Intronic Point Mutation C19G Alters Tau Exon 10 Splicing via RNA Secondary Structure Rearrangement. [*Biochemistry*](#), 58, 1565-1578. (= These authors contributed equally to this work).
4. J. Keszy, K. M. Patil, S. R. Kumar, Z. Shu, H. Y. Yong, L. Zimmermann, A. A. L. Ong, D. K. Toh, M. S. Krishna, **L. X. Yang**, J. L. Decout, D. Luo, M. Prabakaran, G. Chen, E. Kierzek. (2019) A Short Chemically Modified dsRNA-Binding PNA (dbPNA) Inhibits Influenza Viral Replication by Targeting Viral RNA Panhandle Structure. [*Bioconjug Chem*](#), 30, 931-943.
5. R. Y. Puah, = H. Jia, = M. Maraswami, = D. F. K. Toh, = R. Ero, = **L. X. Yang**, K. M. Patil, A. A. L. Ong, M. S. Krishna, R. M. Sun, C. L. Tong, M. Huang, X. Chen, T. P. Loh, Y. G. Gao, D. X. Liu, and G. Chen. (2018) Selective Binding to mRNA Duplex Regions by Chemically Modified Peptide Nucleic Acids Stimulates Ribosomal Frameshifting. [*Biochemistry*](#), 57, 149-159. (= These authors contributed equally to this work).
6. Z. S. Zhong, **L. X. Yang**, H. P. Zhang, J. H. Shi, J. J. Vandana, D. T. Lam, R. C. L. Olsthoorn, L. Y. Lu, and G. Chen. (2016) Mechanical unfolding kinetics of the SRV-1 gag-pro mRNA pseudoknot: possible implications for -1 ribosomal frameshifting stimulation. [*Sci Rep*](#), 6, 39549.