Unravelling the roles of non-Watson-Crick and triplex structures in regulating RNA folding dynamics and function

Zhong, Zhensheng

2016

Zhong, Z. (2016). Unravelling the roles of non-Watson-Crick and triplex structures in regulating RNA folding dynamics and function. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/66532>

<https://doi.org/10.32657/10356/66532>

**Unravelling the roles of non-Watson-Crick and triplex
structures in regulating RNA folding dynamics and
function**

ZHONG ZHENSHENG

SCHOOL OF PHYSICAL & MATHEMATICAL SCIENCES

2015

**Unravelling the roles of non-Watson-Crick and triplex
structures in regulating RNA folding dynamics and
function**

ZHONG ZHENSHENG

School of Physical and Mathematical Sciences

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

2015

ACKNOWLEDGEMENT

First and foremost, I would like to express my sincere gratitude to my supervisor, Professor Chen Gang, for his strong support, meticulous suggestions, continuous guidance and inspiring criticisms during my graduate study in the Division of Chemistry and Biological Chemistry, School of Physical and Mathematical Science at Nanyang Technological University.

I am grateful to Dr. Gitali Devi, Dr. Ranjani Narayanan, Dr Zhou Yuan, Ms. Tong Cailing, Dr. Kiran Patil, Mr. Manikantha Maraswami, Ms. Yuan Zhen, Mr. Liu Yiran, Ms. Desiree Toh Kaixin, Ms. Yang Lixa and Mr. Shi Jiahao in our research group, for their great help during my graduate study.

I would like to acknowledge Professors Ryszard Kierzek, René Olsthoorn, Shao Fangwei, Xing Bengang, Li Tianhu and Liu Xuewei for kindly allowing my access to their instruments and providing research materials.

I am also indebted to Dr. Steven B. Smith and Mr. Tok Weelee for their assistance on the optical tweezers maintenance. I would like to express my thanks for all administrative and technical staff in our school.

Last but not least, I would like to show my deepest gratitude to my parents and my wife. I would like to dedicate this thesis to them.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENT	I
TABLE OF CONTENTS	II
ABSTRACT	VI
LIST OF ABBREVIATIONS	IV
 CHAPTER I Introduction: Optical Tweezers as a Tool for RNA Single-Molecule Force Spectroscopy Measurement	
1.1 Structural diversity and biological importance of RNA	1
1.2 Mechanical properties of RNAs are critical to biological functions	3
1.3 Introduction of single-molecule manipulation techniques based on optical tweezers	4
 CHAPTER II Methodology: Single-Molecule Force Spectroscopy Measurement Using MiniTweezers	
2.1 Design of MiniTweezers	7
2.2 Calibration of MiniTweezers	12
2.3 Measurement modes and data analysis	14
 CHAPTER III A U·U Pair to U·C Pair Mutation in a Human Telomerase RNA Hairpin Causes Structural Destabilization and Stretching Force-Induced Misfolding	
3.1 Introduction	23
3.2 Materials and methods	25
3.3 Results and Discussion	29

3.4 Conclusions	52
3.5 Perspectives	54
 CHAPTER IV Minor-Groove Base Triple Formation in SRV-1 RNA Pseudoknot Increases Mechanical Stability and –1 Ribosomal Frameshifting Efficiency	
4.1 Introduction	55
4.2 Materials and methods	59
4.3 Results and Discussion	61
4.4 Conclusions	71
4.5 Perspectives	72
 CHAPTER V Conclusions and Future Perspectives	73
REFERENCES	75
LIST OF PUBLICATIONS	88

ABSTRACT

It is not well understood how the RNA folding and unfolding dynamics, mechanical properties, and biological functions are affected by non-Watson-Crick pairs and base triples. Using a human telomerase RNA hairpin containing four consecutive non-canonical base pairs as a model system, we studied the effect of U·U pair to U·C pair substitutions on thermodynamic and mechanical properties. The ensemble thermal denaturation of the hairpins shows a 3-state pathway with an on-pathway intermediate with the disruption of the non-canonical base paired region as well as two apical Watson-Crick base pairs. In single molecule force spectroscopic studies, a misfolded hairpin at forces below 13 pN is observed. The single U·U to U·C substitution lowers the melting temperature of the first transition in thermal denaturation and significantly lowers the mechanical unfolding rupture force without obviously affecting the unfolding transition-state position. Furthermore, the substitution increases the frequency of misfolding and the mechanical folding rates, and shifts the folding transition-state position towards the native hairpin structure. Our results clearly demonstrate the non-nearest-neighbor effects of the U·U to U·C substitution.

Many RNA viruses including Simian Retrovirus type 1 (SRV-1) utilize minus-one ribosomal frameshifting as an alternative translation mechanism to generate stable ratios of structural and catalytic proteins. The detailed mechanism of minus-one ribosomal frameshifting is still not well understood. The SRV-1 frameshifting pseudoknot is located in the overlapping region of gag and pol genes and plays critical roles in stimulating frameshifting. Using optical tweezers, we studied the mechanical properties of a series of pseudoknots derived from the wildtype SRV-1 frameshifting pseudoknot. In addition, we carried out ensemble thermal denaturation and native polyacrylamide gel electrophoresis studies. Our studies reveal that the minor-groove base triple formation in the pseudoknot enhances the structural compactness and

mechanical stability. The -1 frameshifting efficiency is positively correlated with one-step unfolding rupture force and inversely correlated with the unfolding rate.

LIST OF ABBREVIATIONS

bp	Base pair(s)
CCD	Charge-coupled Device
DNA	Deoxyribonucleic acid
dsRNA	Double stranded RNA
EDTA	Ethylenediaminetetraacetic acid
nt	Nucleotide(s)
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase Chain Reaction
PD	Photodiode
pN	PicoNewton
PSD	Position sensitive detector
PBS	Polarizing beam-splitter cube
RNA	Ribonucleic acid
ssRNA	Single stranded RNA
TRIS	Tris(hydroxymethyl)aminomethane
WLC	Wormlike chain
EWLC	Extensible wormlike chain
FEC	Force versus extension curve
M	Misfolded state
S	Single stranded state
P	Pseudoknot
N	Native hairpin
NA	Numerical aperture
I	Intermediate hairpin

CHAPTER I

Introduction: Optical Tweezers as a Tool for RNA Single-Molecule Force Spectroscopy Measurement

1.1 Structural diversity and biological importance of RNA

Ribonucleic acids (RNAs) are important biomolecules that play critical roles in essential biological processes including DNA replication, RNA transcription, protein translation, RNA alternative splicing, and RNA interference.^[1] In recent years, it has been found that more than 80% of human genomic DNA sequences is transcribed to functional RNAs, while only around 2% of that directly encodes protein sequence.^[2] Besides the classical message RNA (mRNA), transfer RNA (tRNA) and ribosomal RNA (rRNA) which carry genetic information and are involved in protein synthesis, many new types of noncoding RNAs have been discovered playing important roles.^[3, 4]

RNA structure is highly hierarchical. The basic structural unit of RNA is ribonucleotide, which is composed of a heterocyclic base (purine or pyrimidine), a five-carbon sugar and a phosphate group. There are four major types of bases in RNA: guanine (G), adenine (A), cytosine (C) and uracil (U). Canonical linear RNA backbone is composed of ribose and phosphate groups, by forming 3', 5'-phosphodiester bonds. Some intronic RNAs may exit as lariat form by forming a 2', 5'-phosphodiester bond between the 5' end nucleotide and an internal adenine nucleotide. Recent studies show that circular RNAs are widely present in mammalian cells with diverse important biological functions.^[5, 6] Both the backbone and the bases may be modified during or after transcription *in vivo*. More than 100 types of natural RNA

Figure 1.1 An illustration of hierarchy of RNA structure. (A) Major-groove base triples composed of Watson-Crick base pairs and tertiary Hoogsteen base pairs. (B) Secondary and tertiary structure of an RNA pseudoknot. Watson-Crick pairs are indicated by short lines. Dashed lines and dots represent the non-Watson-Crick base pairs. (C) Three-dimensional structure of the RNA pseudoknot shown in panel B (PDB 3P22).^[18]

1.2 Mechanical properties of RNAs are critical to biological functions

Numerous studies indicate that the RNA structural stability may play key roles in its biological functions.^[15-17, 19-22] The thermodynamics of RNA structure is successfully predicted by a nearest-neighbor model.^[23-25] However, thermodynamic properties alone is not sufficient to describe the biological behavior of RNA, as recent studies have shown that mRNAs are usually less structured *in vivo* than *in vitro* at similar buffer and temperature conditions.^[26] What's more, in an ATP-depleted cell, the mRNAs are more structured than usual.^[26] Such phenomenon indicates that RNA structures *in vivo* are usually substrates of helicases and affected by other RNA binding proteins, which require ATP as energy source, so that the RNA structures *in vivo* may be far away from thermodynamic equilibrium. In those cases, the information of folding and unfolding kinetics of RNA is essential to explain and predict their biological functions. Also, helicase-facilitated RNA unfolding *in vivo* occurs at relatively constant temperature.^[19-21, 27-33] Thus, local successive unzipping of RNA structures by applying mechanical stretching force in near physiological buffer condition at constant temperature may better mimic *in vivo* RNA helicase activity than by increasing temperature or denaturant concentration.^[15, 19-22, 27, 28, 30-32, 34-38] In addition, as the end-to-end distance of the RNA molecule is a well-defined reaction coordinate, mechanical unfolding facilitates the unravelling of the unfolding and folding dynamics of complex RNA in real time.^[22, 34]

Previous single-molecule force spectroscopy researches reveal that the mechanical properties of RNA structures play direct and important roles in various biological processes, including prokaryotic intrinsic transcription termination and -1 ribosomal frameshifting.^[15, 39-42] In such cases, local structures in RNA sequence may serve as a mechanical sensor against the unfolding forces generated by helicases, and/or provide stalling force against the linear molecular motors (such as RNA polymerase and ribosome) moving along the RNA.

Previous studies also showed that temperature and salt condition can affect the mechanical properties of RNAs, which allow us to better understand how the natural RNAs respond to the environmental and metabolic stimulation by altering their mechanical stabilities and consequently regulate the relevant biological functions.^[9, 43, 44]

1.3 Introduction of single-molecule manipulation techniques based on optical tweezers

Many approaches have been applied for single-molecule manipulation. Single-molecule studies of RNA folding have been done using nanopore force spectroscopy, atomic force microscopy, single-molecule Förster resonance energy transfer (smFRET), magnetic tweezers and optical tweezers.^[45, 46]

Optical tweezers is an instrument that use highly focused laser beam(s) to create a mechanical trap that is able to stably hold a microscopic transparent particle in three dimensions. Consider one ray of light refracted by the particle, its momentum changes as the moving direction of the ray changes (Fig. 1.2). According to the law of conservation of momentum, a gradient force loaded on the particle is generated.

However in uniform optical field, the total force is zero since the momentum changes from the refraction of each light ray cancel each other out. In a non-uniform optical field, such as in a highly focused Gaussian beam, the total force loaded on the particle points to the focus of the beam, if the refractive index of the surrounding buffer is smaller than that of the particle. Besides refraction, the particle also scatters and absorbs a fraction of photons, which generates a scattering force points to the downstream of the laser focus. The scattering forces can be offset if the optical trap were generated by two well-aligned laser beams with opposite directions.

Figure 1.2 Fundamental principles of optical tweezers. (A) Momentum change of a ray of light passing through a transparent particle. (B) A transparent particle trapped by a focused Gaussian beam.

Optical tweezers allows researchers to observe and manipulate micron-size particles by applying an accurate force range from 0.1 to more than 100 pN. Also, with high resolution piezo motor stage or position sensitive detector (PSD), the position of particle can be measured with accuracy of 0.1 nm. The PSDs used in MiniTweezers are optical position sensors that can measure the position of a light spot in two dimensions on a silicon sensor surface. Optical tweezers has been successfully used in studying

single-cell and single-molecule biological systems *in vitro* and even *in vivo* in recent decades.^[9, 15, 16, 28, 34-36, 40, 41, 44, 46-49]

CHAPTER II

Methodology: Single-Molecule Force Spectroscopy

Measurement Using MiniTweezers

2.1 Design of MiniTweezers

The MiniTweezers was invented by Steven B Smith and coworkers.^[32, 50] By direct measurement of light momentum, the MiniTweezers can accurately measure trapping force without calibration of trap stiffness. The MiniTweezers has a compact design for the optical and mechanical parts, which allows it to be hanged with a bungee cord inside a soundproof box (Fig. 2.1). Such design greatly reduces the environmental noise and vibration.

Figure 2.1 The MiniTweezers system.

2.1.1 Optical path and mechanical design

The 845 nm fiber-coupled diode lasers (Lumic LU0845M150-1G06F10A) are used to generate single mode laser. The fiber tip positions are controlled by fiber wigglers (Fig. 2.2). About 5% of laser light is reflected at a pellicle (Thorlabs BP108) before the lasers interact with the trapping particle, so that trap position signals are amplified by light lever and detected by position sensitive detectors (PSD). The laser light that pass through the pellicle is reflected by a polarizing beam-splitter cube (PBS), and then circularly polarized by quarter wave plates at foci. The laser lights are focused by 60X water-immersed objective lenses to create a strong optical trap. The reflection forces of the two counter-propagating lasers on the particle are cancelled out. Then each laser light passes through a second objective lens and reflected by PBS. Then its three-dimensional pattern is detected by a PSD in x and y direction and a photodiode in z direction, so that the trapping force at each dimension is measured. A blue LED is used to generate background light and the image is recorded by a CCD camera, which is connected to a monitor for real-time control.

The chamber is mounted by a chamber frame, whose movement is driven by a motor stage (Newport 561D-XYZ) with 3 piezo motors (Thorlabs Z606). Optical trap movement in x and y direction is controlled by fiber wigglers. The two laser beams can be aligned by carefully adjusting the tilt and three-dimensional position of the fiber tips through the wigglers.

Figure 2.2 Optical path and mechanical components of Minitweezers. Here, laser is 845 nm fiber-coupled diode laser (Lumic LU0845M150-1G06F10A); pellicle is pellicle mirror (Thorlabs BP108); pbs is polarizing beam-splitter; npbs is non-polarizing beam-splitter; $\lambda/4$ is quarter-wave plate; OBJ is water immersed objective lens (Olympus UPlanApo-60x-Water); PSD is position-sensitive photo detector (UDT DL-10); PD is photodiode (UDT 10-DI); motorized stage is Newport 561D-XYZ stage.

2.1.2 Principle of force measurement

The linear momentum flux of a ray of light is $|dP/dt|=nW/c$. Here n is the index of refraction of medium, W is the power of the ray and c is the light speed. When a transparent particle with high index of refraction is trapped at a laser focus, photons are refracted and their momenta are changed. According to the law of conservation of momentum, if an external force F drags the trapped particle out of its equilibrium trap position, then the sum of the change of the light flux over all rays at the back side of the focus equals to the force, namely $\Sigma dP/dt = F$. By measuring changes of momenta of the rays coming out the back side of the trap in three dimensions, the external force F loaded on the bead is measured. If the trapped particle is static or moves in a constant

and low speed, the Stokes drag force, gravity and buoyant force on the particle are negligible, so that the external force is the tension of molecule tether, and the value of external force equals to the trap force.

In practice, a spherical particle with a relatively large radius (3 μm) is used so that the scatter light rays are kept in a narrow angle, and the application of the 2 high numerical aperture objective lenses (NA=1.2) leads to a light collection efficiency of 99%, as high NA lens has a large collection angle.^[50]

As shown in Fig. 2.3, PSDs measure the external force component in directions perpendicular to the optic axis (x and y direction), according to equation 2.1 and 2.2,^[50]

$$F_x = \frac{D_x R_D}{c \Psi_x R_L} = s_x D_x \quad (2.1)$$

$$F_y = \frac{D_y R_D}{c \Psi_y R_L} = s_y D_y \quad (2.2)$$

Here, F_x and F_y are external forces in x and y direction, respectively, c is the speed of light, R_L is the focal length of the lens, R_D is the half-width of the PSD, Ψ_x and Ψ_y are the responsivities of the PSD in x and y direction, and D_x and D_y are the electronic difference signals in x and y direction, respectively. As c , R_L , R_D , Ψ_x and Ψ_y are constant, they can be replaced by amplification factors s_x and s_y in the equations. The amplification factors are measured by force calibration.

Similarly, PDs measure the component in direction parallel to the optic axis (z direction) based on equation 2.3 (Fig 2.3),

$$F_z = s_z D_z \quad (2.3)$$

Where F_z is external force in z direction, D_z is the electronic difference signals in z, and s_z is the amplification factor in z direction.

Figure 2.3 (A) Laser beam passes through the first objective lens and traps a particle at focus. Then the beam is collimated by a second objective lens and projected onto a detector surface. The regions that rays are able to be collected by objective lenses are indicated by dashed lines. The solid lines indicate the edges of the focused laser beam. (B,C) When an external force (red arrow) perpendicular to the optic axis applies to the particle, the particle moves slightly out of the trap center so that a trap force generates to balance the external force, as well as the rays leaving the beam deflects. The deflection is transformed by the second objective lens, and the displacement of beam spot is detected by the PSD. (D,E) External force parallel to the axis changes the area of light spot as well as the intensity projected on the photodiode detector. Force pushing particle towards the first objective lens expands the light spot on detector, and force pushing particle towards the second objective lens concentrates the light spot on the detector. Figures are reproduced from S. B. Smith et al.^[51]

2.1.3 Instrument Performance

The theoretical force resolution is determined by the single-bit resolution of PSD (0.013 pN). The single-bit resolution of PSD and the trap stiffness (typically ~ 0.1 pN/nm) determine that the theoretical spatial resolution of MiniTweezers is 0.1 nm.^[50] However, the experimental spatial resolution, limited by Brownian noise and sampling bandwidth, is about 2 nm at 1000 Hz. Similarly, the actual force resolution is about 0.1 pN, limited by the Brownian noise. The Brownian noise is introduced through thermal fluctuation of the trapped bead, the handles of the single-molecule construct and the micropipette. Wen et al discussed how the handle length of the RNA/DNA construct affects the signal-noise ratios.^[28] A longer and softer handle introduces larger thermal fluctuation than a shorter and stiffer one.

The Minitweezers can directly load and measure a force on the trapped particle ranging from 0.1 pN to 150 pN, which is sufficient to break the noncovalent interactions within and between nucleic acids and proteins. The maximum holding force is determined by the laser power used (100 mW), the numerical aperture of objective lenses, and the trapped particle diameter (3 μm).^[51]

2.2 Calibration of MiniTweezers

2.2.1 Force calibration

Several methods are used to calibrate the force measurement to obtain the amplification factors in equation 2.1 to 2.3. The first method is Stokes test.^[52] In this method, a spherical particle with known radius is trapped by optical tweezers. By moving the optical trap, the trapped particle moves in the still water inside the chamber,

and both the trap force and a Stokes drag force are loaded on the particle in opposite direction. According to Stokes' Law (equation 2.4),

$$F_d = -6\pi\mu r v \quad (2.4)$$

Here F_d is the drag force, μ is the dynamic viscosity of water at a certain temperature, r is the radius of the particle, and v is the velocity of the particle. If the particle moves in a constant speed, the value of drag force equals to trap force, and we can calculate the radius of the particle. By comparing the calculated radius to the known radius, we can briefly calibrate the measured trap force by adjust the force sensitivities for each direction per laser. The relative error of this method is about 10%.

The second method is stretching long dsDNA.^[53] The transition force range of overstretching long dsDNA with known GC content is narrow and invariant in certain buffer conditions. For example, lambda DNA is overstretched at a force range from 65 to 70 pN, in a buffer with 0.5 M NaCl, 10 mM Tris, 1 mM EDTA, pH 7.5. This method is used to test whether the force is well calibrated.

2.2.2 Distance calibration

Once the force measurement is calibrated, distance calibration can be carried out. During distance calibration, a particle is held by the micropipette as well as optical trap with constant force feedback. Under constant-force feedback mode, the distance between trap center and particle center is constant. The particle moves in x and y direction driven by motors for a known displacement s with vector components s_x and s_y recorded. As the feedback mode is on, the optical trap automatically traces the particle and also has the same displacement. As the movement of optical trap in x and y direction is driven by bending the fiber tip, displacement l of laser focus with vector components l_x and l_y is also recorded by PSD with light lever amplification for each

laser beam. By adjusting the x and y light lever sensitivities per laser, the two displacements should be the same, so that the distance measurement is calibrated.

2.3 Measurement modes and data analysis

2.3.1 Experimental setup

Typically, molecular biological methods are applied to synthesize an RNA/DNA hybrid construct with the RNA structure to be studied, flanked by two double helix handles with biotin and digoxigenin labels at the two terminal ends. The synthesis of the RNA constructs will be described in Chapter 3 and Chapter 4.

During the single-molecule optical tweezers experiment, the biotin group and the digoxigenin group of the construct bind specifically to a streptavidin coated $\sim 1.8\ \mu\text{m}$ particle and an anti-digoxigenin antibody coated $\sim 3\ \mu\text{m}$ particle, respectively (Fig. 2.4). Firstly, the construct is incubated with anti-digoxigenin beads for about 5 minutes at room temperature. Then the mixture is diluted with buffer and injected into the chamber. The biotin group construct is then attached to a streptavidin bead with the aid of optical trap. During the optical tweezers experiment the streptavidin bead is held by micropipette and the anti-digoxigenin bead is held by optical trap (Fig. 2.4).

Figure 2.4 Experimental setup of single-molecule force spectroscopic experiments on RNA using optical tweezers. The drawing is not to scale.

2.3.2 Mechanical behaviors of semi-flexible polymers

For semi-flexible biopolymers including peptide, ssRNA, dsRNA, dsDNA and RNA/DNA hybrid, their end-to-end distance (ΔX) at low force region is well predicted by the worm-like chain (WLC) model (equation 2.5).^[37, 54, 55]

$$F = \frac{k_B T}{P} \left[\frac{1}{4 \left(1 - \frac{\Delta X}{L_0} \right)^2} - \frac{1}{4} + \frac{\Delta X}{L_0} \right] \quad (2.5)$$

Here, k_B is Boltzmann constant, T is absolute temperature, P is persistence length, and L_0 is contour length.

As it is assumed that the covalent bonds of the polymer chain do not change under force for the WLC model, the extensions predicted at high force region is significantly shorter than experiment results.^[34, 53] Taken bond enthalpy change into consideration, an extensible worm-like chain (EWLC) model better fits the experimental results at high force region (equation 2.6).^[56]

$$F = \frac{k_B T}{P} \left[\frac{1}{4 \left(1 - \frac{\Delta X}{L_0} + \frac{F}{K} \right)^2} - \frac{1}{4} + \frac{\Delta X}{L_0} - \frac{F}{K} \right] \quad (2.6)$$

Here K is stretch modulus. For ssRNA, L_0 , P and K are 0.59 nm per nucleotide, 1 nm, and 1500 pN, respectively.^[34, 37, 54, 55, 57] For the RNA/DNA handle, L_0 is 0.3 nm per base pair, and P and K are acquired from least-square fitting.^[34]

2.3.3 Trap stiffness measurement

The trap stiffness is linearly correlated to the laser power used. Other factors contribute to the trap stiffness include diameter of trapped particle, refraction indices of liquid medium and particle, and numerical aperture of objective lens.^[51]

The light lever-coupled PSDs are used to measure the displacement of optical trap X_{trap} . The displacement of the trapped particle is calculated using equation 2.7,

$$X_{particle} = X_{trap} - \frac{F}{\epsilon} \quad (2.7)$$

Here F is the trap force and ϵ is the trap stiffness. For a tethered particle (Fig. 2.4), as the polystyrene particles used in the experiments are rigid, and the pipette position is constant during data acquisition, the extension change of the molecular tether equals to the displacement of the trapped particle. Thus, to calculate the accurate extension of the tether, we need to measure the trap stiffness.

For dragging a tethered particle using optical trap, the slope of the force-position curve of the particle is the stiffness of the trap spring and the molecule tether spring. As the trap stiffness is dependent on both the laser trap and the trapped particle, to obtain the stiffness of the trap alone *in situ*, a method is applied based on the design that the optical trap of MiniTweezers is created by dual counter-propagating lasers.

For a tethered particle trapped by the lasers, changing the trap position will change the length of the molecular tether, as well as the tether force loading on the trapped particle, which drags the particle away from the trap center with an unknown distance. However, for a dual trap, the effective trap force is the sum of forces from trap A and trap B. If one trap moves toward the pipette and the other trap moves in the opposite direction, while effective trap force is kept constant and the average particle position remains fixed. In such case, the length of the tether molecule is invariant, so that the stiffness for each trap can be calculated using the displacement of each trap measured by light levers and the force change for each trap measured by PDs and PSDs (Fig. 2.5). The sum of the two stiffness values is the effective stiffness of the dual trap.

Figure 2.5 Force versus time and position versus time traces for each trap in trap stiffness measurement. The dashed lines indicate the maximum and minimum forces and positions for each trap during trap stiffness measurement.

2.3.4 Force-clamp and position-clamp experiments

In constant-force experiments, the force is kept constant by an electronic feedback,^[28, 35, 36] so that the distance between the trap center and the trapped particle is constant. The extension increases when the RNA molecule unfolds, and decreases when it refolds. The change of the extension is only corresponding to the transitions of the (un)folding of the RNA structure to be investigated. For RNA structures with large hysteresis between folding and unfolding pathways, force-jump methods is applied to unfold and refold the RNA repeatedly in two separate constant forces, respectively.^[35, 36]

In constant-position experiments, the trap position is passively kept constant without electronic feedback signals,^[58] so the trapped particles moves towards the trap center when the RNA unfolds, resulting in the trapped force decreasing, and moves away from the trap center when the RNA folds, resulting in the trapped force increasing. In constant-position experiments the change of the extension is corresponding to both the transitions of the (un)folding of the RNA structure and the change of length of handles.

In constant-force experiments, for extension versus time traces at a certain force, the extensions at unfolded and folded states are determined by Gaussian fitting of extension versus time traces,^[28]

$$y_X = a_1 e^{-\left(\frac{X-X_F}{c_1}\right)^2} + a_2 e^{-\left(\frac{X-X_U}{c_2}\right)^2}, \quad (2.8)$$

Here y_X is the frequency of each binned extension X ; a_1 and a_2 are amplitudes of the peaks; c_1 and c_2 are width of the peaks; X_U and X_F are the extensions at unfolded and folded state, respectively.

For the standard deviation of the extension at each state,

$$\delta X_F = \frac{c_1}{\sqrt{2}} \quad (2.9)$$

$$\delta X_U = \frac{c_2}{\sqrt{2}} \quad (2.10)$$

For the standard deviation of the change of extension between two states, the difference between two neighbor extension data points X' are fitted to a single Gaussian distribution,

$$y = ae^{-\left(\frac{X'-b}{c}\right)^2} \quad (2.11)$$

Here y is the frequency of each binned X' ; a is the amplitude of the peak; c is the width of the peak; b is the average of all X' , respectively. The standard deviation $\delta\Delta X$, is calculated from

$$\delta\Delta X = \frac{c}{2} \quad (2.12)$$

Similarly, in constant-position experiments, the forces at folded and unfolded states are determined by Gaussian fitting of force versus time traces,^[28]

$$y_F = a_1 e^{-\left(\frac{F-F_U}{c_1}\right)^2} + a_2 e^{-\left(\frac{F-F_F}{c_2}\right)^2}, \quad (2.13)$$

Here y_F is the frequency of each binned force F , and F_U and F_F are the forces at unfolded and folded state, respectively.

A hidden Markov modeling algorithm is applied to passive mode force versus time traces to get the lifetime of each unfolded and folded state.^[59] Then the sorted lifetimes are fitted using single exponential decay curve to get the force dependent rate constant,

$$P = e^{-kt} \quad (2.14)$$

where P is the probability that the unfolded or folded state lasts to time t , and k is the rate coefficient.

The force dependent rate coefficient is fitted to the Bell model, ^[28, 60]

$$k(F) = k_m k(0) e^{\left(\frac{FX^\ddagger}{k_B T}\right)} \quad (2.15)$$

where k_m is the instrumental factors to rate coefficient, $k(0)$ is the rate coefficient at zero force, $X^\ddagger$ is the distance from unfolded or folded state to the transition state, k_B is Boltzmann constant, and T is absolute temperature.

The Bell model assumes a constant $X^\ddagger$ at different force. For reactions with significant Hammond effect, namely the transition state position significantly shifts towards the unfolded state with force increasing, ^[61] more complex fitting models are needed. ^[62-64]

With the assumption that unfolding reaction can be described as escape from a deep one-dimensional free-energy well, as well as the energy barrier is with linear-cubic or cusp free-energy surfaces, the force dependent kinetic rates can be described as

$$k(F) = k(0) \left(1 - \frac{\Delta X^\ddagger F \mu}{\Delta G^\ddagger}\right)^{\frac{1}{\mu}-1} e^{\left(\frac{\Delta G^\ddagger}{k_B T}\right) \left[1 - \left(1 - \frac{\Delta X^\ddagger F \mu}{\Delta G^\ddagger}\right)^{\frac{1}{\mu}}\right]} \quad (2.16)$$

Here $\Delta X^\ddagger$ is extension change between reaction initial state and transition state, $\Delta G^\ddagger$ is activation energy, $k(0)$ is the unfolding kinetic rate at 0 force, k_B is Boltzmann constant and T is absolute temperature. For linear-cubic barrier μ equals to 2/3, and for cusp-like barrier μ equals to 1/2. If μ equals to 1, the equation recovers to Bell model.

The critical force F_c is the force at which unfolded and folded state are equally distributed. The critical force is obtained from the cross point of the Bell fittings for both unfolding and refolding kinetics. The free energy G_0 of unfolding a single molecule RNA structure without external force is calculated from

$$G_o = F_c \cdot \Delta X(F_c) \cdot N_A - W_{EWLC}(F_c) \cdot N_A \quad (2.17)$$

Here N_A is Avogadro Constant, $\Delta X(F_c)$ is extension change at critical force in constant force experiments, and $W_{EWLC}(F_c)$ is the work of stretching the unfolded ssRNA to the critical force, by numerical integration of the EWLC equation.

2.3.5 Force-ramp experiment

If the optical trap moves in a constant speed periodically, and the trap stiffness is approximately constant, the force loading rate on the RNA structure is also approximately constant. The RNA structure unfolds during the force increase and refolds during the force decrease. The forces at which the RNA structure starts to unfold and refold are called rupture force and folding force, respectively.

The total number of nucleotides folded or disrupted for each transition is calculated according to the extension change (ΔX) values at the folding/rupture force using EWLC model, with the length of the initial and terminal state of the transitions taken into consideration, which may be the diameter of hairpin stem (2 nm) or the end-to-end distance of a pseudoknot.

The unfolding or refolding rates can be extracted from the unfolding or refolding force distribution, respectively, using Szabo-Hummer-Dudko method.^[62] For unfolding force distribution,

$$k(F) = \frac{\dot{F}(F)p(F)}{\int_F^\infty p(f)df} = \frac{h_k \dot{F}(F_0 + (k-1/2)\Delta F)}{(h_k/2 + \sum_{i=k+1}^N h_i)\Delta F} \quad (2.18)$$

Here p is the probability density, and $\dot{F}(F)$ is force loading rate of the system at force F , h_k is the height of k th bin in the unfolding force distribution histogram that start at F_0 , ΔF is the bin size of the histogram, i and k are the bin numbers, and N is the total bin number of the histogram. Poisson noise is used to estimate the standard deviation of

data in the unfolding force distribution histogram.^[65, 66] For a column in a histogram with a data number of X , the Poisson noise of the data number is $X^{1/2}$, as an estimation of standard deviation. Then the standard deviations of the force dependent unfolding rates are estimated based on the Poisson noises and the standard deviation of force loading rate through error propagation.

The free energy of RNA (un)folding can be extracted from the FECs using Crooks fluctuation theorem.^[44] The works for folding and unfolding are measured by numerical integration of the FECs. The integration ranges are chosen so that the work of stretching native hairpin to ssRNA is measured. EWLC model can be used to fit the FECs for the RNA/DNA handle. The traces significantly deviating from EWLC model may be discarded. The stretching work for the RNA/DNA handle is extracted by numerically integrating the EWLC fitting curves. The ssRNA stretching work is calculated based on the EWLC model. A maximum likelihood estimator described by Shirts et al can be used to obtain the free energies and the statistical errors.^[67]

CHAPTER III

A U·U Pair to U·C Pair Mutation in a Human Telomerase RNA Hairpin Causes Structural Destabilization and Stretching Force-Induced Misfolding

3.1 Introduction

The diverse structures and dynamic properties of RNAs determine their biological functions. An RNA secondary structure usually consists of various loops and double-stranded helices,^[17, 68, 69] which is well predicted by a nearest-neighbor thermodynamic model based on the stacking of each of the adjacent base pairs.^[25, 69, 70]

Non-canonical pairs (e.g., U·U, U·C, and G·A pairs) in RNA loop and stem regions stabilize or destabilize RNA secondary structures, and contribute to tertiary structures formation, which provides specific structure features for ligand binding.^[14, 17, 24, 27, 31, 37, 71-73] Ensemble thermal melting experiments of short RNAs have been performed to extract the nearest-neighbor thermodynamic parameters. The current nearest-neighbor model only considers the non-canonical base pairs adjacent to the stems for the prediction of RNA secondary structure stability.^[69] It is not well understood that how the non-canonical pairs not adjacent to a stem affect RNA folding thermodynamics and kinetics, however.^[72, 74-80]

Consecutive U·U and U·C base pairs exist in many functional RNAs.^[14, 18, 80-86] Previous NMR and X-ray crystallography studies^[80-83, 87-90] show that a U·U pair typically forms two hydrogen bonds between the bases, while a U·C pair usually forms only one hydrogen bond. U·U and U·C base pairs may also be stabilized by water bridge formation. As revealed by previous ensemble optical melting studies, when adjacent to a stem, a U·U pair is more stable than a U·C pair by about 0.5-1 kcal/mol of change in free energy.^[24, 71, 74, 80, 91-94] However, how U·U and U·C base pairs affect the (un)folding dynamics and mechanical properties of an RNA structure is still not clear.

Single-molecule force spectroscopy based on optical tweezers has been used to investigate the kinetic and thermodynamic features of RNA mechanical folding and unfolding in near-physiological solutions at constant temperature.^[15, 27, 28, 31, 34-38] Here, using both temperature dependent ultraviolet absorbance spectroscopy and dual-laser optical tweezers, we studied the effect of U·U to U·C substitution on the unfolding and folding dynamics and kinetics of a hairpin truncated from human telomerase RNA (Fig. 3.1).^[80] Our results reveal that a single U·U to U·C pair substitution affects the stability and folding dynamics of the native hairpin.

Figure 3.1 (A) RNA constructs. Hairpin 1 is truncated from human telomerase RNA.^[80]

The mutated nucleotides in hairpins 2 and 3 are highlighted by the red squares. (B)

U·U and U·C pairs from the three dimensional structure of hairpin 1(PDB: 1NA2). (C) Experimental setup for single molecule force spectroscopic experiments on the RNA hairpins (not to scale). The RNA construct with a hairpin of interest flanked by two DNA/RNA hybrid handles were attached to two polystyrene beads through non-covalent interactions between biotin-streptavidin and digoxigenin-anti-digoxigenin, respectively. The streptavidin-coated bead was attached to the tip of a micropipette by suction and the anti-digoxigenin-coated bead was held by the optical trap. The three-dimensional structure of hairpin 1 (PDB: 1NA2) is shown, with the U·U and U·C pairs shown in blue color.

3.2 Materials and methods

3.2.1 Thermal denaturation

HPLC purified RNA hairpin samples for the thermal denaturation experiments were purchased from Sigma. The sequences of 29-nt native hairpins for thermal melting are 5'-GGGCUGUUUUUCUCGCUGACUUUCAGCCC-3' (hairpin 1), 5'-GGGCUGUUUUUCUCGCUGACUCCAGCCC-3' (hairpin 2), and 5'-GGGCUGUCUUUCUCGCUGACUCCAGCCC-3' (hairpin 3). The sequences of proposed 15-nt misfolded hairpins are 5'-CGCUGACUUUCAGCC-3' (misfolded 5-bp triplloop hairpin 1) and 5'-CGCUGACUCCAGCC-3' (misfolded 4-bp pentaploop hairpin 2), which are truncated from the native hairpins (residues 14 to 28). The sequences of HPLC purified RNA duplexes (SABio) are listed in Table 2, which contain same 4×4 non-canonical base pairs as the corresponding hairpins. All the RNA samples were verified by mass spectroscopy by the companies.

A Shimadzu 2550 spectrometer were used to conduct the RNA ensemble thermal melting experiments with a temperature ramping rate of 0.5 °C/min. Both the

absorptions at both 260 and 280 nm at different temperatures were recorded simultaneously. 5 μ M native hairpins or 10 μ M misfolded hairpins were melted at buffer with 200 mM or 1 M NaCl, 20 mM HEPES, 0.1 mM EDTA, pH 7.3, respectively. Before the thermal denaturation experiments, all hairpins were heated at 95 $^{\circ}$ C, then slow-cooled to 20 $^{\circ}$ C in more than 2 hours. For each condition, the measurements were repeated 3 times.

The first derivative of absorbance at 260 nm with respect to temperature smoothed with a 5 $^{\circ}$ C window using an averaging method are used to extract thermodynamic parameters using an algorithm developed by previous studies.^[95-98] The algorithm is briefly described below.

For thermal melting curves with multiple transitions, with the two-state assumption, the equilibrium constant of *i*th transition K_i is determined by Van't Hoff equation,

$$K_i = e^{\frac{\Delta H_i^{\circ} \left(\frac{1}{T_{M,i}} - \frac{1}{T} \right)}{R}} \quad (3.1)$$

Here, $T_{M,i}$ and ΔH_i° are the melting temperature in Kelvin and the unfolding enthalpy in kcal/mol, respectively. T is thermodynamic temperature and R is the ideal gas constant. For a melting curve with only two transitions, ignoring the baseline effects, the observed absorbance A , is calculated from

$$A = \frac{A_1 K_1 + (A_1 + A_2) K_1 K_2}{q} \quad (3.2)$$

Here A_1 and A_2 are the fractional hyperchromicities of each transition, the partition function q is determined by

$$q = 1 + K_1 + K_1 K_2 \quad (3.3)$$

For two-transition curves, the first derivatives of the absorbance to temperature

$$\frac{\partial A_t}{\partial T} = \frac{A_1(K_1\Delta H^\circ_1 + (\Delta H^\circ_1 + \Delta H^\circ_2)K_1K_2) + A_2(\Delta H^\circ_1 + \Delta H^\circ_2)K_1K_2}{qT^2R} - \frac{(A_1(K_1 + K_1K_2) + A_2K_1K_2)(\Delta H^\circ_1K_1 + (\Delta H^\circ_1 + \Delta H^\circ_2)K_1K_2)}{q^2T^2R} \quad (3.4)$$

The parameters A_1 , A_2 , ΔH°_1 , ΔH°_2 , $T_{M,1}$ and $T_{M,2}$ are extracted using least-square fitting method. For the i^{th} transition, the entropy and temperature-dependent free energy changes ΔS°_i and $\Delta G^\circ_i(T)$ are calculated from

$$\Delta S^\circ_i = \frac{\Delta H^\circ_i}{T_{M,i}} \quad (3.5)$$

$$\Delta G^\circ_i(T) = \Delta H^\circ_i - T\Delta S^\circ_i \quad (3.6)$$

The total change of enthalpy ΔH° and total change of free energy $\Delta G^\circ(T)$ are calculated from

$$\Delta H^\circ = \Delta H^\circ_1 + \Delta H^\circ_2 \quad (3.7)$$

$$\Delta G^\circ(T) = \Delta G^\circ_1(T) + \Delta G^\circ_2(T) \quad (3.8)$$

For the duplexes and misfolded hairpin melting curves, MeltWin 3.5 program^[99] was used to extract thermodynamic parameters according to two-state model. In detail, the two-state model was applied to the melting curve (normalized absorption versus temperature) at each sample concentrations to extract the thermodynamic parameters including melting temperature and enthalpy. Then together with effective concentrations, the entropy and free energy were calculated using van't Hoff equation (equation 3.9). For duplexes, the melting temperatures at different effective concentrations were also fit using the van't Hoff equation (equation 3.9) to achieve the thermodynamic parameters.^[25, 99, 100]

$$\frac{1}{T_M} = \frac{R}{\Delta H^\circ} \ln \frac{C_T}{4} + \frac{\Delta S^\circ}{\Delta H^\circ} \quad (3.9)$$

Here C_T is the total RNA strand concentration.

3.2.2 Optical tweezers experiments

The optical tweezers experiments were carried out using a high resolution dual-laser optical tweezers with similar design to Minitweezers.^[53] The laser powers used for both lasers are 200 mW.

Single-molecule constructs were made as described previously.^[15, 35] Briefly, the recombinant plasmids were constructed by inserting chemically synthesized DNAs into p2luc vector^[101]. The ~1200-nt RNA molecules were made by PCR amplification of the plasmids followed by in vitro transcription by T7 RNA polymerase. The RNAs contain the 29-nt native hairpin sequences, which are flanked by ~600-nt sequences. The RNAs were annealed with complementary strands of PCR-generated ~600-bp dsDNAs to generate RNA/DNA hybrid handle A and handle B, respectively. The handle A was labelled at the 3' end of the DNA strand by introducing biotin-modified dUTP using T4 DNA polymerase. Handle B was labelled at the 5' end of the DNA strand by digoxigenin labeled primer during PCR.

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.^[15, 34, 35] The micropipette was connected to a piezoelectric stage. In a force-ramp experiment, the force was changed by moving the piezo stage at 100 nm/s, which resulted in a force-ramp rate of about 11 pN/s (with an approximately constant stiffness of 0.11 pN/nm). The force was maintained by an electronic force feedback in constant-force experiments. The data acquisition rate was 100 Hz, which sets the limit of the time resolution to be 10 ms. For the constant-force experiments, we chose the forces so that

the measured lifetimes are longer than 0.05 s, which is about the electronic feedback response time^[28, 35]. All of the optical tweezers experiments were carried out at 22 ± 1 °C (mean $\pm$ limit error). The buffer contains 200 mM NaCl, 10 mM Tris-HCl, 0.1 mM EDTA, pH 7.3. For each hairpin, data were collected for at least two different molecular tethers.

3.3 Results and Discussion

3.3.1 A reversible pathway with two transitions and the destabilizing effect of a U·U pair substituted to a U·C pair are revealed by ensemble thermal denaturation experiments

According to NMR studies, the structure of a 29-nucleotide (nt) hairpin (hairpin 1, Fig. 3.1) truncated from human telomerase RNA contains a 4 consequential non-canonical base pairs region with one U·C pair and three U·U pairs and 8 Watson-Crick pairs.^[80] Each U·U pair has 2 hydrogen bonds, while the U·C pair has 1 hydrogen bond between amine of cytosine and O4 of uracil, as well as a potential water-bridged hydrogen bond between N3 of cytosine and H3 of uracil.^[80] In this study we conduct ensemble thermal denaturation experiments for 3 hairpins (Fig. 3.1). The U7·U23 pair in hairpin 1 is substituted to a U7·C23 pair in hairpin 2, and the U8·U22 in hairpin 2 is substituted to a C8·U22 pair in hairpin 3.

For all the RNA hairpin constructs, the heating and cooling curves overlapped and did not show obvious hysteresis, which suggest that thermal unfolding and folding reactions are in equilibrium. Two thermal melting transitions are observed in experiments for all three native hairpins, indicating a three-state unfolding pathway (Fig. 3.2 and Table 3.1). Such a thermal melting pattern is consistent with previous thermal denaturation studies and the temperature-dependent imino proton NMR spectra

studies on the wild-type RNA hairpin.^[80, 102] In buffer with either 200 mM or 1 M NaCl, the U·U to U·C substitutions significantly lower T_{m1} but do not apparently affect T_{m2} .

Figure 3.2 UV absorptions at 260 nm for the native hairpins at 200 mM NaCl during thermal denaturation experiments. (A) Normalized absorption trajectories of hairpins 1, 2 and 3. Heating and cooling trajectories are overlapped. Only the heating trajectories are shown. (B) First derivatives of the heating trajectories for hairpins 1, 2 and 3, respectively. A similar thermal denaturation pathway was observed for hairpin 1 in previous studies.^[80, 102]

For all the native hairpins per buffer, the hyperchromicity of the second transition is significantly higher at 280 nm than at 260 nm, indicating the melting of a G-C base pair-rich region (Fig. 3.3).^[80, 103] All of these evidences suggest that the first transition from native hairpin to an intermediate structure (N-I) is caused by the cooperative melting of the 4 sequential non-canonical base pairs additional with the two apical Watson-Crick base pairs, and the second transition from the intermediate structure to completely unfolded RNA (ssRNA, I-U) is caused by the melting of the 6-bp G-C rich stem at the bottom.

Figure 3.3 UV absorptions at 280 nm for native hairpins at 200 mM NaCl during thermal denaturation experiments. (A) Normalized UV melting trajectories of hairpins 1 (red), 2 (green) and 3 (blue), respectively, recorded at 280 nm. Heating and cooling trajectories are overlapped. Only the heating trajectories are shown. (B) A comparison of first derivative curves of hairpin 2 at 260 nm (open triangles) to 280 nm (solid triangles), respectively. The solid lines are the fitting curves according to equation 3.4. A relatively higher hyperchromicity (with a larger dA/dT value) is observed at 280 nm than at 260 nm for the second transition involving the disruption of the G-C rich bottom stem. The first transition involves the disruption of the U-rich internal loop.

The U·U to U·C substitution lowers the change of free energy values for the native hairpin to single stranded RNA, which is consistent with previous denaturation studies for smaller non-canonical U·U/U·C base pair regions.^[24, 71, 74, 91-94] The U·U to U·C substitutions also broaden the first transition with the hyperchromicity decreasing. Upon the U·U to U·C substitution, the free energy values of either transition change (Table 3.1), implying that the two transitions may not be independent of each other or well defined, and indicating that the thermal denaturation of the U·U/U·C pairs may be coupled to the two flanking stems. Such results reveal the non-nearest-neighbour effects of U·U and U·C pairs on the structural stability of the hairpins.

Table 3.1 Thermodynamic parameters obtained for denaturation of hairpins at 260 nm in 200 mM or 1M NaCl, 20 mM HEPES, 0.1 mM EDTA, pH 7.3.

RNA molecule	reaction	200 mM NaCl			1 M NaCl			$-\Delta G^\circ$ (37 °C predicted) ^a (kcal/mol)
		$-\Delta H^\circ$ (kcal/mol)	T_m (°C)	$-\Delta G^\circ$ (22 °C) (kcal/mol)	$-\Delta H^\circ$ (kcal/mol)	T_m (°C)	$-\Delta G^\circ$ (37 °C) (kcal/mol)	
Hairpin 1	N-I	30.2 ± 0.2	50.3 ± 0.1	2.6 ± 0.1	31.5 ± 6.4	54.0 ± 1.0	1.4 ± 0.1	1.9
	I-U	61.3 ± 2.3	72.1 ± 0.3	8.9 ± 0.3	71.3 ± 9.2	77.4 ± 1.6	8.3 ± 0.7	8.7
	Sum	91.5 ± 2.1	-	11.5 ± 0.3	102.8 ± 5.4	-	9.7 ± 0.6	10.6
Hairpin 2	N-I	32.3 ± 0.6	45.6 ± 0.4	2.4 ± 0.1	30.7 ± 0.7	49.2 ± 0.8	1.2 ± 0.1	1.9
	I-U	55.3 ± 0.8	70.3 ± 0.6	7.8 ± 0.2	62.9 ± 3.4	76.4 ± 1.8	7.1 ± 0.3	8.1
	sum	87.6 ± 0.7	-	10.2 ± 0.2	93.6 ± 3.4	-	8.3 ± 0.3	10.0
Hairpin 3	N-I	21.9 ± 0.6	40.7 ± 0.7	1.3 ± 0.1	21.2 ± 0.3	43.9 ± 0.5	0.5 ± 0.1	1.9
	I-U	48.2 ± 3.4	69.8 ± 1.1	6.7 ± 0.5	50.7 ± 5.2	76.5 ± 1.8	5.7 ± 0.4	8.1
	sum	70.1 ± 3.5	-	8.0 ± 0.5	71.9 ± 4.9	-	6.2 ± 0.4	10.0
Misfolded hairpin 1	M-U	46.3 ± 0.7	67.7 ± 0.1	6.2 ± 0.1	51.2 ± 1.9	70.0 ± 0.8	4.9 ± 0.2	5.2
Misfolded hairpin 2	M-U	41.5 ± 0.3	65.5 ± 0.5	5.3 ± 0.1	42.0 ± 1.6	67.3 ± 0.6	3.7 ± 0.1	4.5

^a The values of free energies change in 1 M NaCl at pH 7 at 37 °C were predicted by RNAstructure 5.3^[69]. A previous thermal melting study reveals a free energy value of -9.8 kcal/mol and enthalpy value of -92.0 kcal/mol for hairpin 1 at 1 M KCl.^[80] The transitions N-I and I-U are for the melting of the native hairpin to the intermediate structure, and the intermediate structure to single-stranded random coil, respectively (Fig. 3.10C). The misfolded hairpin constructs were made based on the proposed misfolded intermediates (Fig. 3B). All the experimental data are reported as mean ± standard deviation from 3 repeated experiments.

3.3.2 The destabilizing effect of replacing U-U base pairs to U-C base pairs are observed in ensemble thermal denaturation of the duplex constructs

We also constructed three duplexes with the same four sequential non-canonical base pairs for ensemble thermal denaturation studies. For all the three duplexes, the heating and cooling curves are overlapped and do not show obvious hysteresis, indicating that thermal unfolding and folding reactions are in equilibrium. Only one transition is observed for each duplex, implying the denaturation of the four sequential non-canonical base pairs together with the two flanking canonical base paired stems are cooperative.^[23, 75, 104]

Fitting with both the van't Hoff equation and the two-state model, we extracted the thermodynamic parameters for the duplexes at 200 mM and 1 M NaCl, respectively (Table 3.2).^[25, 99, 100]

Table 3.2 Folding thermodynamic parameters extracted from thermal denaturation experiments for the duplexes, recorded at 260 nm in 200 mM NaCl or 1 M NaCl, 20 mM HEPES, 0.1 mM EDTA, pH 7.3.

Sequence	NaCl concentration (mM)	T_M^{-1} versus $\ln(C_T/4)$ plots				Average of melting curve fits			
		$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (eu)	$-\Delta G^\circ$ (kcal/mol)	T_m (°C) ^a	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (eu)	$-\Delta G^\circ$ (kcal/mol)	T_m (°C) ^a
Duplex 1 GCUGUUUUUCC GA-3' ACGACUUUCAG GC-5'	200	97.0 ± 1.8	284.0 ± 5.8	9.0 ± 0.1	44.9	93.9 ± 4.6	274.0 ± 14.5	8.9 ± 0.1	45.2
	1000	115.8 ± 6.5	335.3 ± 20.1	11.8 ± 0.3	51.9	95.7 ± 3.0	272.8 ± 9.6	11.1 ± 0.1	52.4
	200	97.7 ± 6.6	289.3 ± 21.3	8.0 ± 0.1	41.6	85.4 ± 3.6	249.6 ± 11.3	8.0 ± 0.1	42.3
	1000	103.6 ± 2.9	301.2 ± 9.1	10.1 ± 0.1	48.2	92.5 ± 3.9	266.3 ± 12.4	9.9 ± 0.1	48.9
Duplex 2 GCUGUUUUUCC GA-3' ACGACCUUCAG GC-5'	200	78.6 ± 1.7	230.9 ± 5.7	7.0 ± 0.1	38.9	64.5 ± 1.6	184.8 ± 5.5	7.2 ± 0.1	40.2
	1000	61.6 ± 7.9	171.0 ± 25.2	8.5 ± 0.2	47.4	70.1 ± 5.7	198.3 ± 17.8	8.6 ± 0.2	46.5
	200	78.6 ± 1.7	230.9 ± 5.7	7.0 ± 0.1	38.9	64.5 ± 1.6	184.8 ± 5.5	7.2 ± 0.1	40.2
	1000	61.6 ± 7.9	171.0 ± 25.2	8.5 ± 0.2	47.4	70.1 ± 5.7	198.3 ± 17.8	8.6 ± 0.2	46.5

^aThe T_m in °C is calculated at a total RNA strand concentration (C_T) of 100 μ M. T_m is the melting temperature in absolute temperature. The ΔG° values are folding free

energies at 37 °C. The ΔG° values predicted by RNAstructure 5.3 program^[69] are –12.0, –11.4 and –11.4 kcal/mol for duplexes 1, 2 and 3, respectively, in a buffer condition of 1M NaCl at pH 7 at 37 °C.

Each U·U-to-U·C substitution destabilizes the duplex structure by 1-2 kcal/mol at 37 °C, which is consistent with the thermal denaturation results for the hairpins (Table 3.1). In contrast to the current nearest-neighbour model,^[69, 105, 106] which takes only the non-canonical pairs adjacent to the canonical base paired stems into account for RNA secondary structure stability prediction, our results suggest that, a U·U pair which does not adjoin a canonical base paired stem also affects the thermodynamic stability. Interestingly, our values of change in free energy for both hairpin 1 and duplex 1 at 1 M NaCl buffer (Tables 3.1 and 3.2), as well as the previously measured value of change in free energy for hairpin 1 at 1 M KCl measured by thermal denaturation assay are well predicted by RNAstructure 5.3,^[69, 80, 105, 106] With the U·U to U·C substitutions, the structure is less stable than predicted for both the hairpin and duplex constructs. Our results suggest that replacing a single U·U to U·C pair may disrupt the remaining U·U pairs, possibly due to incompatibility of base stacking and/or the backbone conformation. The accuracy of RNA structure stability prediction model may increase by including the non-nearest-neighbor effects of consecutive non-canonical base pairs observed in this and previous studies^[72, 74-79].

3.3.3 Single-molecule force-ramp experiments reveal a two-state unfolding pathway with the formation of a misfolded hairpin

To further test how the U·U to U·C substitutions affect the mechanical unfolding and folding kinetics of the native hairpins, we carried out single-molecule force spectroscopic studies on the three hairpins constructs using optical tweezers. For each

construct, hysteresis was observed between the unfolding and refolding cycles of the force versus extension curves (FEC) during the force-ramp experiments (Fig. 3.4), suggesting the mechanical folding or unfolding pathway is far away from equilibrium at this condition.

Figure 3.4 Representative FECs for hairpin 3 (Fig. 3.1A). The FECs for hairpin 1 and hairpin 2 are similar to panels A and B. The force increasing and decreasing curves are shown in red and black, respectively. The hairpin unfolding and folding forces are indicated by horizontal red and black arrows, respectively. The grey, green and magenta dashed lines are EWLC fitting curves for DNA/RNA handles with ssRNA, misfolded hairpin and native hairpin, respectively. For most of the force decreasing traces (black), before the folding transition into the native hairpin, rapid hopping and gradual extension shortening (corresponding to the transition between ssRNA and the misfolded state) were observed at around 8-13 pN with the force decreasing (indicated by thick black arrows in black traces). (A) A typical pulling cycle with relatively large extension changes for the unfolding transitions (native hairpin to ssRNA (N-S)) at relatively high forces. (B) A typical pulling cycle with relatively small extension changes for the unfolding transitions (native hairpin to misfolded intermediate (N-M)) at relatively low forces. (C) A typical pulling cycle with relatively slow hopping dynamics (between native hairpin and the misfolded intermediate (N-M)) when the force is increased (occurrence frequency: 9.2%). The slow hopping was not observed for hairpins 1 and 2.

In the unfolding trajectories, the native hairpin (N) unfolds into stretched single strand (S) directly if the unfolding rupture force is above or around 13 pN, as confirmed by the nucleotides disrupted calculated from the change of extension values using EWLC model.^[34, 37, 54, 55, 57] Here the stretched single strand (S) state in mechanical unfolding experiment (295 K) is different from the random coil (U) in thermal denaturation assay.

If the unfolding force is below 13 pN, the measured ΔX values are smaller than the predicted extension change value corresponding to a native hairpin (N) unfolding to a single strand (S) (Fig. 3.4), suggesting an intermediate is formed. Such intermediate is observed in the refolding cycles. For all the 3 sequences, almost all of the traces fold into intermediate states before folding back to the native hairpin. Such transitions are indicated by the thick black arrows in Fig. 3.4. The rapid hopping traces suggest a fast and reversible reaction between the intermediate structure and the ssRNA. Compared to the data sampling frequency of 100 Hz, the folding rate of the intermediate structure is much faster.

We excluded the possibility that the intermediate state is an on-pathway semi-stable structure as has been observed in mechanical unfolding pathway of a protein.^[107] It's likely the intermediate structure has only one stem, which is probable as the change of extension of the intermediate to ssRNA are very small (around 2 nm) (Fig. 3.5). Then the nucleotides folded in the intermediate structure is around 13 ± 4 nt. The formation of an on-pathway intermediate RNA hairpin structure containing the loop and the apical two canonical base pairs of the native hairpin (9 nt in total) or additional with the four sequential non-canonical base pairs (17 nt in total) is not stable even without any stretching force.^[35] Moreover, as the rupture forces for the transition of intermediate to ssRNA are lower than that of unfolding a ssRNA helix, the possibility that the intermediate structure is a single-stranded helix is excluded.^[108] We reason that

the intermediate may be a 13-nt misfolded hairpin structure (M) (Fig. 3.5), which contains residues 15-27 of the construct. Such a misfolded hairpin is semi-stable and consistent with the previous studies.^[35, 109, 110] Other stable or semi-stable structures predicted by RNAstructure 5.3 are not in agreed with the ΔX values observed.

Figure 3.5 Possible structures predicted by RNAstructure 5.3 for the hairpin 1 (A) , 2 (B) and 3 (C) sequences, respectively. Predicted folding free energies for each structure

at 1 M NaCl at 37 °C is shown. The mutated bases in hairpins 2 and 3 are labeled by the red squares. 1 The optimal structures with lowest free energy predicted by the program. 2 A predicted suboptimal structure with two stems. The hairpins on the 5' side are not stable with a folding free energy value of only -1.7 kcal/mol. The 3' side hairpins are the assumed 13-nt misfolded hairpins. 3 and 4 Predicted suboptimal structures with unstable top stem-loop structures.

We did not further characterize the M-S transitions in the force-ramp data because of the fast hopping and the small change of extension. If the native hairpin structure is unfolded at a force significantly below 13 pN, the M-S transition is also expected to occur in the unfolding trajectories (indicated by thick red arrow for the red traces of Fig. 3.4B and C).

3.3.4 The destabilizing effect of a U·U pair to a U·C pair substitution is revealed by force-ramp experiment

In the force-ramp experiments, replacing a single U·U to U·C pair alters the unfolding rupture force distribution of the native hairpin (Figs. 3.4 and 3.6). For each U·U to U·C pair substitution, the mean unfolding rupture forces lower by ~ 1 pN, at a force loading rate of about 11 pN/nm. The decrease of mechanical stability of native hairpin is consistent with our thermal denaturation results.

In the unfolding trajectories, U·U to U·C base pair substitutions also increase the probability of hopping events, and of the intermediate formation at low force region (Fig. 3.6).

Figure 3.6 Rupture force versus extension change of unfolding transitions. A total of 920, 308 and 978 unfolding transitions were obtained, respectively, for hairpin 1 (red circles) (A), hairpin 2 (green triangles) (B), and hairpin 3 (blue squares) (C). The rapid hopping between ssRNA and misfolded state and increased Brownian noise at low forces may cause the broadening of the distribution of the extension change values (see panel C). The EWLC fitting curves are shown for the N-M, N-S, and M-S transitions, respectively.

We extracted the folding free energy values by applying Crooks fluctuation theorem on the FECs (Fig. 3.7 and Table 3.3).^[44, 49] The force ranges for the integration of the force-extension curves were chosen to ensure the mechanical work of (un)folding reactions between the native hairpins and ssRNAs was measured. For all the three constructs, the trend of values of free energies extracted from the force-ramp data is consistent with the trend of values of free energies from our thermal denaturation experiments. Compared to ensemble thermal denaturation method, the values of free energies extracted from force-ramp experiments are less accurate due to the large hysteresis in mechanical unfolding. We estimated the systematic error to be 15.0%, mainly because the folding transitions at low force region of the FECs are with large Brownian noise.

Figure 3.7 Extraction of the unfolding free energies from the force-ramp experiments using Crooks fluctuation theorem. (A) A typical force-ramp cycle for hairpin 1. Force increasing and decreasing curves are shown in red and gray, respectively. The EWLC fitting for the native, misfolded and ssRNA states of hairpin 1 are shown as red, magenta and black dashed lines, respectively. (B-D) Folding (solid) and unfolding (dashed) work distributions of hairpins 1 (B), 2 (C) and 3 (D) obtained from the force-ramp experiments. The force ranges for the integration of the force-extension curves were chosen to ensure the mechanical work of (un)folding reactions between the native hairpins and ssRNAs was measured. The bin size is 5 kcal/mol. The folding free energy values are shown in Table 3.3.

Table 3.3 Folding free energies for the native hairpin obtained from the force-ramp data base on the Crooks fluctuation theorem (Fig. S4). The free energy values at 22 °C are reported as mean $\pm$ standard error.

RNA molecule	Number of pulling cycles	Integration range (pN)	Work for stretching ssRNA (kcal/mol)	$-\Delta G^\circ$ (kcal/mol) ^a
Hairpin 1	741	4.5 – 22.0	10.0	15.2 ± 2.3
Hairpin 2	296	4.0 – 18.0	8.6	10.2 ± 1.5
Hairpin 3	890	3.0 – 16.5	8.0	9.1 ± 1.4

As both the folding and unfolding kinetics of misfolded hairpin (M) is very fast, we assumed the rate limiting step of the unfolding pathway is native hairpin (N) to single strand (S).^[111] With the assumption we extracted the apparent unfolding rates by applying the Szabo-Hummer-Dudko method (equation 2.18) on the unfolding rupture force distributions (Fig. 3.8 and Table 3.4).^[62] The U·U to U·C pair substitutions increase the unfolding kinetics. The force-dependent unfolding rates extracted are well fitted to the Bell model.^[28, 34-36] The force-dependent unfolding rates are not well fitted to more complex models (data not shown), which suggest a weak Hammond effect on the unfolding of hairpins, despite the U·U to U·C pair substitutions.^[62-64, 111]

Figure 3.8 Analysis of unfolding force distribution. (A) Unfolding force distributions for hairpins 1 (red), 2 (green) and 3 (blue), respectively. In panel A the unfolding transitions may include a mixture of native hairpin to ssRNA (N-S) and native hairpin to misfolded intermediate (N-S-M) transitions. The solid lines are the predicted

distributions with the parameters extracted from panel B by fitting to the Bell model. Analyzing the unfolding force distributions excluding the data points deviating from the EWLC model for N-S unfolding transitions does not affect the unfolding kinetics parameters significantly (Table S2). (B) Force dependent kinetics extracted from panel A and C using Szabo-Hummer-Dudko method are fit to the Bell model. Errors for the rate constants are estimated based on Poisson noise.

We also tested the effect of the unfolding transitions significantly deviating from the theoretical values corresponding to N-S transition. The unfolding transitions with $F < 14$ pN and $\Delta X(F) + 2 < \Delta X_{\text{EWLC}}(F)$ are excluded and the unfolding rates were extracted by applying Dudko's method to the rupture force distributions of the remaining data. Here $\Delta X_{\text{EWLC}}(F)$ is theoretical value of extension change corresponding to N-S transition predicted by EWLC model. The kinetics parameters are not significantly affected by the exclusion (Table 3.4).

Table 3.4 Unfolding kinetics parameters obtained for the native hairpin constructs from force-ramp experiments by fitting the unfolding rupture force distributions (Fig. 5). All uncertainties are reported as standard errors.

	RNA molecule	$\ln k_0$ (s ⁻¹)	$\Delta X^\ddagger$ (nm)	Adjusted r^2
Including all data points in Figure 3.6	Hairpin 1	-9.5 ± 0.7	2.8 ± 0.2	0.98
	Hairpin 2	-9.0 ± 0.5	3.0 ± 0.2	0.98
	Hairpin 3	-4.7 ± 0.7	2.0 ± 0.3	0.92
Excluding the data points deviating from EWLC for the N-S transition	Hairpin 1	-12.5 ± 1.6	3.5 ± 0.5	0.94
	Hairpin 2	-8.8 ± 1.2	2.9 ± 0.4	0.94
	Hairpin 3	-6.3 ± 0.5	2.5 ± 0.2	0.96

3.3.5 The presence of force-induced misfolded intermediate and stabilization effect of U·U to U·C pair substitutions are further confirmed by constant-force experiments

The constant-force experiments were applied to further characterize the mechanical properties of the RNA hairpins. A hopping method was applied to study the (un)folding of hairpins 2 and 3 at constant forces within the range of 8-11 pN, respectively (Fig. 3.9A).^[35, 36] For the hairpin 1, a force-jump method was applied to measure the folding and unfolding kinetics within the force ranges of 8.5-10.0 pN and 13.5-15.0 pN, respectively (Fig. 3.9B)^[35, 36].

Figure 3.9 Constant-force experiment results. Here the “unfolded” state is a mixture of ssRNA (S) and a misfolded hairpin (M). (A) A representative constant-force (8.5 pN) hopping trace of extension versus time and force versus time curves of hairpin 3. (B) A representative force-jump trace of hairpin 1. Folding and unfolding occur at 9.5 and 14.5 pN, respectively. (C) Folded state (red) and unfolded state (green) lifetimes sorted

in descending order are fitted to single exponential distributions for extracting unfolding and folding kinetics, respectively. We also compared the fitting parameters using descending ordering method to that using binning method (data not shown). The values of kinetic rates from both methods are close to that of reciprocal averaged lifetime. We did not observe any possible non-zero maximum larger than the time resolution. (D) Summary of force-dependent folding and unfolding rates, which are shown in open and closed symbols (red circle, hairpin 1; green triangle, hairpin 2; blue square, hairpin 3), respectively. Data are fit to Bell model.

In the constant-force experiments, with the forces above 13 pN, the ΔX values corresponded to the prediction of the EWLC model for the N-S transitions (Table 3.5), which is consistent with the force-ramp results. When the force is below 13 pN, the ΔX values are smaller than the prediction for the N-S transitions but larger than that for the N-M transitions. We reason that constant forces below 13 pN, the hairpin constructs may also fold into the misfolded hairpins in force-ramp experiments. However, we didn't directly observe the (un)folding transitions of the misfolded hairpins, presumably because of the fast kinetics rates are beyond the data acquisition rate of 100 Hz with force feedback response of about 0.05 s. Our results are consistent with the fast exchange dynamics on the millisecond time scale observed for a relatively small 6-bp tetraloop DNA hairpin under stretching force.^[112] Moreover, the Δx values for the M-S transitions may be concealed by Brownian noise and/or baseline drifting. As a result, the “unfolded” state at forces below 13 pN in the constant-force traces may contain an unresolved mixture of ssRNA (S) and misfolded small hairpin (M).

As the unfolding of the native hairpins is much slower than the hopping between ssRNA (S) and the misfolded hairpin (M) at forces below 13 pN, the apparent mechanical kinetics of the native hairpins to the ‘unfolded state’ is able to be obtained

from the extension versus time traces (Fig. 3.9 C,D and Table 3.5). The force-dependent kinetic rates constants were extracted from life span distributions using single exponential fitting. Consistent with our force-ramp experiments, the mechanical unfolding kinetics of hairpin 1 is the slowest. Hairpin 2 unfolding kinetics is slightly slower than hairpin 3.

The (un)folding transition state positions can be obtained from the fitting of (un)folding kinetics data using Bell model. As shown in Figure 3.9D and Table 3.6, the U·U to U·C mutations do not affect the slope of the force-dependent unfolding kinetics, i.e., the mechanical unfolding transition state position ($\Delta X_{\text{unfolding}}^{\ddagger}$), but lower the mechanical unfolding barrier (unfolding force). Based on the $\Delta X_{\text{unfolding}}^{\ddagger}$ values and EWLC model, there are about 7 nucleotides disrupted at the mechanical unfolding transition state (Table 3.6). Thus, at the mechanical unfolding transition state, 3 or 4 base pairs at the bottom stem are disrupted. Thus, our mechanical unfolding results directly reveal that the U·U to U·C mutations affect the stability of the 6-bp bottom stem and/or the transition state structure with 3-4 Watson-Crick base pairs disrupted (i.e., with 2-3 Watson-Crick base pairs formed below the 4 sequential non-canonical base pairs). The mechanical unfolding transition state position obtained from the constant-force data is consistent with the analysis for the force-ramp data.

Table 3.5 Constant-force experiment results for the native hairpin constructs in 200 mM NaCl, 10 mM Tris-HCl, 0.1 mM EDTA, pH 7.3^a.

Force (pN)	ΔX measured (nm) ^a	ΔX predicted (nm) ^b	k (s ⁻¹)	$\langle \tau \rangle^{-1}$ (s ⁻¹)
Hairpin 1 unfolding				
13.5	10.4 ± 0.4	10.2 (6.7)	0.42 ± 0.01	0.45
14	10.7 ± 0.4	10.3 (6.8)	0.60 ± 0.02	0.52
14.5	10.8 ± 0.3	10.4 (6.8)	0.84 ± 0.01	0.79
15	11.0 ± 0.5	10.5 (6.9)	1.15 ± 0.04	1.11
Hairpin 1 folding				

8.5	6.4 ± 0.4	8.7 (5.9)	0.40 ± 0.01	0.39
9	6.7 ± 0.4	8.9 (6.0)	0.24 ± 0.01	0.26
9.5	7.0 ± 0.6	9.1 (6.1)	0.24 ± 0.01	0.26
10	7.6 ± 0.7	9.3 (6.2)	0.16 ± 0.01	0.17
Hairpin 2 unfolding				
9	7.4 ± 0.7	8.9 (6.0)	0.07 ± 0.01	0.07
9.5	7.8 ± 0.8	9.1 (6.1)	0.12 ± 0.01	0.13
10	8.1 ± 0.8	9.3 (6.2)	0.11 ± 0.01	0.11
10.5	8.8 ± 0.9	9.4 (6.3)	0.21 ± 0.01	0.16
Hairpin 2 folding				
9	7.4 ± 0.7	8.9 (6.0)	0.34 ± 0.01	0.35
9.5	7.9 ± 0.9	9.1 (6.1)	0.16 ± 0.01	0.16
10	8.0 ± 0.9	9.3 (6.2)	0.16 ± 0.01	0.16
10.5	8.7 ± 1.0	9.4 (6.3)	0.10 ± 0.01	0.09
Hairpin 3 unfolding				
8.5	7.0 ± 0.6	8.7 (5.9)	0.51 ± 0.01	0.51
9	6.9 ± 0.6	8.9 (6.0)	0.75 ± 0.01	0.73
9.5	7.6 ± 0.7	9.1 (6.1)	1.14 ± 0.02	1.09
10	8.4 ± 0.7	9.3 (6.2)	1.42 ± 0.02	1.34
Hairpin 3 folding				
8.5	7.0 ± 0.6	8.7 (5.9)	2.47 ± 0.05	2.14
9	7.1 ± 0.6	8.9 (6.0)	1.60 ± 0.04	1.44
9.5	7.9 ± 0.6	9.1 (6.1)	0.65 ± 0.01	0.60
10	8.4 ± 0.8	9.3 (6.2)	0.36 ± 0.01	0.36

^a All experimental data are reported as mean ± standard error, except that the uncertainties for k are reported as 95% confidence interval for the single-exponential fitting.

^b Calculated using EWLC model, for the complete unfolding of the 29-nt native hairpins (N) to ssRNAs (S) with the diameter of the hairpin stem (2 nm) taken into account. Values shown in parentheses are for the predicted extension change upon the unfolding of the 29-nt native hairpins to the 13-nt misfolded hairpins (M).

Table 3.6 Bell model fitting parameters of constant-force experiment results.

RNA molecule	$\ln k_0$ (s ⁻¹)	$\Delta X^\ddagger$ (nm)	Number of	Adjusted r ²

	nucleotides folded or unfolded ^a			
Hairpin 1 (unfolding)	-10.1 ± 0.2	2.8 ± 0.1	7 ± 1	0.99
Hairpin 1 (folding)	4.1 ± 1.4	-2.4 ± 0.6	N/A	0.84
Hairpin 2 (unfolding)	-9.5 ± 2.3	3.1 ± 1.0	8 ± 3	0.73
Hairpin 2 (folding)	6.0 ± 1.6	-3.2 ± 0.7	N/A	0.87
Hairpin 3 (unfolding)	-6.7 ± 0.4	2.9 ± 0.2	7 ± 1	0.99
Hairpin 3 (folding)	12.1 ± 0.8	-5.3 ± 0.3	N/A	0.99

^a The number of nucleotides disrupted at the unfolding transition state is calculated using EWLC model. The specific number of nucleotides folded at the folding transition state is not shown here since the “unfolded” state may be a mixture of ssRNA (S) and misfolded intermediate (M). All the folding forces for the constant-force experiments are below 11 pN, thus the 13-nt off-pathway misfolded structure may form at the “unfolded” state. The diameter of the hairpin stem is cancelled out if the off-pathway intermediate hairpin forms at the “unfolded” state. We can estimate the ranges of the number of nucleotides folded at the transition state based on EWLC model to be 6-19, 8-21, and 14-27, respectively, for hairpins 1, 2, and 3.

The mechanical folding rates obtained from the constant-force experiments are similar for all three hairpin constructs, with a slight increase upon the substitution of the second U·U pair with a U·C pair. Thus, the substitution of U·U with U·C pairs destabilizes the RNA native hairpin mainly through increasing mechanical unfolding kinetics. The destabilization effect of the U·U to U·C mutation may results in an increased folding rate of the native telomerase RNA pseudoknot, which requires the disruption of the U·U and U·C pairs and the hairpin loop structure.^{[15, 35, 80, 102, 109, 110, 113-}

116]

The mechanical folding transition state position varies significantly for the three constructs, as indicated by the varied slopes for the force-dependent folding kinetics curves (Fig. 3.9D). In consideration of the presence of a mixture of misfolded intermediate (M) and ssRNA (S) at the “unfolded” state in the constant-force experiments (with the force below 13 pN), the mechanical folding transition state position could be 0-13 nt further toward the fully-folded native hairpin state (N) (Table 3.5). At the mechanical folding transition state, the non-Watson-Crick U·U and U·C pairs may be formed as evidenced by the facts that (1) All three hairpins show relatively slow apparent folding kinetics and low folding forces, and (2) Hairpin 3 and hairpin 1 constructs have the apparent folding transition state positions closest and farthest to the fully-folded native hairpin (N), respectively. Slow mechanical folding kinetics due to the presence of a non-Watson-Crick structure at the mechanical folding transition state has been reported previously.^[35, 36, 38] Significantly, we demonstrate in this work that a subtle U·U to U·C mutation results in (1) the destabilization of the stem in the native hairpin structure, (2) a small increase in apparent folding kinetics of the native structure, and (3) shifting of the folding transition state.

The crossing points of the curves of the force-dependent (un)folding kinetics (Fig. 3.9D) give the critical forces (F_c , with equal unfolding and folding rates). The critical forces are 11.1 ± 0.3 , 10.0 ± 0.3 and 9.3 ± 0.2 pN, respectively, for hairpins 1, 2 and 3 (mean $\pm$ standard deviation). We obtained the free energies required to unfold the native hairpins at the critical forces. Apparently, the U·U to U·C mutation destabilizes the native hairpin structure.

3.3.4 A comparison between single-molecule mechanical (un)folding results and ensemble thermal denaturation results

Single-molecule stretching experiments facilitate the direct revelation of the misfolding events. To further confirm the proposed misfolded small hairpin is reasonable in terms of folding stability, we constructed two truncated 15-nt sequences (M1 and M2, from residues 14-28) containing the proposed non-native 5-bp hairpin structures (Fig. 3.10B) for ensemble thermal melting studies at both 200 mM and 1 M NaCl (Table 3.1). The thermal denaturation of two truncated 15-nt constructs M1 and M2 show only one two-state transition. The melting temperature (T_m) values for M1 and M2 are in between T_{m1} and T_{m2} for the native hairpin constructs (Table 3.1). Thus, in the reversible thermal melting of the native hairpin constructs, the off-pathway misfolded intermediates may not be present. Thus, the misfolded intermediate is induced by the relatively low stretching force in mechanical (un)folding reactions, as observed before for other RNAs and DNAs.^[38, 117]

Figure 3.10 The comparison of thermal and mechanical (un)folding pathways and the intermediate and transition state structures. (A) On-pathway intermediate state (I) in thermal (un)folding. The mutated residues in hairpins 2 and 3 are highlighted by the

red squares. (B) The proposed misfolded state (M) in mechanical (un)folding reactions. The misfolded hairpin constructs for the ensemble thermal melting experiment (Table 1) are truncated from the native hairpins (residues 14 to 28). The mutated residues in hairpins 2 and 3 are highlighted by the red squares. (C) Thermal (un)folding involves the native hairpin (N), on-pathway intermediate (I, see panel A), and unfolded single-stranded random coil (U). (D) Mechanical unfolding at a relatively high force (e.g., 15 pN) involves the native hairpin (N) and stretched single strand (S). (E) Mechanical (un)folding at a relatively low force (e.g., 10 pN) may involve the native hairpin (N), stretched single strand (S), and misfolded off-pathway intermediate (M, see panel B). The mechanical unfolding transition state ($\ddagger$) of the N to S reaction is shown with the bottom 3-4 base pairs disrupted.

Based on the free energies at zero force obtained from the thermal melting experiments and EWLC model for stretching ssRNA, we can relate the mechanical unfolding with thermal melting data (equation 3.10).

$$\Delta G_{F_c,MS} = \Delta G_{MU}^{\circ} + \Delta G_{\text{stretching},13 \text{ nt}}(F_c) \quad (3.10)$$

Here, $\Delta G_{F_c,MS}$ is the unfolding free energy at the critical force (F_c). ΔG_{MU}° is the unfolding free energy at zero force at 295 K, obtained from the ensemble thermal melting experiments (Table 3.1). $\Delta G_{\text{stretching}}$ is the free energy of stretching an ssRNA from 0 pN to a certain force obtained by numerical integration of the EWLC model. We obtained the predicted critical forces (F_c) for the (un)folding of the proposed misfolded intermediate structures M1 and M2 to be 11.9 ± 0.1 and 10.6 ± 0.1 pN, respectively (without considering the diameter of hairpin stem) (mean $\pm$ standard deviation). Thus, the proposed off-pathway intermediate structures (M) are consistent with the fact they are in dynamic exchange (hopping) with the stretched single strands (S), when the forces are close to about 11 pN.

In addition, we calculated the critical forces (F_c) and free energies at the calculated critical forces (ΔG_{Fc}) for the unfolding of native hairpins to single strands (N-S) and native hairpins to misfolded intermediates (N-M), respectively, based on the free energy values at zero force (ΔG°) determined by ensemble thermal melting results (Table 3.1) and the EWLC model (equations 3.11 and 3.12).

$$\Delta G_{Fc,NS} = \Delta G^\circ_{NU} + \Delta G_{\text{stretching},29 \text{ nt}}(F_c) \quad (3.11)$$

$$\Delta G_{Fc,NM} = \Delta G^\circ_{NU} - \Delta G^\circ_{MU} + \Delta G_{\text{stretching},16 \text{ nt}}(F_c) \quad (3.12)$$

Here, ΔG°_{NU} and ΔG°_{MU} are the unfolding free energies at zero force at 295 K, obtained from the ensemble thermal melting experiments (Table 3.1). For all the three hairpin constructs, the measured unfolding free energies at critical forces (ΔG_{Fc}) from the constant-force experiments are larger than the calculated $\Delta G_{Fc,NM}$, but smaller than $\Delta G_{Fc,NS}$ (Table 3.7). The results further suggest that, with the stretching force around 8-13 pN, the misfolded hairpin structure (M) and ssRNA (S) are in rapid exchange, and the “unfolded” states in the constant-force experiment are a mixture of ssRNA (S) and misfolded intermediate (M).

Table 3.7 Comparison of the measured and calculated values for the critical forces, and extension changes and unfolding free energies at the corresponding critical forces.^a

RNA molecule	Reaction	F_c (pN)	ΔX_c (nm)	ΔG_{Fc} (kcal/mol)
Hairpin 1	Observed	11.1 ± 0.3	$8.2 \pm 0.5b$	13.0 ± 0.9
	N-S (calculated)	12.5 ± 0.2	10.0 ± 0.2	17.9 ± 0.3
	N-M calculated)	9.2 ± 0.4	6.0 ± 0.2	8.0 ± 0.3
Hairpin 2	Observed	10.0 ± 0.3	8.1 ± 0.9	11.6 ± 1.3

	N-S (calculated)	11.6 ± 0.1	9.7 ± 0.1	16.2 ± 0.2
	N-M calculated)	8.7 ± 0.3	6.0 ± 0.2	7.4 ± 0.2
Hairpin 3	Observed	9.3 ± 0.2	7.9 ± 0.6^c	10.5 ± 0.8
	N-S (calculated)	10.0 ± 0.4	9.2 ± 0.4	13.2 ± 0.5
	N-M calculated)	6.0 ± 0.7	5.1 ± 0.6	4.4 ± 0.5

---

^a The constant-force experiments were done in 200 mM NaCl, 10 mM Tris-HCl, 0.1 mM EDTA, pH 7.3 at 22 °C. At the critical forces, F_c , the “unfolded” and folded states are equally populated. The calculated values are based on the free energies for thermal unfolding reactions (ΔG°_{N-U} and ΔG°_{M-U} , see Table 3.1). ΔG_{Fc} is the free energy for unfolding at critical force ($\Delta G_{Fc} = N_A \cdot F_c \Delta \cdot X_{Fc}$), i.e., the free energies for stretching ssRNA are not subtracted. N_A is Avogadro constant. All uncertainties are reported as standard deviations. The observed values for all three sequences obtained from the constant-force experiment are in between the calculated values for the N-S and N-M reactions, suggesting that the “unfolded state” in the constant-force traces may be a mixture of ssRNA (S) and misfolded intermediate (M).

^b The value was obtained from the force-ramp and constant-force data at 11.0 pN (Data not shown).

^c The value was obtained from the constant-force measurement at 9.5 pN.

3.4 Conclusions

Our combined ensemble thermal denaturation and single-molecule mechanical unfolding studies show that a single U·U to U·C substitution within the four sequential non-canonical base pairs region destabilizes the native hairpin structure. Due to the decrease of the native hairpin unfolding forces, during the mechanical unfolding by force-ramp, the stretching force-induced misfolding is observed to occur more frequently upon the thermodynamically destabilizing U·U to U·C substitutions.

Compared to the hairpin 1 construct with the off-way intermediate forming a 5-bp triplloop, hairpin 2 and hairpin 3 constructs have the same destabilized off-pathway intermediate structure forming a 4-bp pentaloop (Fig. 3.10B). Interestingly, compared to hairpins 1 and 2, hairpin 3 shows slightly faster apparent folding kinetics, indicating that, under the stretching force (8-13 pN), disruption of the misfolded small hairpin may not be the rate-determining step for the formation of the native hairpin. Thus, the native hairpin is mainly stabilized by the native interactions and by slowing down the unfolding kinetics. Our results suggest that a mutation in non-Watson-Crick base pairs of an RNA native structure may significantly modulate the RNA (un)folding dynamics. Interestingly, previous single-molecule FRET studies of the telomerase RNA pseudoknot domain suggest that replacing two U·U pairs with U·C pairs cause more frequent misfolding involving a similar nonnative hairpin structure.^[110]

The U·U and U·C pairs and adjacent structures present in the telomerase RNA hairpin need to be disrupted for the formation of functionally important pseudoknot and triplex structures.^[15, 80, 102, 116] Because the telomerase activity is highly up-regulated in most cancer cells, the unique structural and kinetic feature of thermodynamically-coupled Watson-Crick and non-Watson-Crick regions of the telomerase RNA hairpin structure may be the drug targets. Similarly, the expression and nuclear retention element (ENE) present in some viral and cellular long noncoding RNAs (lncRNAs) is also a hairpin structure containing a U-rich internal loop, which can form pseudoknot and triplex structures with the downstream A-rich tracts.^[14, 18, 84-86] It is expected that the folding dynamics and functions of RNAs may also be modulated by other non-canonical pairs such as the consecutive G·A and A·A pairs in kink turns and other structural motifs.^[72, 75, 79, 118]

The long-range effect on structural stability and dynamics is commonly found in proteins for allostery and signal transduction. The single-molecule mechanical

manipulation approach has been applied to directly probe the long-range effect on protein structural stability and dynamics.^[119] However, the detailed structural features of the mechanical (un)folding transition state are often not easily derived for proteins.^[119-121] Allostery through a DNA duplex has also been observed to result in the regulation of the binding dynamics of multiple ligands to the same DNA in a distance.^[122, 123] Our single-molecule mechanical unfolding/unzipping studies on the telomerase RNA hairpin directly reveal that 3-4 base pairs are disrupted at the unfolding transition state, and that a non-Watson-Crick base pair may affect the structural stability/dynamics at least 2-3 Watson-Crick base pairs away. Further detailed experimental and computational studies are needed to understand how the long-distance coupling may occur through RNA helical structures. The non-nearest-neighbor residues may affect the biological functions involving the disruption of pre-formed RNA structures. For example, the helicase activity may be affected by the RNA structures in a distance for the regulation of pre-mRNA splicing and mRNA translation.^[19, 20, 29, 32, 33]

3.5 Perspectives

Non-Watson-Crick base pairs widely exist in RNA structures. Our single-molecule and ensemble methods can be used to study other RNA structures containing non-Watson-Crick base pairs, such as A·G, A·C and A·A pairs. Also, non-Watson-Crick base pairs are often ligand binding sites. Our single-molecule and ensemble methods may be used to study the effect of ligand binding on Non-Watson-Crick base pairs containing structures.

CHAPTER IV

Minor-Groove Base Triple Formation in SRV-1 RNA Pseudoknot Increases Mechanical Stability and –1 Ribosomal Frameshifting Efficiency

4.1 Introduction

Minus-one Ribosomal frameshifting is an alternative translational mechanism widely utilized by many viruses to maintain a certain ratio of gag and gag-pol protein expression,^[124-129] which is crucial for the replication and viability of the viruses.^[130] Alteration of frameshifting efficiency can dramatically affect the virus reproduction.^[131] During –1 ribosomal frameshifting process,^[132-134] a ribosome is positioned at a certain sequence of mRNA called slippery sequence, and has a certain probability to move one nucleotide backward, before continuing translation in the new reading frame. Upon frameshifting, the stop codon of the original frame (0 frame) is bypassed, and thus a fusion protein may be synthesized. The slippery sequence is usually X XXY YYZ, where XXX are any 3 identical nucleotides, YYY are either AAA or UUU, and Z is usually A, C or U. An RNA structure located downstream of the slippery sequence separated by a 2 to 8 nt single-stranded spacer stimulates –1 frameshifting. The RNA structure can be a hairpin, a three-way junction or a pseudoknot. All the three components, namely slippery sequence pattern, spacer length and downstream RNA structure can significantly affect the –1 frameshifting efficiency.^[132]

Several steps are involved in –1 frameshifting.^[135, 136] Firstly, the downstream RNA structure blocks the ribosome at the entrance of the downstream tunnel, resulting in the

stalling and pausing of the ribosome over the slippery sites. Secondly, with the coupled rearrangement of tRNA and mRNA from the A/T hybrid state to the A/A state, the anticodon loop of the aminoacylated tRNA, which is base paired with the mRNA, moves about 0.9 nm towards the 5' direction of the mRNA, resulting in a tension force generated in the mRNA spacer region between the A-site and the downstream structure at the entrance of the ribosome tunnel. Here the A/T state denotes that tRNA is in the A-site on the 30S subunit as well as interacts with elongation factor Tu (EF-Tu). The A/A state denotes that tRNA is in the A-site on both the 30S and 50S subunits.^[137] The tension force stretches both the downstream structure and the tRNA-bound codons. To release the tension, the reading frame may shift 1 nucleotide towards the 5' direction, if the downstream structure is mechanically stable enough against the tension force. Finally, the ribosome unfolds the downstream RNA structure, and the translation continues in the new -1 frame. Thus, it is critical to understand how the mechanical properties of the downstream structure affect -1 frameshifting.

In spite of recent progress, the detailed mechanism of how the RNA structure determines -1 frameshifting efficiency is still unclear. Mechanical pulling studies using single-molecule force spectroscopy have provided insight into the possible correlation between frameshifter RNA structures and -1 frameshifting efficiency.^[15, 41, 138-140] Two studies showed a correlation between the frameshifting efficiency and the mechanical strength of a pseudoknot,^[15, 138] while such correlation was not observed in the other studies.^[41, 139, 140] Both major-groove and minor-groove base triples in an RNA pseudoknot were found to increase -1 frameshifting efficiency and mechanical stability.^[15] Green et al suggested that a slower unfolding rate correlates with higher -1 frameshifting efficiency over a narrow force range.^[41] Woodside et al proposed a mechanism that the frequency of the formation of alternative structures is proportional to -1 frameshifting efficiency.^[139]

An H-type pseudoknot structure is formed in the overlapping region of gag and pol genes of Simian retrovirus type-1 (SRV-1). An NMR study revealed that minor groove base triples form between stem 1 and loop 2 of the pseudoknot.^[141] Extensive mutagenesis assays have revealed that the frameshift efficiencies *in vivo* and *in vitro* are highly dependent on the loop residues in the pseudoknot.^[141, 142]

Here we studied the mechanical properties, thermal stability, and electrophoresis migration rates of SRV-1 RNA frameshifting pseudoknots, and compared our results to the *in vitro* frameshifting efficiencies measured previously (Figure 4.1).^[142] All the constructs contain the same slippery sequence (GGGAAAC) and a 7-nt spacer. The frameshifting pseudoknots chosen in our studies include the ‘wildtype’ pseudoknot (SF206), whose three-dimensional structure is revealed by NMR, and the mutants investigated by mutagenesis assays. The structure of ‘wildtype’ pseudoknot unravels the minor-groove base triples formation between stem 1 and loop 2, as well as the stacking of bases in loop 2. Mutations expected to disrupt the base triples and the base stacking are with low frameshifting efficiencies.^[142] So we used the ‘wildtype’ pseudoknot, as well as 6 mutants with *in vitro* frameshifting efficiencies in the range between 21.0% and 0%, to investigate the possible coupling effect between stem 1 and loop 2 on thermal and mechanical stabilities of the frameshifting pseudoknots, and the possible correlation to the frameshifting efficiency.

Figure 4.1 (A) secondary structure of pseudoknot SF206 and mutants with the flanking handles. The mRNA strand is highlighted in gray. The DNA strands complementary to the mRNA in the handles are highlighted in dark green. The same slippery sequence (GGGAAAC, indicated by a black line) is present in all sequences. A 7-nt spacer (UGAUCGA) is present in SF206, SF209 and SF211, and a slightly different spacer (UAGUUGA) is present in the remaining 4 constructs. (B) Experimental setup of single-molecule force spectroscopy using Minitweezers. The drawing is not to scale. (C) Minor-groove base triples of SF206. Triple interaction are formed between A28 of loop 2 and C6-G14 of stem 1 (colored blue), between U27 of loop 2 and C5-G15 of stem 1 (colored green) and between A26 of loop 2 and G4-C16 of stem 1 (colored magenta).

4.2 Materials and methods

4.2.1 Thermal denaturation experiments

The HPLC purified RNA oligonucleotides for thermal denaturation experiments were purchased from Sigma. All the samples were verified by mass spectroscopy by the company. The ensemble UV thermal denaturation experiments for the pseudoknots and a control hairpin were conducted using a Shimadzu 2550 spectrometer. The control hairpin (residues 1 to 18) is truncated from the pseudoknot SF206 (residues 1 to 34). The temperature ramping rate is 0.5 °C/min. Absorption values at 260 nm were recorded. The buffers contain 20 mM HEPES, 0.1 mM EDTA, pH 7.3 with varying NaCl concentrations (<0.1 mM, 200 mM or 1 M NaCl).

The RNA sample concentrations were 3 µM. All pseudoknots and the control hairpin were heated at 85 °C then slowly cooled to 15 °C in more than 2 hours before the thermal melting experiments. The melting temperatures were obtained from dA/dT versus temperature curves. The measurement for each sample at each buffer condition was repeated three times.

4.2.2 Native PAGE

RNA oligonucleotides for native PAGE were annealed in the buffer with 200 mM NaCl, 20 mM HEPES, 0.1 mM EDTA, pH 7.3. The pseudoknots were heat at 95°C for 5min, then slowly cooled down to room temperature for 3h and equilibrated at 4°C overnight. The control hairpin was heat at 95°C for 5min, then snap-cooled on ice and equilibrated at 4°C overnight. Each sample was mixed with 6× loading dye (iDNA) and diluted with buffer to 1 µM. Then 20 µL of each sample was loaded onto a 15%

native polyacrylamide gel with 1× TBE (10 mM EDTA, 90 mM Tris/Borate) running buffer. The gel was run at 4°C and 250 V for 10 h, then was stained with Ethidium bromide (Bio-Rad).

4.2.3 Sample preparation for single-molecule optical tweezers experiments

Single-molecule constructs were made as described previously,^[15, 35] using the recombinant pSFCASS5 plasmids containing the pseudoknot sequences.^[142] The ~1160-nucleotide (nt) RNA molecules were made by PCR amplification of the plasmids followed by *in vitro* transcription by T3 RNA polymerase (Promega). The RNAs contain a 475-nt upstream sequence, a 2-nt single stranded linker (GA), the pseudoknot sequences, a downstream 2-nt single stranded linker (AU), and a 630-nt downstream sequence. The RNAs were annealed with complementary strands of PCR-generated 475-bp and 630-bp double-stranded DNAs (dsDNAs) to generate RNA/DNA hybrid handles (handle A and handle B). T4 DNA polymerase (NEB) was used to label the 3' end of the DNA strand of handle A by introducing biotin-16-dUTP (Roche). Handle B was labelled at the 5' end of the DNA strand by digoxigenin labeled primer during PCR. 1.8-μm Streptavidin coated polystyrene beads were purchased from Spherotech. Recombinant protein G (Thermo Scientific) was coated on 3-μm carboxyl coated polystyrene beads (Spherotech) using EDC (Sigma) and sulfo-NHS (Thermo Scientific). Dimethyl pimelimidate (Thermo Scientific) was used to cross-link anti-digoxigenin antibody to the protein G-coated beads. Only appropriately annealed DNA/RNA hybrids are able to form tether with both types of beads.

4.2.4 Optical tweezers experiments

Measurements were performed using a dual-beam Minitweezers.^[58] The laser powers used for both lasers are 100 mW. The anti-digoxigenin-coated polystyrene bead was trapped in the optical trap and the streptavidin-coated polystyrene bead was held on a micropipette by suction. In a force-ramp experiment, the force was changed by moving the optical trap at different constant speeds, with a constant stiffness of about 0.07 pN/nm, yielding approximate constant force loading rate from 4 to 24 pN/s. The force was maintained by a ~200 Hz electronic force feedback in constant-force experiments. The data acquisition rate was 1000 Hz. All of the optical tweezers experiments were carried out at 22 ± 1 °C. The buffer contains 200 mM NaCl, 10 mM Tris-HCl, 0.1 mM EDTA, pH 7.3. For each sample, data were collected from at least 3 individual tethers.

4.3 Results and discussion

4.3.1 Pulling experiments reveal that SRV-1 pseudoknots mainly unfold in one step

In our pulling experiments, all sequences refolded into pseudoknot structures as indicated by the subsequent unfolding traces. More than 95% of pulling cycles showed a one-step unfolding transition during stretching (Figure 4.2A, Table 4.1). The extension changes of one-step unfolding transition are consistent with the prediction of EWLC model for unfolding of pseudoknots into stretched single strands. In rare cases, we observed two-step unfolding traces (Figure 4.2B). A typical one-step pulling cycle has a large hysteresis, while a typical two-step pulling cycle shows decreased unfolding force and has a small hysteresis or is reversible.

Figure 4.2 Pulling cycles of SF206 with (A) typical one-step unfolding and (B) rare two-step unfolding. Red curves are unfolding traces and black curves are refolding traces.

Table 4.1 Summary of the pulling experiments.

Molecule	Number of 1-step unfolding traces	Extension change of 1-step unfolding transitions (nm)	Nucleotides unfolded for 1-step unfolding transitions ^a	Number of 2-step unfolding traces	Occurrence frequency of 2-step traces
SF206	888	12.4 ± 0.9	33.1 ± 1.9	1	0.1%
SF209	774	12.6 ± 1.0	33.7 ± 2.1	21	2.6%
SF211	511	12.5 ± 0.9	33.7 ± 1.8	1	0.2%
SF217	908	12.8 ± 1.1	34.9 ± 2.4	20	2.2%
SF220	601	13.0 ± 1.0	34.7 ± 2.0	10	1.6%
SF229	774	12.7 ± 1.1	34.8 ± 2.3	39	4.8%
SF348	484	8.9 ± 0.9	26.9 ± 1.7	23	4.5%

All uncertainties are reported as standard deviations.

^a Calculated using EWLC model. The end-to-end distance (4 nm) of the pseudoknots is considered.

4.3.2 SRV-1 pseudoknots mechanically fold in two steps

More than 99% of the refolding trajectories contain two transitions. The near-reversible two-step force-ramp transitions in both unfolding and refolding trajectories indicate that the unfolding and refolding intermediates are actually the same structure. We performed the constant-force experiments to measure the extension changes for both transitions during unfolding and refolding, respectively (Figure 4.3).^[36] Consistent with the force-ramp experiment, the constant-force experiment reveals one-step unfolding and two-step folding traces. In the constant-force folding traces, the first step corresponds to fast and reversible transition between single-stranded and intermediate state with about 6-7 nm extension change. The second step corresponds to the transition from intermediate state to pseudoknot with about 4-5 nm extension change. The extension change of the single-stranded state to intermediate state transition indicates the formation of a 19-nt hairpin containing stem 1 of the pseudoknots (Figs. 4.1A and 4.3C). After the formation of the pseudoknot structure at low forces, the constant-force trajectory is stable without hopping, which is consistent with the fact that the pseudoknots have relatively high unfolding forces. We observed that one of the relatively unstable pseudoknots, SF217, has a lifetime of tens of seconds for the pseudoknot form at 15 to 17 pN (Data not shown). SF348 has a similar folding pattern with a relatively smaller extension change for the intermediate state to pseudoknot state due to the presence of a significantly shorter loop 2. We did not observe the intermediate with stem 2-containing hairpin forming. The formation of stem 2-containing hairpin from single strand would result in an extension decrease of about 8 to 10 nm at the force range of 10 to 30 pN. Our results are consistent with the simulation study by Cho et al.^[143]

Figure 4.3 (A) Force-jump experiment results for SF211. Data are subsampled to 100 Hz. (B) Distribution of extension data points between the dashed lines shown in (A). Gaussian fittings indicate 3 states, namely the pseudoknot (P), intermediate (I) and single strand (S). (C) Mechanical unfolding and folding mechanisms. The drawing is not to scale.

4.3.3 The terminal base pair of stem 1 affects the unfolding kinetics and frameshifting

We tested the importance of the terminal G-C base pair of stem 1, which is directly in contact with ribosome, and may affect the mechanical unfolding and frameshifting.^[142]

According to nearest-neighbour model, changing G-C (SF206) to C-G (SF209) and A-U (SF211) pair destabilize the stem 1, by 0.5 or 1.9 kcal/mol (penalty for terminal A-U pair is considered) , respectively, at 1 M NaCl and 310 K (predicted using RNAstructure 5.3).^[69, 105, 106] Consistently, the $\ln k_0$ values of SF206, SF209 and SF211 are -15.9 ± 0.6 , -15.1 ± 0.9 and $-12.0 \pm 1.1 \text{ s}^{-1}$, respectively. The frameshifting efficiencies are 21.0%, 15.3% and 7.8%, respectively. The mean one-step unfolding forces of SF206, SF209 and SF211 with 4~5 pN/s loading rate are within a small range ($32.5 \pm 2.4 \text{ pN}$, $31.9 \pm 1.9 \text{ pN}$ and $30.8 \pm 3.2 \text{ pN}$, respectively). Thus, mechanical unfolding kinetics shows better correlation with frameshifting efficiency.

4.3.4 Minor-groove base triples enhance the pseudoknot mechanical stability and thereby stimulate frameshifting

Compared to SF206, A26C (SF217) and U27C (SF220) have a single residue mutation in loop 2, and show lower frameshifting efficiency and mechanical stability, consistent with the NMR structure of SF206 that A26 and U27 form minor-groove base triples with stem 1, respectively (Figure 4.1B). U27C mutation in SF220 is not expected to disrupt the single hydrogen bond formed between U27 and G15. The decrease in mean one-step unfolding force (from $32.5 \pm 2.4 \text{ pN}$ to $30.2 \pm 2.8 \text{ pN}$) with force loading rate of 4~5 pN/s and frameshifting efficiency (from 21.0% to 10.1%) as well as the increase in $\ln k_0$ values (from -15.9 ± 0.6 to $-14.0 \pm 0.7 \text{ s}^{-1}$) might be due to unfavourable stacking of C27 with A26. Thus, single-molecule mechanical unfolding studies can reveal the effect of subtle U to C substitution. As expected, A26C mutation in SF217 results in the disruption of A26·G4·C16 base triple and base stacking among A25, A26 and U27, which is consistent with the results that relatively larger reduction in mean unfolding force (from $32.5 \pm 2.4 \text{ pN}$ to $25.7 \pm 2.0 \text{ pN}$) and frameshifting efficiency

(from 21.0% to 6.6%) as well as the increase in $\ln k_0$ values (from -15.9 ± 0.6 to $-12.5 \pm 0.5 \text{ s}^{-1}$).

In SF229, the 9-nt adenine-rich loop 2 is substituted with a pyrimidine-rich sequence. Remarkably, SF229 shows the similar mean unfolding force ($25.6 \pm 2.3 \text{ pN}$) as SF217, suggesting the critical role of A26 in stabilizing the stem 1-loop 2 triplex and the pseudoknot. SF348 has a significantly shortened loop 2 with only three A residues. SF348 forms a pseudoknot and has a similar mean unfolding force ($25.9 \pm 2.1 \text{ pN}$) as SF229 and SF217, indicating that a shortened loop may cause the disruption of the minor-groove base triples. Taken together, our results suggest that loop 2 sequence and length can significantly affect pseudoknot mechanical unfolding and frameshifting efficiency.

4.3.5 One-step unfolding rupture force and unfolding rate are correlated to *in vitro* frameshifting efficiency

We compared the mean one-step unfolding rupture forces at 4~5, 10~13 and 18~21 pN/s to *in vitro* frameshifting efficiency, respectively (Figure 4.4A). The mean one-step unfolding rupture forces have a positive correlation with *in vitro* frameshifting efficiency for all force loading rates. The unfolding rate constant extracted from Bell model shows an inverse correlation with *in vitro* frameshifting efficiency (Figure 4.4B,C and Table 4.2). The distances between unfolded state and unfolding transition state ($x^\ddagger$) are about 2 nm for all pseudoknots (Figure 4.4D), which is in agreement with our previous results on unfolding of human telomerase RNA pseudoknot.^[35]

Our results indicate that although the mechanical stretching force applied by optical tweezers may not be the same as the mechanical action of a ribosome on a pseudoknot,

it facilitates the cooperative mechanical unfolding and unveils the stabilizing effects of the stem-stem coaxial stacking and loop-stem interactions on the component stems in a pseudoknot. Our studies suggest the cooperative mechanical unfolding properties determine the frameshifting efficiency.

Figure 4.4 (A) *In vitro* frameshifting efficiency^[142] versus mean unfolding force at around 5 pN/s force loading rate. (B) Unfolding kinetics of SF206 extracted from force distributions at different force loading rates using Dudko's method.^[62] Bell model is used to fit the force-dependent unfolding rates to extract rate constant at zero force and unfolding transition state positions. (C) *In vitro* frameshifting efficiency versus $\ln k_0$. (D) *In vitro* frameshifting efficiency versus unfolding transition positions.

Table 4.2 Summary of one-step mechanical unfolding kinetic parameters

Molecule	$\ln k_0$ (s ⁻¹)	$x^\ddagger$ (nm)	Adjusted r^2
SF206	-15.9 ± 0.6	2.1 ± 0.1	0.94
SF209	-15.1 ± 0.9	2.1 ± 0.1	0.91
SF211	-12.0 ± 1.1	1.7 ± 0.1	0.84
SF217	-12.5 ± 0.5	2.0 ± 0.1	0.96
SF220	-14.0 ± 0.7	2.0 ± 0.1	0.93
SF229	-13.3 ± 0.6	2.2 ± 0.1	0.94
SF348	-11.4 ± 0.8	1.9 ± 0.1	0.90

All uncertainties are reported as standard errors.

4.3.6 Thermal denaturation results suggest no correlation between thermal stability and *in vitro* frameshifting efficiency

In thermal denaturation experiments, all the melting curves are superimposable for heating and cooling curves, suggesting the transitions are reversible. In all three buffer conditions, two apparent transitions are observed for all pseudoknots (except SF348 in buffer with 0 mM NaCl). The pseudoknots and the control hairpin do not completely melt (Figure 4.5A,B) in buffer with 200 mM or 1 M NaCl. The melting profiles of the control hairpin imply that the transition with higher melting temperature corresponds to the thermal unfolding of stem 1 of the pseudoknots.

In buffer with 0 mM NaCl, pseudoknot structure does not form for SF348 as no obvious low-temperature transition is observed. Except for SF209 and SF211, the $T_{m, \text{stem 1}}$ of the pseudoknots are in a narrow range between 78 and 82 °C, which is close to the melting temperature of control hairpin (80.1 ± 0.1 °C). Consistent with the fact

that there are destabilizing mutations in the first base pair of stem 1, the $T_{m, \text{stem 1}}$ of SF209 and SF211 are significant lower than those of other pseudoknots. In 0 mM NaCl buffer, hyperchromicity of the first transition is about 2 folds as that of the second transition.

Our thermal melting experiment results indicate that the unfolding of stem 1 and stem 2, as well as the unfolding of tertiary structures and stem 1, are decoupled (Figure 4.5A,C), and the low-temperature transition of the pseudoknots corresponds to the melting of tertiary triplex structures and stem 2. Our thermal melting experiment results are in agreement with the course-grained simulation by Cho et al, although an apparent one-step melting transition was observed in the simulation.^[143] We reason that such inconsistency may be due to the fact that the simulation by Cho et al only considered the stability of stems and neglected the tertiary interactions between stems and loops, but the base triples may play important roles in the thermal melting pathway of SRV-1 frameshifting pseudoknots. We compared the $T_{m, \text{stem 1}}$ to *in vitro* frameshifting efficiency (Figure 4.5D), and no obvious correlation was observed. Due to the overlap of the tertiary structure and stem 2 melting transitions, and the limited data quality, we did not further extract the thermodynamic parameters.^[127, 133, 144]

Figure 4.5 (A) Normalized UV melting trajectories of pseudoknot SF206 in the buffers with 1M (red), 200 mM (green) and 0 mM (blue) NaCl, respectively. (B) Normalized UV melting trajectories of control hairpin in the buffers with 1M (red), 200 mM (green) and 0 mM (blue) NaCl, respectively. (C) Normalized UV melting trajectories of pseudoknots and control hairpin in the buffer with 0 mM NaCl. (D) Comparison of $T_{m,stem 1}$ to *in vitro* frameshifting efficiency.

4.3.7 Native PAGE results suggest that base triple formation makes the pseudoknot more compact

We carried out native PAGE experiment for pseudoknots and the control hairpin (Figure 4.6). The migration rate of control hairpin is significantly larger than the pseudoknots, which is consistent with the fact that control hairpin (stem 1 hairpin) is smaller than the pseudoknots. The migration rates of SF206, SF209 and SF211 are similar, presumably due to the fact that the loop 2 sequence of these three pseudoknots are the same with the stem 1-loop 2 triplex structure forming. SF220 has unfavorable

stacking in loop 2 and thus has a slightly slower migration rate. The migration rate of SF217 and SF229 are lowest due to the disruption of the base triple structures, which is consistent with the single-molecule experiment results. SF348 has a similar migration rate as SF220, probably because it has a shortened loop 2 resulting in a more compact pseudoknot structure even without tertiary interaction with stem 1. Our native PAGE experiment clearly suggests that base triple formation is correlated to the stem 1-loop 2 structural compactness.

Figure 4.6 Native PAGE results for RNA oligonucleotides. Samples in Lane 1 to 8 are SF206, SF209, SF211, SF217, SF220, SF229, SF348 and control hairpin, respectively.

4.4 Conclusions

Our results reveal that (1) the structural compactness and stability are correlated with the stem 1-loop 2 triplex formation, (2) unfolding rupture force has positive correlation with -1 frameshifting efficiency, and (3) unfolding rate constant is inversely proportional to frameshifting efficiency. Though the way mechanical unfolding executed by optical tweezers may not be the same as the RNA structure encounters ribosome in -1 frameshifting mechanism, the optical tweezers experiments unravel the

coupling unfolding of base pairs in stem 1 and the minor groove base triples, while in the thermal melting experiments the base triples and the stem 1 do not melt cooperatively. Together with the *in vitro* frameshifting efficiency measurement performed by our collaborators,^[142] our single-molecule and bulk experiment results suggest that such cooperative effect also exists in the ribosome-induced mechanical unfolding in the -1 ribosomal frameshifting process. Our studies provided detailed insights into the determinants of a pseudoknot structure affecting frameshifting efficiency, which may help us understand and manipulate the -1 frameshifting process.

4.5 Perspectives

The direction of mechanical stretching force loaded on a pseudoknot is determined by the handles labelling sites. The RNA/DNA hybrid handles are strictly linked to the 5' and 3' end of the backbone of the pseudoknot by the current sample synthesis method. To better mimic the helicase activity of ribosome on a pseudoknot, we may develop RNA-DNA conjugation methods that allow the linkage of dsDNA handles to alternative sites of an RNA structure. Similar strategies have been applied in the studies on single-molecule anisotropic mechanical unfolding pathways of protein and DNA.^[119, 145]

CHAPTER V

Conclusions and Future Perspectives

In this thesis, we introduced the biological importance of RNA mechanical properties, the basic physics of optical tweezers, and methodology of single-molecule force spectroscopy measurement using dual-laser optical tweezers. Two cases of studies using optical tweezers based single-molecule manipulation as well as bulk experiment techniques were also described in detail in Chapter 3 and 4.

In Chapter 3, the non-nearest neighbor structural destabilization effect of non-Watson-Crick U·U pair to U·C pair mutations in a human telomerase RNA hairpin was investigated using single-molecule force spectroscopy and thermal melting experiments. During the mechanical unfolding by force-ramp, a stretching force-induced misfolding is observed, while during the thermal melting experiments, a structural intermediate different from the force-induced misfolding structure is observed. The mechanical unfolding experiments unravel that the force-induced misfolding structure occurs more frequently upon the thermodynamically destabilizing U·U to U·C substitutions.

In Chapter 4, we compared our results of mechanical properties, thermal stability, and electrophoresis migration rates of SRV-1 RNA frameshifting pseudoknots, to the *in vitro* frameshifting efficiencies measured previously.^[142] This work suggests that the one-step mechanical unfolding rate constant of frameshifting pseudoknots is inversely proportional to frameshifting efficiency, while the melting temperatures from the step-wise thermal melting pathway do not show obvious correlation to frameshifting efficiency. Our studies suggest that the mechanical stability of SRV-1 RNA

frameshifting pseudoknots is determined by the coupling of base pairs in stem 1 and the minor groove base triples. Such kinetic cooperative effects may also exist in the ribosome-induced mechanical unfolding during the -1 ribosomal frameshifting process, and determine the frameshifting efficiency.

Conclusively, by using single-molecule and bulk experiments, this thesis illustrates the chemical physical and biological importance of non-Watson-Crick and triplex structures in RNAs. In future, we may also investigate the single-molecule force spectroscopies of other biological functional RNAs, which may contain other non-Watson-Crick base pairs besides U·U and U·C, or other form of tertiary structures, such as G quadruplex. We may also develop microfluidic and bio-conjugation methods, to study single-molecule RNA-ligand binding force spectroscopy, as well as the anisotropic effect on mechanical unfolding of RNA structures.

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] E. Pennisi, DNA study forces rethink of what it means to be a gene. *Science* **2007**, *316*, 1556-1557.
- [3] S. Guil, M. Esteller, RNA-RNA interactions in gene regulation: the coding and noncoding players. *Trends Biochem Sci* **2015**, *40*, 248-256.
- [4] M. Guttman, I. Amit, M. Garber, C. French, M. F. Lin, D. Feldser, M. Huarte, O. Zuk, B. W. Carey, J. P. Cassady, M. N. Cabili, R. Jaenisch, T. S. Mikkelsen, T. Jacks, N. Hacohen, B. E. Bernstein, M. Kellis, A. Regev, J. L. Rinn, E. S. Lander, Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **2009**, *458*, 223-227.
- [5] K. S. Kosik, Molecular biology: Circles reshape the RNA world. *Nature* **2013**, *495*, 322-324.
- [6] T. B. Hansen, T. I. Jensen, B. H. Clausen, J. B. Bramsen, B. Finsen, C. K. Damgaard, J. Kjems, Natural RNA circles function as efficient microRNA sponges. *Nature* **2013**, *495*, 384-388.
- [7] H. Grosjean, Modification and editing of RNA: historical overview and important facts to remember. in *Fine-Tuning of RNA Functions by Modification and Editing*, Springer, **2005**, pp. 1-22.
- [8] G. Jia, Y. Fu, X. Zhao, Q. Dai, G. Zheng, Y. Yang, C. Yi, T. Lindahl, T. Pan, Y.-G. Yang, N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat Chem Biol* **2011**, *7*, 885-887.
- [9] W. Stephenson, A. Colvin, A. A. Chen, P. T. Li, 102 Force and temperature dependent folding of a 2-base-pair RNA kissing complex. *J Biomol Struct Dyn* **2015**, *33 Suppl 1*, 65.
- [10] P. T. Li, Analysis of diffuse K⁺ and Mg²⁺ ion binding to a two-base-pair kissing complex by single-molecule mechanical unfolding. *Biochemistry* **2013**, *52*, 4991-5001.

- [11] W. Stephenson, P. N. Asare-Okai, A. A. Chen, S. Keller, R. Santiago, S. A. Tenenbaum, A. E. Garcia, D. Fabris, P. T. Li, The essential role of stacking adenines in a two-base-pair RNA kissing complex. *J Am Chem Soc* **2013**, *135*, 5602-5611.
- [12] P. T. Li, I. Tinoco, Jr., Mechanical unfolding of two DIS RNA kissing complexes from HIV-1. *J Mol Biol* **2009**, *386*, 1343-1356.
- [13] P. T. Li, C. Bustamante, I. Tinoco, Jr., Unusual mechanical stability of a minimal RNA kissing complex. *Proc Natl Acad Sci U S A* **2006**, *103*, 15847-15852.
- [14] G. Devi, Y. Zhou, Z. Zhong, D. F. Toh, G. Chen, RNA triplexes: from structural principles to biological and biotech applications. *Wiley Interdiscip Rev RNA* **2015**, *6*, 111-128.
- [15] 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.
- [16] K. L. Frieda, S. M. Block, Direct observation of cotranscriptional folding in an adenine riboswitch. *Science* **2012**, *338*, 397-400.
- [17] I. Tinoco, Jr., C. Bustamante, How RNA folds. *J Mol Biol* **1999**, *293*, 271-281.
- [18] R. M. Mitton-Fry, S. J. DeGregorio, J. Wang, T. A. Steitz, J. A. Steitz, Poly(A) tail recognition by a viral RNA element through assembly of a triple helix. *Science* **2010**, *330*, 1244-1247.
- [19] W. Cheng, S. Dumont, I. Tinoco, Jr., C. Bustamante, NS3 helicase actively separates RNA strands and senses sequence barriers ahead of the opening fork. *Proc Natl Acad Sci U S A* **2007**, *104*, 13954-13959.
- [20] 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.
- [21] C. Bustamante, W. Cheng, Y. X. Mejia, Revisiting the central dogma one molecule at a time. *Cell* **2011**, *144*, 480-497.
- [22] I. Tinoco, Jr., Force as a useful variable in reactions: unfolding RNA. *Annu Rev Biophys Biomol Struct* **2004**, *33*, 363-385.

- [23] J. A. Jaeger, D. H. Turner, M. Zuker, Improved predictions of secondary structures for RNA. *Proc Natl Acad Sci U S A* **1989**, *86*, 7706-7710.
- [24] M. Wu, J. A. McDowell, D. H. Turner, A Periodic-Table of Symmetrical Tandem Mismatches in RNA. *Biochemistry* **1995**, *34*, 3204-3211.
- [25] 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.
- [26] S. Rouskin, M. Zubradt, S. Washietl, M. Kellis, J. S. Weissman, Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature* **2014**, *505*, 701-705.
- [27] I. Tinoco, G. Chen, X. Qu, RNA reactions one molecule at a time. *Cold Spring Harb Perspect Biol* **2010**, *2*, a003624.
- [28] J. D. Wen, M. Manoskas, 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.
- [29] 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.
- [30] I. Tinoco, Jr., R. L. Gonzalez, Jr., Biological mechanisms, one molecule at a time. *Genes Dev* **2011**, *25*, 1205-1231.
- [31] 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.
- [32] 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.
- [33] 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.

- [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] 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. 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.
- [37] 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.
- [38] 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.
- [39] M. H. Larson, W. J. Greenleaf, R. Landick, S. M. Block, Applied force reveals mechanistic and energetic details of transcription termination. *Cell* **2008**, 132, 971-982.
- [40] R. V. Dalal, M. H. Larson, K. C. Neuman, J. Gelles, R. Landick, S. M. Block, Pulling on the nascent RNA during transcription does not alter kinetics of elongation or ubiquitous pausing. *Mol Cell* **2006**, 23, 231-239.
- [41] L. Green, C. H. Kim, C. Bustamante, I. Tinoco, Characterization of the mechanical unfolding of RNA pseudoknots. *J Mol Biol* **2008**, 375, 511-528.
- [42] S. Yan, J. D. Wen, C. Bustamante, I. Tinoco, Jr., Ribosome excursions during mRNA translocation mediate broad branching of frameshift pathways. *Cell* **2015**, 160, 870-881.
- [43] 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.
- [44] J. Viereggs, 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.

- [45] X. Zhuang, Single-molecule RNA science. *Annu Rev Biophys Biomol Struct* **2005**, *34*, 399-414.
- [46] W. J. Greenleaf, M. T. Woodside, S. M. Block, High-resolution, single-molecule measurements of biomolecular motion. *Annu Rev Bioph Biom* **2007**, *36*, 171-190.
- [47] M. C. Zhong, L. Gong, J. H. Zhou, Z. Q. Wang, Y. M. Li, Optical trapping of red blood cells in living animals with a water immersion objective. *Opt Lett* **2013**, *38*, 5134-5137.
- [48] M. C. Zhong, X. B. Wei, J. H. Zhou, Z. Q. Wang, Y. M. Li, Trapping red blood cells in living animals using optical tweezers. *Nat Commun* **2013**, *4*, 1768.
- [49] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco, Jr., C. Bustamante, Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. *Nature* **2005**, *437*, 231-234.
- [50] 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.
- [51] S. B. Smith, Y. J. Cui, C. Bustamante, Optical-trap force transducer that operates by direct measurement of light momentum. *Method Enzymol* **2003**, *361*, 134-162.
- [52] K. Svoboda, S. M. Block, Biological Applications of Optical Forces. *Annu Rev Bioph Biom* **1994**, *23*, 247-285.
- [53] 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.
- [54] M. D. Wang, H. Yin, R. Landick, J. Gelles, S. M. Block, Stretching DNA with optical tweezers. *Biophys J* **1997**, *72*, 1335-1346.
- [55] M. Manosas, F. Ritort, Thermodynamic and kinetic aspects of RNA pulling experiments. *Biophys J* **2005**, *88*, 3224-3242.
- [56] C. Bustamante, J. F. Marko, E. D. Siggia, S. Smith, Entropic elasticity of lambda-phage DNA. *Science* **1994**, *265*, 1599-1600.

- [57] Y. Seol, G. M. Skinner, K. Visscher, Elastic properties of a single-stranded charged homopolymeric ribonucleotide. *Phys Rev Lett* **2004**, *93*, 118102.
- [58] P. J. Elms, J. D. Chodera, C. J. Bustamante, S. Marqusee, Limitations of constant-force-feedback experiments. *Biophys J* **2012**, *103*, 1490-1499.
- [59] S. A. McKinney, C. Joo, T. Ha, Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophys J* **2006**, *91*, 1941-1951.
- [60] I. Tinoco, C. Bustamante, The effect of force on thermodynamics and kinetics of single molecule reactions. *Biophys Chem* **2002**, *101*, 513-533.
- [61] J. M. Matthews, A. R. Fersht, Exploring the energy surface of protein folding by structure-reactivity relationships and engineered proteins: observation of Hammond behavior for the gross structure of the transition state and anti-Hammond behavior for structural elements for unfolding/folding of barnase. *Biochemistry* **1995**, *34*, 6805-6814.
- [62] O. K. Dudko, G. Hummer, A. Szabo, Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proc Natl Acad Sci U S A* **2008**, *105*, 15755-15760.
- [63] O. K. Dudko, G. Hummer, A. Szabo, Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys Rev Lett* **2006**, *96*, 108101.
- [64] 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.
- [65] Y.-J. Wu, C.-H. Wu, A. Y.-C. Yeh, J.-D. Wen, Folding a stable RNA pseudoknot through rearrangement of two hairpin structures. *Nucleic Acids Res* **2014**, *42*, 4505-4515.
- [66] J. Kim, C.-Z. Zhang, X. Zhang, T. A. Springer, A mechanically stabilized receptor-ligand flex-bond important in the vasculature. *Nature* **2010**, *466*, 992-995.
- [67] M. R. Shirts, E. Bair, G. Hooker, V. S. Pande, Equilibrium free energies from nonequilibrium measurements using maximum-likelihood methods. *Phys Rev Lett* **2003**, *91*.
- [68] N. B. Leontis, A. Lescaute, E. Westhof, The building blocks and motifs of RNA architecture. *Curr Opin Struct Biol* **2006**, *16*, 279-287.

- [69] 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.
- [70] 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.
- [71] J. Santalucia, R. Kierzek, D. H. Turner, Stabilities of consecutive A·C, C·C, G·G, U·C, and U·U mismatches in RNA Internal loops - Evidence for stable hydrogen-bonded U·U and C·C+ pairs. *Biochemistry* **1991**, *30*, 8242-8251.
- [72] G. Chen, B. M. Znosko, S. D. Kennedy, T. R. Krugh, D. H. Turner, Solution structure of an RNA internal loop with three consecutive sheared GA pairs. *Biochemistry* **2005**, *44*, 2845-2856.
- [73] 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.
- [74] G. Chen, B. M. Znosko, X. Jiao, D. H. Turner, Factors affecting thermodynamic stabilities of RNA 3 x 3 internal loops. *Biochemistry* **2004**, *43*, 12865-12876.
- [75] G. Chen, D. H. Turner, Consecutive GA pairs stabilize medium-size RNA internal loops. *Biochemistry* **2006**, *45*, 4025-4043.
- [76] G. Chen, S. D. Kennedy, J. Qiao, T. R. Krugh, D. H. Turner, An alternating sheared AA pair and elements of stability for a single sheared purine-purine pair flanked by sheared GA pairs in RNA. *Biochemistry* **2006**, *45*, 6889-6903.
- [77] N. Shankar, S. D. Kennedy, G. Chen, T. R. Krugh, D. H. Turner, The NMR structure of an internal loop from 23S ribosomal RNA differs from its structure in crystals of 50S ribosomal subunits. *Biochemistry* **2006**, *45*, 11776-11789.
- [78] 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.
- [79] 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.

- [80] 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.
- [81] M. K. Lee, M. Gal, L. Frydman, G. Varani, Real-time multidimensional NMR follows RNA folding with second resolution. *Proc Natl Acad Sci U S A* **2010**, *107*, 9192-9197.
- [82] Y. X. Wang, S. Huang, D. E. Draper, Structure of a U·U pair within a conserved ribosomal RNA hairpin. *Nucleic Acids Res* **1996**, *24*, 2666-2672.
- [83] A. Kumar, H. Park, P. Fang, R. Parkesh, M. Guo, K. W. Nettles, M. D. Disney, Myotonic dystrophy type 1 RNA crystal structures reveal heterogeneous 1 x 1 nucleotide UU internal loop conformations. *Biochemistry* **2011**, *50*, 9928-9935.
- [84] K. T. Tycowski, M. D. Shu, S. Borah, M. Shi, J. A. Steitz, Conservation of a triple-helix-forming RNA stability element in noncoding and genomic RNAs of diverse viruses. *Cell Rep* **2012**, *2*, 26-32.
- [85] J. A. Brown, M. L. Valenstein, T. A. Yario, K. T. Tycowski, J. A. Steitz, Formation of triple-helical structures by the 3'-end sequences of MALAT1 and MENbeta noncoding RNAs. *Proc Natl Acad Sci U S A* **2012**, *109*, 19202-19207.
- [86] J. E. Wilusz, C. K. JnBaptiste, L. Y. Lu, C. D. Kuhn, L. Joshua-Tor, P. A. Sharp, A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev* **2012**, *26*, 2392-2407.
- [87] K. J. Baeyens, H. L. De Bondt, S. R. Holbrook, Structure of an RNA double helix including uracil-uracil base pairs in an internal loop. *Nat Struct Biol* **1995**, *2*, 56-62.
- [88] S. E. Lietzke, C. L. Barnes, J. A. Berglund, C. E. Kundrot, The structure of an RNA dodecamer shows how tandem U-U base pairs increase the range of stable RNA structures and the diversity of recognition sites. *Structure* **1996**, *4*, 917-930.
- [89] B. Furtig, C. Richter, J. Wohnert, H. Schwalbe, NMR spectroscopy of RNA. *Chembiochem* **2003**, *4*, 936-962.
- [90] Z. Du, J. Yu, N. B. Ulyanov, R. Andino, T. L. James, Solution structure of a consensus stem-loop D RNA domain that plays important roles in regulating translation and replication in enteroviruses and rhinoviruses. *Biochemistry* **2004**, *43*, 11959-11972.

- [91] S. Schroeder, J. Kim, D. H. Turner, G·A and U·U mismatches can stabilize RNA internal loops of three nucleotides. *Biochemistry* **1996**, *35*, 16105-16109.
- [92] B. N. Bourdelat-Parks, R. M. Wartell, Thermodynamics of RNA duplexes with tandem mismatches containing a uracil-uracil pair flanked by C·G/G·C or G·C/A·U closing base pairs. *Biochemistry* **2005**, *44*, 16710-16717.
- [93] M. E. Christiansen, B. M. Znosko, Thermodynamic characterization of the complete set of sequence symmetric tandem mismatches in RNA and an improved model for predicting the free energy contribution of sequence asymmetric tandem mismatches. *Biochemistry* **2008**, *47*, 4329-4336.
- [94] M. E. Christiansen, B. M. Znosko, Thermodynamic characterization of tandem mismatches found in naturally occurring RNA. *Nucleic Acids Res* **2009**, *37*, 4696-4706.
- [95] L. G. Laing, T. C. Gluick, D. E. Draper, Stabilization of RNA structure by Mg ions. Specific and non-specific effects. *J Mol Biol* **1994**, *237*, 577-587.
- [96] C. A. Theimer, Y. Wang, D. W. Hoffman, H. M. Krisch, D. P. Giedroc, Non-nearest neighbor effects on the thermodynamics of unfolding of a model mRNA pseudoknot. *J Mol Biol* **1998**, *279*, 545-564.
- [97] P. L. Nixon, D. P. Giedroc, Equilibrium unfolding (folding) pathway of a model H-type pseudoknotted RNA: the role of magnesium ions in stability. *Biochemistry* **1998**, *37*, 16116-16129.
- [98] S. Nakano, A. L. Cerrone, P. C. Bevilacqua, Mechanistic characterization of the HDV genomic ribozyme: classifying the catalytic and structural metal ion sites within a multichannel reaction mechanism. *Biochemistry* **2003**, *42*, 2982-2994.
- [99] J. A. McDowell, D. H. Turner, Investigation of the structural basis for thermodynamic stabilities of tandem GU mismatches: solution structure of (rGAGGUCUC)₂ by two-dimensional NMR and simulated annealing. *Biochemistry* **1996**, *35*, 14077-14089.
- [100] P. N. Borer, B. Dengler, I. Tinoco, Jr., O. C. Uhlenbeck, Stability of ribonucleic acid double-stranded helices. *J Mol Biol* **1974**, *86*, 843-853.
- [101] 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.

- [102] L. R. Comolli, I. Smirnov, L. Xu, E. H. Blackburn, T. L. James, A molecular switch underlies a human telomerase disease. *Proc Natl Acad Sci U S A* **2002**, *99*, 16998-17003.
- [103] J. D. Puglisi, I. Tinoco, Jr., Absorbance melting curves of RNA. *Methods Enzymol* **1989**, *180*, 304-325.
- [104] K. M. Weeks, D. M. Crothers, Major groove accessibility of RNA. *Science* **1993**, *261*, 1574-1577.
- [105] J. S. Reuter, D. H. Mathews, RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics* **2010**, *11*, 129.
- [106] D. H. Mathews, RNA secondary structure analysis using RNAstructure. *Curr Protoc Bioinformatics* **2014**, *46*, 12.16.11-12.16.25.
- [107] C. Cecconi, E. A. Shank, C. Bustamante, S. Marqusee, Direct observation of the three-state folding of a single protein molecule. *Science* **2005**, *309*, 2057-2060.
- [108] Y. Seol, G. M. Skinner, K. Visscher, A. Buhot, A. Halperin, Stretching of homopolymeric RNA reveals single-stranded helices and base-stacking. *Phys Rev Lett* **2007**, *98*, 158103.
- [109] S. Cao, S. J. Chen, Biphasic folding kinetics of RNA pseudoknots and telomerase RNA activity. *J Mol Biol* **2007**, *367*, 909-924.
- [110] M. Hengesbach, N. K. Kim, J. Feigon, M. D. Stone, Single-molecule FRET reveals the folding dynamics of the human telomerase RNA pseudoknot domain. *Angew Chem Int Ed Engl* **2012**, *51*, 5876-5879.
- [111] A. Garai, Y. Zhang, O. K. Dudko, Conformational dynamics through an intermediate. *J Chem Phys* **2014**, *140*, 135101.
- [112] 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.
- [113] 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.

- [114] N. K. Kim, Q. Zhang, J. Zhou, C. A. Theimer, R. D. Peterson, J. Feigon, Solution structure and dynamics of the wild-type pseudoknot of human telomerase RNA. *J Mol Biol* **2008**, *384*, 1249-1261.
- [115] Y. G. Yingling, B. A. Shapiro, Dynamic behavior of the telomerase RNA hairpin structure and its relationship to dyskeratosis congenita. *J Mol Biol* **2005**, *348*, 27-42.
- [116] J. L. Chen, C. W. Greider, Functional analysis of the pseudoknot structure in human telomerase RNA. *Proc Natl Acad Sci U S A* **2005**, *102*, 8080-8085; discussion 8077-8089.
- [117] A. Bosco, J. Camunas-Soler, F. Ritort, Elastic properties and secondary structure formation of single-stranded DNA at monovalent and divalent salt conditions. *Nucleic Acids Res* **2014**, *42*, 2064-2074.
- [118] K. T. Schroeder, P. Daldrop, S. A. McPhee, D. M. Lilley, Structure and folding of a rare, natural kink turn in RNA with an A*A pair at the 2b*2n position. *RNA* **2012**, *18*, 1257-1266.
- [119] B. Jagannathan, P. J. Elms, C. Bustamante, S. Marqusee, Direct observation of a force-induced switch in the anisotropic mechanical unfolding pathway of a protein. *Proc Natl Acad Sci U S A* **2012**, *109*, 17820-17825.
- [120] R. B. Best, S. B. Fowler, J. L. Toca-Herrera, J. Clarke, A simple method for probing the mechanical unfolding pathway of proteins in detail. *Proc Natl Acad Sci U S A* **2002**, *99*, 12143-12148.
- [121] H. Li, M. Carrion-Vazquez, A. F. Oberhauser, P. E. Marszalek, J. M. Fernandez, Point mutations alter the mechanical stability of immunoglobulin modules. *Nat Struct Biol* **2000**, *7*, 1117-1120.
- [122] X. Xu, H. Ge, C. Gu, Y. Q. Gao, S. S. Wang, B. J. Thio, J. T. Hynes, X. S. Xie, J. Cao, Modeling spatial correlation of DNA deformation: DNA allostery in protein binding. *J Phys Chem B* **2013**, *117*, 13378-13387.
- [123] S. Kim, E. Brostromer, D. Xing, J. Jin, S. Chong, H. Ge, S. Wang, C. Gu, L. Yang, Y. Q. Gao, X. D. Su, Y. Sun, X. S. Xie, Probing allostery through DNA. *Science* **2013**, *339*, 816-819.
- [124] T. Jacks, H. E. Varmus, Expression of the Rous-Sarcoma Virus Pol Gene by Ribosomal Frameshifting. *Science* **1985**, *230*, 1237-1242.

- [125] T. Jacks, H. D. Madhani, F. R. Masiarz, H. E. Varmus, Signals for Ribosomal Frameshifting in the Rous-Sarcoma Virus Gag-Pol Region. *Cell* **1988**, *55*, 447-458.
- [126] Y. C. 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.
- [127] 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.
- [128] I. Brierley, F. J. Dos Ramos, Programmed ribosomal frameshifting in HIV-1 and the SARS-CoV. *Virus Res* **2006**, *119*, 29-42.
- [129] T. Jacks, K. Townsley, H. E. Varmus, J. Majors, Two efficient ribosomal frameshifting events are required for synthesis of mouse mammary tumor virus gag-related polyproteins. *Proc Natl Acad Sci U S A* **1987**, *84*, 4298-4302.
- [130] J. D. Dinman, M. J. Ruiz-Echevarria, S. W. Peltz, Translating old drugs into new treatments: ribosomal frameshifting as a target for antiviral agents. *Trends Biotechnol* **1998**, *16*, 190-196.
- [131] I. Brierley, P. Digard, S. C. Inglis, Characterization of an Efficient Coronavirus Ribosomal Frameshifting Signal - Requirement for an RNA Pseudoknot. *Cell* **1989**, *57*, 537-547.
- [132] S. Cao, S. J. Chen, Predicting ribosomal frameshifting efficiency. *Phys Biol* **2008**, *5*, 016002.
- [133] D. P. Giedroc, P. V. Cornish, Frameshifting RNA pseudoknots: structure and mechanism. *Virus Res* **2009**, *139*, 193-208.
- [134] J. D. Dinman, Mechanisms and implications of programmed translational frameshifting. *Wiley Interdiscip Rev RNA* **2012**, *3*, 661-673.
- [135] E. P. Plant, K. L. Jacobs, J. W. Harger, A. Meskauskas, J. L. Jacobs, J. L. Baxter, A. N. Petrov, J. D. Dinman, The 9-Å solution: how mRNA pseudoknots promote efficient programmed -1 ribosomal frameshifting. *RNA* **2003**, *9*, 168-174.
- [136] 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.

- [137] T. M. Schmeing, R. M. Voorhees, A. C. Kelley, Y.-G. Gao, F. V. Murphy, J. R. Weir, V. Ramakrishnan, The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science* **2009**, 326, 688-694.
- [138] 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.
- [139] 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.
- [140] K. H. White, M. Orzechowski, D. Fourmy, K. Visscher, Mechanical unfolding of the beet western yellow virus -1 frameshift signal. *J Am Chem Soc* **2011**, 133, 9775-9782.
- [141] P. J. A. Michiels, A. A. M. Versleijen, P. W. Verlaan, C. W. A. 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.
- [142] 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.
- [143] S. S. Cho, D. L. Pincus, D. Thirumalai, Assembly mechanisms of RNA pseudoknots are determined by the stabilities of constituent secondary structures. *Proc Natl Acad Sci U S A* **2009**, 106, 17349-17354.
- [144] P. V. Cornish, D. P. Giedroc, Pairwise coupling analysis of helical junction hydrogen bonding interactions in luteoviral RNA pseudoknots. *Biochemistry* **2006**, 45, 11162-11171.
- [145] S. Selvam, Z. Yu, H. Mao, Exploded view of higher order G-quadruplex structures through click-chemistry assisted single-molecule mechanical unfolding. *Nucleic Acids Res* **2016**, 44, 45.

LIST OF PUBLICATIONS

1. Devi, G., Zhou, Y., Zhong, Z., Toh, D. F. K., & Chen, G. (2015). RNA triplexes: from structural principles to biological and biotech applications. *Wiley Interdisciplinary Reviews: RNA*, 6(1), 111-128.
2. Zhong, Z., Soh, L. H., Lim, M. H., & Chen, G. (2015). A U·U Pair-to-U·C Pair Mutation-Induced RNA Native Structure Destabilisation and Stretching-Force-Induced RNA Misfolding. *ChemPlusChem*, 80(8), 1267-1278.
3. Zhong, Z., Yang, L., Zhang, H., Shi, J., Vandana, J., Lam, D. T. U. H., Olsthoorn, R. C. L., Lu, L., & Chen, G. (2016) Tertiary Structure Formation and Energetic Coupling in Simian Retrovirus Type 1 RNA Pseudoknot Increase Mechanical Stability and -1 Ribosomal Frameshifting Efficiency. Submitted.