IMPLICATIONS OF THE SHAPE OF TELOMERASE RNA

Vijay Ganesh Sekaran

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Eshelman School of Pharmacy (Pharmaceutical Sciences).

Chapel Hill
2012

Approved By:

Michael B. Jarstfer, Ph.D.

Andrew Lee, Ph.D.

Jian Liu, Ph.D.

Shawn Ahmed, Ph.D.

Howard Fried, Ph.D.

ABSTRACT

VIJAY GANESH SEKARAN: Implications of the SHAPE of Telomerase RNA
(Under the direction of Michael B. Jarstfer, Ph.D.)

Telomerase is a specialized ribonucleoprotein complex that uses its RNA subunit to add six nucleotide G-rich repeats to the telomeric DNA at the end of our chromosomes, subsequently controlling cellular lifespan. Though telomerase is undetectable in somatic cells, it is up-regulated in approximately 90% of cancer cells. Crucial to enzyme activity is telomerase RNA (TR), whose diverse secondary structure is intrinsic for biological function.

Due to the size of long RNAs and limitations of RNA structure determination techniques, high resolution structural data for full length TRs has been unavailable. Utilizing a novel chemical footprinting approach SHAPE, that gives high resolution structural data for every nucleotide in an RNA, we examined the structure of *Tetrahymena thermophila* (tTR) and human (hTR) TRs. The research presented here describes important structural aspects of these two telomerase RNAs.

Previous studies have determined the solution structure of tTR via SHAPE. In the first part of this dissertation, we examined tTR via differential SHAPE to identify nucleotides that serve as molecular switches for RNA folding. We also analyzed two structural motifs from the stem IV helix, a portion of tTR that is essential for TERT

binding and catalytic activity. Results from these studies indicate that local conformational flexibility impacts telomerase function.

In part two, we determined the protein-free structure of human telomerase RNA. We initially interrogated the structure of the pseudoknot, CR4-CR5, and CR7 domains and mutants that give rise to pre-mature aging diseases showing that changes in structure instigate disease progression. We then determined the structure of the full length hTR, demonstrating that the pseudoknot forms an elongated hairpin, the template is base paired, and that the 5' RNA forms a G-quadraplex. We finally used the hTR structural data to identify the binding site of aminoglycosides that inhibit telomerase assemblage and activity by binding to hTR.

As a principal component of telomerase, the contribution of TR structural information to RNA and telomerase biology is extremely important. With these data, comes the potential to develop rational approaches to modulate telomerase activity in aging and cancer drug platforms that take advantage of telomerase as a universal cancer target.

ACKNOWLEDGEMENT

During my time at The University of North Carolina at Chapel Hill, I have been fortunate to be able to pursue my passions for research and science. While much of graduate school and the completion of the doctoral dissertation is an individual journey that requires courage and belief in oneself, without the help and support of my friends and family, I would have never been able to complete this project and achieve my goals. Therefore, I am forever indebted to the many people who have helped shape me as a scientist and as an individual.

I would very much like to thank my advisor, Dr. Michael Jarstfer, for his guidance and support throughout my graduate career. He has motivated and pushed me to be creative, independent, and develop critical thinking skills that will be important for the future. His enthusiastic attitude towards learning and science has created a rich and productive learning environment where I have been challenged to reach my full potential daily. He has also been incredibly supportive of my interest in teaching, allowing me to mentor undergraduates and rotation students essentially every semester during my time in his lab. I have become the well-rounded mentor that I am because I myself have had fantastic and inspiring mentor. I am very grateful to have had Mike as an advisor, and am glad that I have been able to cultivate both a scientific relationship and friendship with him while a graduate student in his lab.

I would like to thank my committee members, Dr. Andrew Lee, Dr. Shawn Ahmed, Dr. Howard Fried, and Dr. Jian Liu for their constructive criticism, valuable insight, and mentorship during my graduate career.

I would like to thank all of the members, past and present, of the Jarstfer Lab for their compassion, friendship, and making work enjoyable. There is a lot more to graduate school and working in the lab and getting results, and having great friends to work with can make all the difference. On days when everything is going wrong, it is nice to look across the lab, see a familiar face, and know that someone else not only knows what you are going through, but also knows exactly what to say to make you feel better. Specifically, I would like to thank Dr. Joana Soares and Carrie-Ann Gordon, both of whom have been incredibly supportive scientifically and personally while I have been at UNC. Over the years, we have all had so much fun at work and outside of it, and I am lucky to have you both as friends.

I would also like to thank the many friends I have made over the years who I have always been able to count on and have been a solid support system to me. I am truly blessed to have met some absolutely amazing people. Whether I know you from high school, Georgia Tech, UNC, or the Chapel Hill area, each one of you has profoundly impacted my life. There is no way that I would be the person I am today without you all, and for that I am very thankful.

Finally, and most importantly, I would like to thank my Dad, Mom, and brother, Anand. You guys have always believed in me in all the decisions I have made throughout my life. I appreciate your unwaivering support when I have faced challenges and ability to brighten my day when I have been down. You have instilled in me the

strength, confidence, and determination to pursue and attain my goals. To my family, who has sacrificed so much for me over the years, I am forever grateful for your continuous encouragement and love.

TABLE OF CONTENTS

LIST OF TABLES xi

LIST OF FIGURES xii

LIST OF ABBREVIATIONS xiv

Chapter

I. TELOMERE AND RNA STRUCTRE 1

A. Telomere Maintenance and Disease 1

B. Telomeres and Telomerase 3

C. Human Telomerase Reverse Transctipase (hTERT) 5

D. Human Telomerase RNA (hTR) 7

E. Dyskeratosis Congenita 13

F. The Importance of RNA Architecture 14

G. Selective 2'-Hydoxyl Acylation Analyzed by Primer Extension (SHAPE) 16

H. Thesis overview 19

II. DIFFERENTIAL SHAPE AND ANALYSIS OF tTR STEM IV MUTANTS 21

A. Introduction 21

B. Results 28

1. Addition of 3' flanking region to tTR for SHAPE experiments 28

2. IA and 1M7 modification of tTR 28

3. 125-127 U Mutants 31

4. 132-138 Mutants 35

C. Discussion 39

D. Materials and Methods 44

1. Oligonucleotides and chemicals 44

2. tTR and tTERT expression plasmids 45

3. PCR construction of dsDNA tTR construct 46

4. *In vitro* transcription and purification of tTR 46

5. Translation of tTERT and telomerase assembly 47

6. Immunoprecipitation 48

7. Direct telomerase activity assay 49

8. IA, NMIA, and 1M7 hit reactions 50

9. Superscript III primer extension 50

10. Sequencing gel electrophoresis 51

11. SAFA data analysis and data normalization 52

III. SHAPE ANALYSIS OF THE PSEUDOKNOT, CR4-CR5, AND CR7 DOMAINS OF hTR 53

A. Introduction 53

B. Results 58

1. Addition of 5' and 3' flanking regions to hTR domains for SHAPE 58

2. Wildtype and ΔU177 mutant adopt pseudoknot structures 58

3. DC pseudoknot is actually a hairpin 60

4. ΔU177 and DC mutations to pseudoknot

diminish telomerase activity 62
5. Wildtype CR4-CR5 domain.....63
6. G305A and G322A CR4-CR5 domains.....64
7. Wildtype CR7 domain 66
8. G408C CR7 domain.....68
C. Discussion.....69
D. Materials and Methods.....81
1. Oligonucleotides and Chemicals.....82
2. hTR and hTERT expression plasmids 82
3. Cells and cell culture.....83
4. PCR construction of dsDNA pseudoknot constructs.....83
5. PCR construction of dsDNA CR4-CR5 domain constructs 84
6. PCR construction of dsDNA CR7 domain constructs.....85
7. *In vitro* transcription and purification of domain RNAs.....86
8. NMIA and 1M7 hit reactions.....86
9. Superscript III primer extension 87
10. Sequencing gel electrophoresis.....88
11. SAFA data analysis and data normalization.....88
12. Super telomerase extracts preparation 89
13. Direct telomerase assay.....90
IV. SHAPE ANALYSIS OF THE SOLUTION STRUCTURE OF hTR 92
A. Introduction.....92
B. Results.....93

1. Addition of 3'flanking region to hTR for SHAPE.....93
2. SHAPE of hTR in solution 94
a. 5' hTR and P1 99
b. Template and extended P1 helix.....100
c. Extended P2 helix, No P3 102
d. Hypervariable region and CR4-CR5 domain.....103
e. Box H/ACA and CR7 domains.....106
3. Aminoglycosides prevent telomerase assemblage.....107
4. Identifying binding sites of Nea 5-O and Kan 6" 108
C. Discussion.....111
D. Materials and Methods.....128
1. Oligonucleotides and chemicals 128
2. hTR and hTERT expression plasmids 129
3. PCR construction of dsDNA hTR construct.....129
4. *In vitro* transcription and purification of flanked hTR.....130
5. SHAPE 1M7 Hit Reactions 131
6. Superscript III primer extension 131
7. Sequencing gel electrophoresis.....132
8. SAFA data analysis and normalization.....133
9. *In vitro* transcription and purification of hTR.....133
10. Telomerase Assemblage Assay.....134
11. Nea 5-O and Kan-6" binding to hTR studies.....136
REFERENCES 137

LIST OF TABLES

Table

2.1 Primers for cloning of tTR 46

2.2 SHAPE primers for mapping tTR 50

3.1 Primers for making pseudoknot constructs 84

3.2 Primers for cloning CR4-CR5 domains 84

3.3 Primers for making CR7 domains 85

4.1 Primers for cloning of hTR 129

4.2 SHAPE primers for mapping hTR 131

LIST OF FIGURES

Figures

1.1 Antagonistic effects of telomerase in aging and cancer 2

1.2 Mechanism of telomere primer extension by telomerase 4

1.3 Organization of hTERT domains and their functions 6

1.4 Phylogenetically derived secondary structure of human telomerase RNA (hTR) 8

1.5 Secondary structure of hTR with sequence 10

1.6 RNA structured cassette for SHAPE 17

1.7 SHAPE from beginning to end 18

2.1 Phylogenetic secondary structure or tTR 23

2.2 tTR secondary structure 27

2.3 SHAPE reagents: 1M7 and IA 29

2.4 SHAPE of tTR with IA and 1M7 30

2.5 Nucleotides in tTR with differential SHAPE reactivity when using IA and 1M7 31

2.6 tTR stem IV wildtype and U-mutants 32

2.7 Direct telomerase activity assay tTR U mutants 32

2.8 SHAPE reactivities of nucleotides neighboring U-mutants 33

2.9 U-mutants impact SHAPE reactivity of nucleotides throughout tTR 34

2.10 tTR stem IV wildtype and 132-138 mutants 36

2.11 Direct telomerase activity assay tTR 132-138 mutants 36

2.12 SHAPE reactivities of stem IV apical loop nucleotides in 132-138 mutants 37

2.13 132-138 mutants impact SHAPE reactivity of nucleotides throughout tTR 38

3.1 Phylogenetic hTR secondary structure 54

3.2 SHAPE construct and hTR wildtype and mutant domains 56

3.3 SHAPE gel of wildtype and mutant pseudoknots 59

3.4 SHAPE analysis of the wildtype and mutant pseudoknots 61

3.5 Telomerase activity assay of super telomerase extracts with hTRs that incorporate pseudoknot mutants....62

3.6 SHAPE analysis of wildtype and mutant CR4-CR5 domains 65

3.7 SHAPE analysis of wildtype and mutant CR7 domains 67

4.1 SHAPE gel of full length wildtype hTR, selected primers 95

4.2 SHAPE analysis of full length hTR 96

4.3 Phylogenetic secondary structure of hTR with SHAPE data superimposed....97

4.4 Proposed RNA secondary structure of protein-free hTR 98

4.5 Effect of aminoglycosides on telomerase activity 107

4.6 Compounds binding to hTR elicit changes in secondary structure....109

4.7 Kan 6" binds to the CR4-CR5 domain of hTR 110

4.8 Secondary structural motifs of 5' hTR....113

4.9 Chemoenzymatic footprinting analysis of hTR nucleotides 1-210....115

4.10 Chemoenzymatic footprinting analysis of hTR nucleotides 211-451....116

LIST OF ABBREVIATIONS

1M7	1-methyl-7-nitroisatoic anhydride
2D-CS	2-Dimensional Combinatorial Screening
A	Adenine
ACS	American Chemical Society
AD	Autosomal dominant
AR	Autosomal recessive
ATP	Adenosine-5'-triphosphate
C	Cytosine
CAB	Cajal Body box
cDNA	complementary DNA
CMCT	1-cyclohexyl-(2-morpholinoethyl)carbodiimide metho-p-toluene
cpm	counts per minute
CTP	Cytidine-5'-triphosphate
CR	Conserved Region
DAT	Dissociates Activities of Telomerase
DC	Dyskeratosis congenita
DCR	Dyskeratosis Congenita Registry
ddC	dideoxycytidine
ddT	dideoxythymidine
DEPC	Diethylpyrocarbonate
DHX36	DEAH box protein 36
DKC1	Dyskeratosis congenita 1

DMS	Dimethyl sulfide
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxyribonucleotide triphosphate
dsDNA	double stranded DNA
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
FITC	fluorescein isothiocyanate
FRET	Fluorescence resonance energy transfer
G	Guanine
G4	G-quadraplex
GAR1	H/ACA ribonucleoprotein complex subunit 1
GTP	Guanosine-5'-triphosphate
H/ACA	hairpin-hinge-hairpin-tail structure
Hsp90	Heat shock protein-90
hTERT	human Telomerase Reverse Transcriptase
hTR	human Telomerase RNA
IA	Isatoic anhydride
IDT	Integrated DNA Technologies
Kan 6"	kanamycin-6" Az-FITC
LC	Loading control
MATLAB	Matrix laboratory

MBU	Molecular biology units
MCE	Mixed cellulose esters
mRNA	Messenger RNA
NAP	Nucleotide addition processitivity
Nea 5-O	neomycin-5-O-(AzEt)-FITC
NHP2	H/ACA ribonucleoprotein complex subunit 2
NMIA	N-methyl isatoic anhydride
NMR	Nuclear Magnetic Resonance
NOP10	H/ACA ribonucleoprotein complex subunit 3
ODN	oligonucleotides
PAGE	Polyacrylamide gel electrophoresis
PCI	Phenol/chloroform/isoamyl alcohol
PCR	Polymerase chain reaction
RAP	Repeat addition processivity
Rb	Retinoblastoma
RHAU	RNA Helicase associated with AU-rich element
RID1 and 2	RNA interaction domain 1 and 2
RNA	Ribonucleic acid
RNase	Ribonuclease
RNP	Ribonucleoprotein
RRL	Rabbit Reticulocyte Lysate
RT	Reverse Transcriptase
SAFA	Semi-automated footprinting analysis

SDS	sodium dodecyl sulfate
SHAPE	Selective 2'-Hydroxyl Acylation analyzed by Primer Extension
scaRNA	small Cajal Body-specific RNA
snoRNA	small nucleolar RNA
T	Thymidine
TBE	Tris/Borate/EDTA
TBE	Template boundary element
TE	TRIS-EDTA
TEN	Tris-EDTA NaCl
TERC	Telomerase RNA component
TERT	Telomerase Reverse Transcriptase
TMG	Trimethylglycine
TR	Telomerase RNA
TRE	Template recognition element
tTER	*Tetrahymena thermophila* telomerase RNA
tTERT	*Tetrahymena thermophila* telomerase reverse transcriptase
tTR	*Tetrahymena thermophila* telomerase RNA
U	Uracil
UNC	University of North Carolina at Chapel Hill
UTP	Uridine-5’-triphosphate
UV	Ultraviolet
WT	Wildtype

CHAPTER 1

TELOMERASE AND RNA STRUCTURE

A. Telomere Maintenance and Disease

There are a wide variety of diseases characterized by telomere shortening (1). The most common, though not customarily considered a "disease," is human aging, whose phenotype is consistent with progressive telomere shortening (1,2). Interestingly, the speed of aging and rate of telomere depletion appears to be due to a wide variety of extrinsic factors including age, sex, socioeconomic status, perceived stress, smoking, and obesity (1,3). Diseases typically associated with aging, such as cardiovascular disease and dementia, also have accelerated telomere shortening (1). In addition, several genetic disorders that are characterized by mutations in telomerase genes including aplastic anemia, dyskeratosis congenita, and idiopathic pulmonary fibrosis, are all characterized by shortened telomeres (1).

Similar to telomere shortening, telomere length maintenance by up-regulated telomerase can also lead to human disease; the most common being cancer of which greater than 85% express up-regulated telomerase. In our somatic cells that do not have active telomerase, telomere shortening and eventual cell senescence after many cell divisions serves as a natural tumor suppressor mechanism (1,4). However, once telomerase has been up-regulated, replicative barriers can be bypassed and cells acquire

the means for limitless proliferation, one of the hallmarks of cancer. The lifespan extension conferred by telomerase cooperates with other activated oncogenes and inactivated tumor suppressor genes to transform once regulated primary cells to tumorigenic cancer cell lines (4). Cancer is the third leading cause of death, only preceded by heart disease and infectious disease, and in the United States 1 in 4 people die from it (5,6).

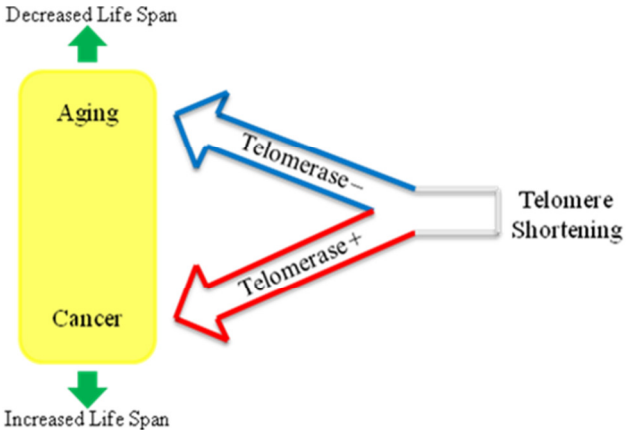


Figure 1.1. Antagonistic effects of telomerase in aging and cancer. Telomerase plays an antagonistic role in the development of both aging and cancer phenotypes. Figure adapted from Blasco 2007 (1).

As an element common to diseases characterized by shortened telomeres and irregular maintenance of telomeres, telomerase is an extremely important research topic. The development of rational approaches to control telomerase activity has the potential to combat cancer, aging, cardiovascular disease, and pre-mature aging syndromes (6). Meanwhile, the development of anti-telomerase drugs and therapies could take advantage of telomerase as a universal cancer target. Despite its importance, current understanding of telomerase is lacking. Therefore, elucidating the structural and functional role of telomerase and its components will make a powerful impression in the fields of cancer, aging, and telomerase biology.

B. Telomeres and Telomerase

Telomeres are specialized chromatin complexes at the ends of eukaryotic chromosomes that protect them from degradation and fusion (7,8). The telomeres are composed of tandem repeats of the sequence TTAGGG in humans and other vertebrates (TTGGGG in *Tetrahymena*) and bound in the shelterin complex (1). In the absence of telomere maintenance mechanisms, the chromosomes shorten with each cell division due to the semi-conservative nature of DNA replication (1,7-9). In our somatic cells, once the telomere reaches a critical length, known as the Hayflick limit, cells reach the limit of their replicative capacity and go into senescence via the p53 and Rb pathways which gives rise to cell proliferation arrest (2,9). If there are mutations in p53 and Rb that allow cells to bypass senescence, cell division continues until the telomeres become critically short and cells enter the crisis stage (2,10). In the crisis stage, extensive telomere shortening to already severely short telomeres usually results in programmed cell death or apoptosis (10). However due to increased genomic instability, some cells can evade and thus escape crisis stage by up-regulating telomerase, which can lengthen and stabilize telomeres giving rise to cells with infinite replicative capability, immortalized cells (10).

Telomerase is a ribonucleoprotein (RNP) complex that extends the length of telomeres by synthesizing telomeric DNA (11). The human telomerase enzyme consists of a protein reverse transcriptase (hTERT), a telomerase RNA (hTR), and a variety of associated proteins including the dyskerin complex (11,12). In cells expressing telomerase, these three components are kept separate from one another until the exact time within the cell cycle when telomeric DNA is to be elongated. hTERT is usually found bound to Pinx1, a tumor suppressor protein, that sequesters hTERT in an inactive

form in distinct nucleoplasmic foci (13,14). At the same time, mature and stabilized functional hTR bound to the dyskerin complex is maintained in the Cajal bodies (15,16). In early S phase, the hTERT protein moves from the nucleolus to within the Cajal bodies where it forms the telomerase holoenzyme complex with hTR/dyskerin (14,17). In mid-S phase, it is thought that the telomerase complex within the Cajal body and the Cajal body itself localizes within the nucleoplasm to the telomere (17,18). The telomerase complex is then thought to emerge from the Cajal body near the telomere, elongate the telomere, and then dissociate and return to the Cajal body showing the constant control that the cell exerts upon telomerase trafficking and regulation (14,17).

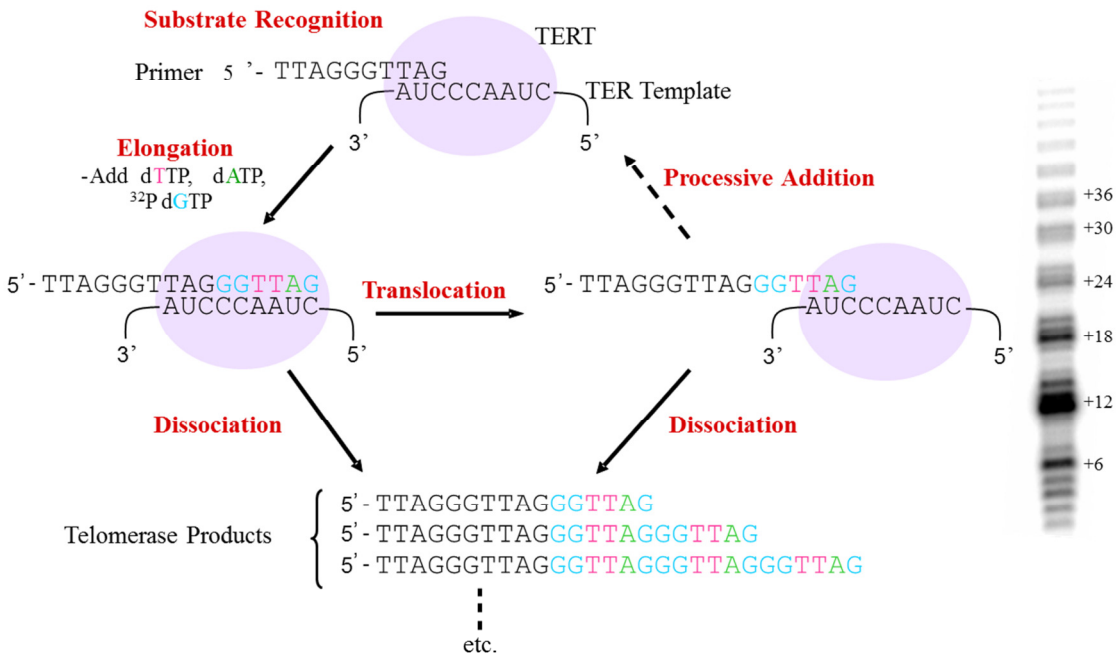


Figure 1.2. Mechanism of telomere primer extension by telomerase. The steps in primer extension or telomere synthesis are substrate recognition, elongation, and translocation. Assembled telomerase can extend telomeric primer by the addition of nucleotides. Translocation allows for further elongation and subsequent processive addition. Note that the dashed arrows represent the figurative transition from the post-translocation state to the state primed for elongation. Processive elongation is interrupted when the extended product dissociates from telomerase. The gel is representative of the telomerase-product ladder typically observed. This figure was adapted Sekaran et al. 2010. (19).

Once brought into the proximity of the telomere, telomerase undergoes a reaction cycle during which it is recruited to and can elongate telomeric DNA at the 3' overhang (15,20). In the first step, the telomerase RNA template active site aligns with and then binds to the 3' DNA overhang (21). The reverse transcriptase element of the TERT protein and the RNA template direct nucleotide addition processivity; this occurs sequentially from the 3' end of the DNA until the 5' end of the RNA template is reached (21). In humans this adds GTTAGG to the telomeric DNA (1,21). Next, the template can dissociate completely from the telomeric DNA, finishing the elongation; or it can dissociate and translocate to the 3' end of the newly synthesized sequence (21). If telomerase translocates, this reaction cycle is repeated until the telomere is sufficiently elongated.

C. Human Telomerase Reverse Transcriptase (hTERT)

The catalytic protein subunit of telomerase, hTERT, is 1132 amino acids and consists of a highly conserved structural organization that can be divided into four functional domains: an N-terminal domain, a telomerase-specific RNA-binding domain (T-motif), the reverse transcriptase (RT) domain, and a C-terminal domain (12,20). The large N-terminal region of hTERT is important for enzyme catalysis, *in vivo* RNP complex formation, RNA binding, multimerization, hTERT nucleolar localization, and template boundary definition (20,22,23). Additionally, this region also contains two smaller motifs: (1) an RNA- interaction domain (RID1) and (2) a dissociates activities of telomerase (DAT) motif involved in telomere length maintenance (22,23). Interestingly enough, the first 20 nucleotides of this domain have also been identified as containing a

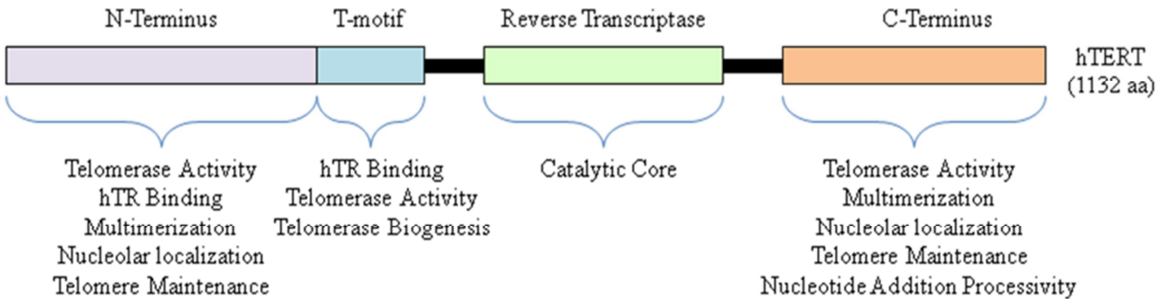


Figure 1.3. Organization of hTERT domains and their functions. hTERT contains four functional domains: an N-terminal domain, a T-motif or RNA binding domain,and RT domain, and C-terminal domain that have many protein and telomerase functions. Figure adapted from Pendino et al. 2006 (12).

mitochondrial localization signaling peptide, which hints at an extra-telomeric role for the hTERT protein (24). Deletion of these nucleotides confers less hydrogen peroxide induced oxidative damage on the mitochondria, suggesting that hTERT might have a role in a pro-apoptotic pathway in this organelle (24-26).

The T-motif, also known as the RNA-binding domain, is a telomerase-specific motif which is critical for telomerase activity *in vitro* and *in vivo*, telomerase biogenesis, proper 5' template boundary definition, and RNA binding via a second RNA-interaction domain (RID2) (21,22,27,28). Studies have shown that mutations in RID2 can reduce hTERT-hTR binding by as much as 50-70%, suggesting that RID2 is the major hTR-binding motif (22,29). The RT domain of hTERT contains all seven RT motifs characteristic of other reverse transcriptases as well as high sequence conservation with HIV-RT (12,20). Incidentally, even specific residues crucial in nucleotide binding, dNTP discrimination, and nucleotide-addition processivity appear conserved between HIV-RT and the hTERT RT domain (20). Like the N-terminal region, the C-terminal residues are not highly conserved. Functionally, the C-terminal region plays pivotal roles in cellular immortalization, multimerization, telomerase repeat processivity, primer binding,

telomerase activity *in vivo*, enzyme nuclear localization and telomerase length maintenance (12,23,30).

D. Human Telomerase RNA (hTR)

The RNA component of telomerase, hTR, is 451 nucleotides, contains the 11 nucleotide long template sequence involved in telomere repeat elongation, and is intrinsic for telomerase biological function (12,31). Though its size and primary sequence homology are highly divergent from other species, the secondary structure of hTR has been demonstrated to be phylogenetically conserved (12,32,33). The four universally conserved elements within its secondary structure include: the template containing pseudoknot domain (CR1-CR3), the CR4-CR5 domain, the Box H/ACA domain (CR6/CR8), and the CR7 domain (Figure 3) (31,33). Structurally speaking, the 5' end of the RNA which includes the pseudoknot and the CR4-CR5 domain is essential for enzyme processivity while the 3' region of the RNA which includes the Box H/ACA and CR7 domains is important for proper RNA cellular localization (31). Current models of hTR secondary structure have been derived from phylogenetic comparison or compensatory mutational analysis due to the large size of this RNA (31). Since comprehensive high-resolution data to define secondary structure of hTR is currently unavailable, we hope to alleviate this dilemma by providing these data in Chapter 4.

The template inclusive pseudoknot domain includes the first 209 nucleotides of hTR and is formed from the close interaction of conserved regions CR1-CR3 (Figure 1.4 and 1.5). Due to the extent of primary sequence conservation in this region and its predicted roles in telomerase biological function and the development of

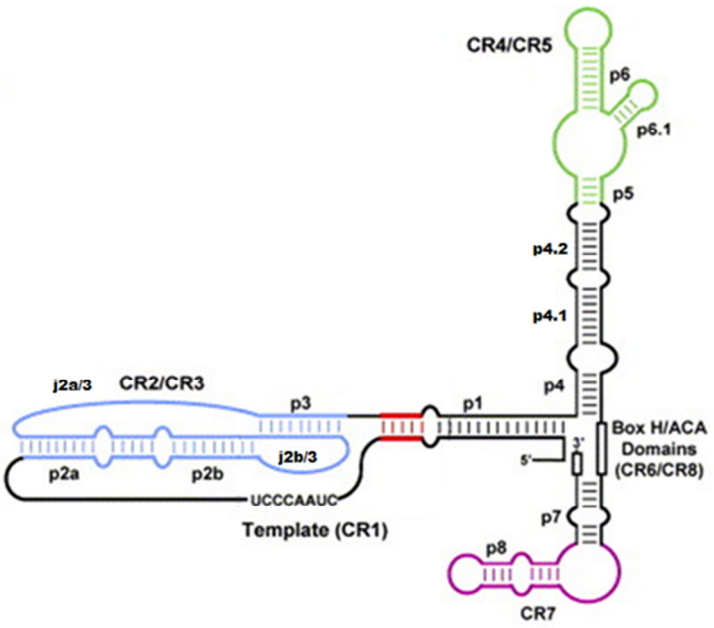


Figure 1.4. Phylogenetically derived secondary structure of human telomerase RNA (hTR). hTR has four highly conserved domains: the pseudoknot domain, CR4-CR5, CR7, and the Box H/ACA domain. Figure adapted from Legassie 2006 (33).

several diseases, the pseudoknot domain has been extensively investigated (31,33). The significance and role of nucleotides 1-17 has yet to be completely understood, though it has been suggested to adopt a G-quadraplex conformation or to interact with the CR4-CR5 domain providing a mechanism for hTR itself to control telomerase activity (Figure 1.5) (34). Though it is not completely conserved in vertebrates, the entire P1 helix does in fact form in hTR *in vivo* and the proximal end of the P1 helix is essential for telomerase activity (35,36). Furthermore, the proximal P1 helix contains the template boundary region and has also been implicated to possess a putative hTERT binding site (33,36). The RNA template region that binds to the telomeric DNA overhang and aids in the synthesis and addition of telomeric repeats appears to be mostly single stranded and not involved in any RNA-RNA or RNA-protein interactions (35). However, probing template structure under *in vivo* conditions insinuates the possibility of some sort of interaction for C50-C52 (35). Studies in yeast suggest close proximity between the

template region and loop J2b/3 to explain these template interactions (37). Additionally, examination of the helix leading into the pseudoknot in hTR structure does not support the formation of the previously predicted helix P2.1a, but does acknowledge P2a development (35).

There is a high correlation between pseudoknot tertiary structural formation and telomerase activity, suggesting an implicit role in biological function for this structural motif (33,38). A pseudoknot is an RNA tertiary structure containing two stem loops in which the first stem's loop (stem 1 is P2b and loop 1 is J2b/3) is part of the second's stem (stem 2 is P3 and loop 2 is J2a/3) and vice versa (32). In recent years, NMR solution structures that represent portions of the pseudoknot domain (P2b, P3, J2b/3, and J2a/3) including nucleotides 93-121 and 166-184, have greatly contributed to comprehending the role of this dynamic tertiary structure in telomerase function (Figure 1.5) (32,33). The pseudoknot structure comes together due to the U-rich J2b/3 loop (U99-U103) sitting in the major helix of P3, forming evolutionarily conserved (through ciliates) U•A•U Hoogstein triples that help maintain the pseudoknot structural integrity and contribute directly to enzymatic activity (32,33,39-44). Meanwhile, A171 and A172 from the J2a/3 loop nestle in the minor groove of helix P2b forming two non-canonical triple base interactions also yielding pseudoknot stability (33,36,40,41). Between these two sets of triple base pairings exists a Hoogstein base pairing between U99 and A173 that causes the stacking of helices P2b and P3, as opposed to a completed triple helix structure (33,40).

The hTR-specific trans-activating domain or hyper-variable region encompasses the next 160 nucleotides (~210-370) following the template-inclusive pseudoknot domain

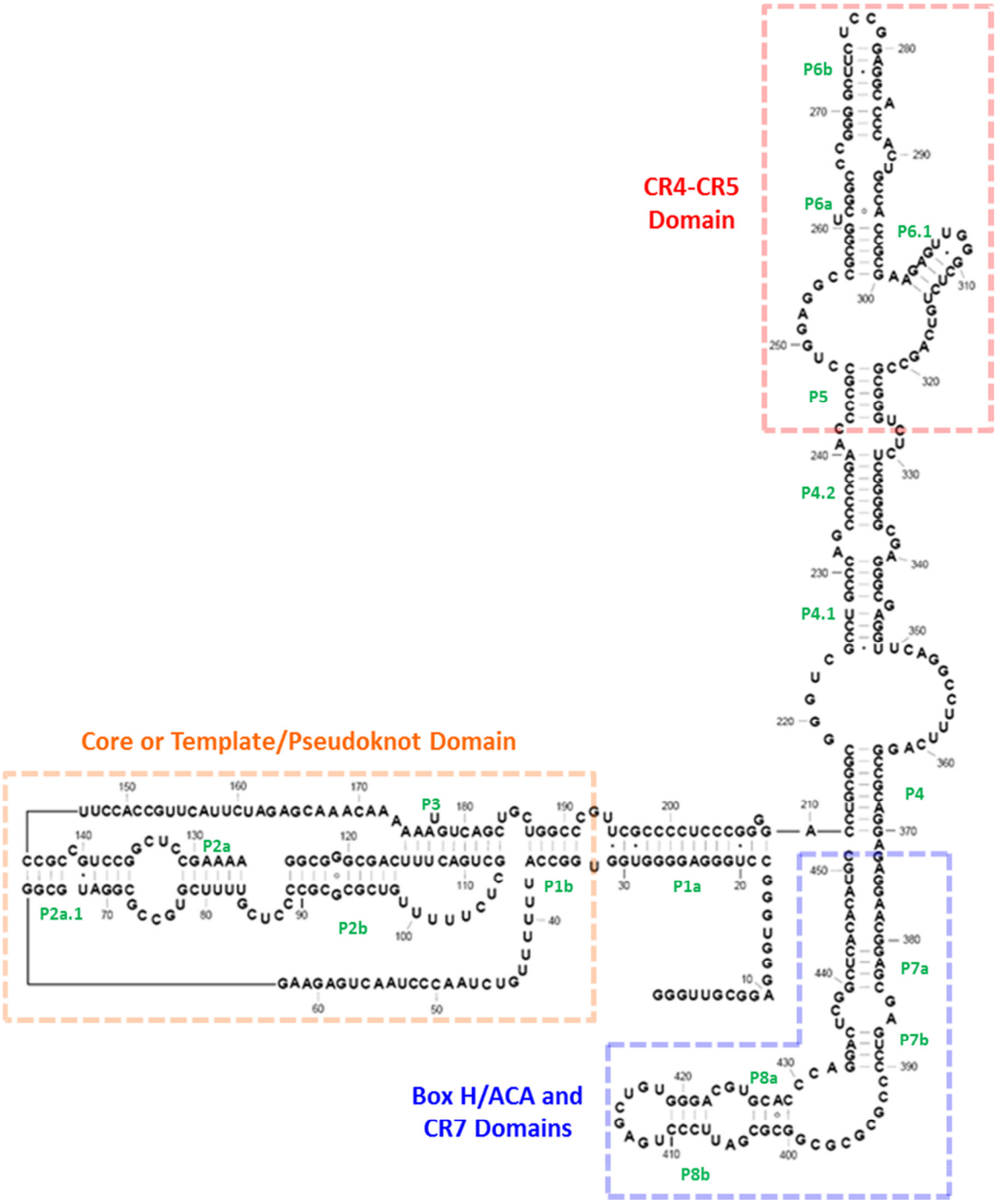


Figure 1.5. Secondary structure of hTR with sequence. Above is the secondary structure, important domains, and nucleotide sequence of hTR. Figure adapted from Chen et al. 2004 (31).

and contains conserved regions CR4 and CR5 (Figure 1.5). Unlike the pseudoknot domain, CR4-CR5 domain structure is less thoroughly characterized despite a high level of phylogenetic conservation (32,33,45). Represented as one long hairpin, the CR4-CR5

domain is composed of several very stable GC-rich helices broken apart by asymmetrical loops and single nucleotide bulges and contains structural elements required for proper TERT binding and therefore telomerase function (35,46). The stem-containing end of the domain which includes base-paired regions P4, P4.1, P4.2, and P5 appear to form tight helices that stack nicely upon one another and are separated by loops and bulges both *in vitro* and *in vivo*; there is currently not too much data available on these regions of hTR (Figure 1.5) (35).

Moving up towards the top to the stem loop, regions P6a and P6b form eight base pair helices that are separated by the asymmetric J6 loop (Figure 1.5) (47). P6a contains a bulged nucleotide at C262 that sits the helix major groove while the base of P6b forms a water-mediated triple base pair between C267 and the G268•C288 base pair (32,35,47). The J6 loop sandwiched between P6a and P6b is phylogenetically conserved among mammals with the first C residue and the last U residue being present across virtually all species (48). Interestingly, surface plotting of J6 reveals a solvent accessible space in the center of this loop suggesting that this region may act as a receptor for possible hTR-hTERT interactions, such as with one of the hTERT RNA-interacting domains (47). Probably the most scrutinized region of the CR4-CR5 domain is the p6.1 helix located between P5 and P6a. The four base pairs, wobble base pair, and three nucleotide loop that compose P6.1 structure are critical for telomerase activity, though the exact sequence of the stem does not appear to be essential (32,45,49). Loop nucleotides 307-309 have been demonstrated by NMR and chemical modification to adopt solvent exposed conformations *in vitro* but are protected *in vivo* suggesting the location of a TERT binding site (35,43,50,51). A more radical idea suggests that the loop of P6.1 might

participate in long range interactions with the template region yielding a major tertiary RNA folding event in hTR architecture (43).

The 3'-terminal structure of hTR, which is comprised of CR6-CR8 includes the H/ACA (hairpin-hinge-hairpin/ACA motif) box and CAB box that are required for *in vivo* hTR 3' end processing and localization and correct telomerase intracellular trafficking (Figure 1.5) (35,52). The hTR H/ACA region, which is important for delineation of the end of hTR for maturation purposes and necessary for accumulation of hTR intracellularly, is a distinct feature of snoRNAs (small nucleolar RNAs) and scaRNAs (small Cajal body-specific RNAs) (35,52). In these RNAs, the H/ACA motif serves to guide post-transcriptional modifications including pseudouridylation, methylation, and cleavage; however, these functions have not been proved for hTR (53). Proteins shown to associate with the H/ACA motif in scaRNAs and snoRNAs including dyskerin, NHP2, NOP10, and GAR1 are also associated with hTR and the active telomerase complex (51,52). Structural studies suggest that both the H box and the ACA box within the H/ACA motif are single stranded and flanked by nice tight helices (P4 and P7a), though, due to the terminal location of these sequences in hTR, absolute configuration is currently unknown (35). In contrast, crystal structures of other H/ACA motif-containing RNAs have identified the box ACA structure and nearby P7a helix and nucleotides in the looped region between P7b and P8a as required for proper dyskerin protein association (54-56).

Past the H/ACA region, is the CR7 region whose tight helices and CAB box containing loop structure are implicit for hTR accumulation in the cell (35,52). Phylogenetic analysis reveals a greater than 90% conservation for CR7 within mammals

and suggests that structural deviation here would prevent accessory protein binding and telomerase activity (42). Solution structures resolved by NMR have identified two wobble base pairings G404•U425 and U411•U417, the latter of which allows nucleotide U416 to bulge creating a kink turn motif that is needed for accessory protein binding and RNA processing (42,57,58). The CAB box (UGAG) found in the stem's end loop region remains single stranded as is required to reap the functional benefits of the sequence-specific element. Though data exists suggesting RNA architecture of specific nucleotides in the CR7 and H/ACA domains of hTR, most has been extrapolated from studies with other H/ACA motif-containing proteins (54). Only by examining the structure of these regions specifically in hTR can a greater understanding of their structure and contribution to hTR be fully realized.

E. Dyskeratosis Congenita

Dyskeratosis congenita (DC) is a rare, genetic, bone marrow failure syndrome that is characterized by shortened telomeres (59,60). Symptoms of this disease include abnormal skin pigmentation, premature graying of the hair, leukoplakia, premature loss of teeth, and nail dystrophy (59,61). Additionally, DC patients can also show signs of testicular atrophy, pulmonary fibrosis, liver cirrhosis, osteoporosis, premature aging and have an increased risk of pulmonary disease and cancer (62). Mutations in three different components of the telomerase holoenzyme can lead to clinical presentation of dyskeratosis congenita (59). The X-linked form of DC, the most common and severe form of this disease, is caused by missense mutations in the DKC1 gene (59). It is expected that these mutations weaken the interactions between dyskerin and the H/ACA

domain of hTR, preventing proper localization of hTR to the Cajal bodies, and creating a telomerase deficiency. The autosomal recessive (AR) DC is caused by homozygous mutations that can be found in any of the protein components of telomerase (59). Currently, cases have only been identified in hTERT and NOP10, though it is thought that AR-DC represents a significant portion of cases of this disease (59). Not much is known about AR-DC, though its severity and symptoms are along the same lines as the X-linked form of DC (59). The autosomal dominant (AD) form of DC is caused by missense mutations or deletions found in *TERC* (63,64). Of the nine families included within the Dyskeratosis Congenita Registry (DCR), each one presented a different mutation in *TERC* giving rise to AD-DC (64,65). The structural consequences of these mutations are unknown, though biologically these mutations lead to telomerase haploinsufficiency, which is not enough activity to maintain telomere lengths (66,67). We plan to examine the structural consequences of certain mutations in hTR that give rise to the autosomal dominant form of dyskeratosis congenita and other premature aging-related in Chapter 3 to better understand the mechanism of disease development.

F. The Importance of RNA Architecture

In recent years, a remarkable number of diverse RNA types have been identified with diverse functions ranging from the control and modulation of gene expression to RNA processing and modification and even serving as a template for chromosome elongation (68-74). Each of these distinct RNA types, though different in length and functional consequence, is encoded by the same four nucleotides: adenine, cytosine, guanine, and uracil (75,76). And unlike DNA which is mostly maintained in a double

helix structure, RNA can form a wide variety of secondary and tertiary structures including hairpins, internal loops, bulges, junction loops, A-minor motifs, kissing hairpins, and pseudoknots as shaped by both canonical and non-canonical Watson-Crick base pairings and electrostatic forces (77-79). This demonstrates that proper biological functionality requires more than primary sequence and is highly dependent upon secondary and tertiary interactions arising from RNA folding and therefore structure (78).

In addition to forming structural motifs, RNAs are also constantly in a state of conformational flux in their environment, meaning that a given RNA can have multiple conformational states (79). This dynamic attribute of RNA is characterized by intermolecular and intramolecular forces such as the hydrophobic effect, hydrogen bonding, metal ion coordination, van der Waal's forces, and even other biological molecules and binding ligands (75,77). Biologically speaking, an RNA's ability to readily switch conformations allows for greater control of function in differing cellular situations and even the ability to toggle between active and inactive states (75). Due to the many roles of RNA in the cell and the clearly evident, though not thoroughly understood, relationship between RNA conformations and biological function, RNA structural biology is an exceptionally important research topic in current molecular biology.

In the past few years, great advances have been made in the development of techniques to interrogate the three-dimensional structural characteristics of small RNAs (<100 nts) required for functional versatility (78). However, finding high-resolution methodologies to examine the global architecture of large RNAs is still challenging due to a wide variety of problems (80). X-ray crystallography, which is capable of capturing

smaller RNAs at 2-4 angstrom resolution, is hindered by large RNAs' numerous transition states, lack of conformational homogeneity, and structural dynamics (78,81). Furthermore, weak tertiary interactions cause more flexibility warranting a higher chance for RNA misfolding, while negatively charged phosphate groups on the surface prevent tight packing causing poorly ordered RNAs to diffract x-rays at much lower resolutions (81). NMR works well with RNAs up to 15 kDa, but beyond that point it presents its own set of problems including the need for pure millimolar sample concentrations and possible aggregation of larger molecules (82). Additionally, as RNAs have fewer NOEs than proteins, overall analysis has a tendency to suffer (83). Informatics and computational structure prediction are also encumbered by a natural "RNA misfolding problem" that arises from non-canonical Watson-Crick base pairings, ignorance of steric hindrances, and most difficult, folding dynamics which can range from milliseconds to minutes (76,84). In contrast, the relative simplicity of synthesizing DNA from RNA via reverse transcription and development of specialized software to analyze results has created a new paradigm for high quality structural information from large RNAs, making RNA footprinting an ideal method for probing local and global structure of large RNAs like hTR (76).

G. Selective 2'-Hydoxyl Acylation Analyzed by Primer Extension (SHAPE)

Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE) is an RNA footprinting technique that takes advantage of the fact that the chemical reactivity of the 2'-ribose position is strongly dependent on the adjacent 3' phosphodiester bond (85). As SHAPE reactivities are evaluated by primer extension, the RNA of interest must

be placed in a "structured cassette" before chemical modification to ensure that data can be quantified at all nucleotide positions, especially both the 5'and 3' ends of the RNA (Figure 1.6) (86). Due to the nature of reverse transcriptases, nucleotides neighboring the primer binding site cannot be measured because of the pauses the enzyme experiences during its initiation phase (86). And if the primer binding site happens to be the 3'-end of the RNA, those sequence data points will be lost as well. Similarly, on the 5' end of the RNA, the presence of intense bands corresponding to the full-length extension product masks SHAPE reactivities of initial nucleotides. The "structured cassette" involves adding structured linkers to both the 5'and 3' ends of the RNA as well as adding a reverse transcriptase (RT) binding site to the very 3' end of the RNA (Figure 1.6) (86). Each of these added regions folds into its own hairpin structure and does not appear to interfere with native RNA folding, but allows for complete analysis of RNA structure (86).

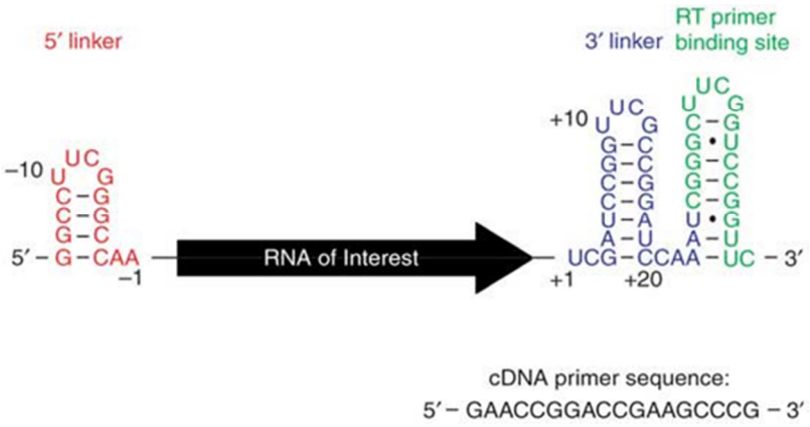


Figure 1.6. RNA structured cassette for SHAPE. Adding these flanking regions to an RNA of interest will facilitate analysis of all nucleotides within the RNA by SHAPE. This figure was taken from Merino et al. 2005 (85).

Chemical modification occurs via the addition of the electrophilic reagent N-methyl-isatoic anhydride (NMIA) (dissolved in DMSO) to the folded RNA which selectively acylates the ribose 2'-hydroxyl position while liberating carbon dioxide

(Figure 1.7) (85,87). The sensitivity of the 2'-ribose to esterification is directed by the targeted nucleotide's flexibility, RNA structural conformation, and therefore, also the proximity of the nearby 3'-phosphodiester anion (85). This suggests that nucleotides in highly flexible conformations will be more heavily modified than nucleotides that adopt more constrained conformations (85,87). Since every nucleotide has a 2'-hydroxyl, structural information is available for every nucleotide position in the RNA in a single chemical modification step, as compared to other chemical or enzymatic modifications that are accessible only to specific targets (78,85). An electrophilic reagent capable of interacting with a 2'-hydroxyl will also face competition from the simultaneously occurring hydrolysis reaction; however, NMIA demonstrates good first-order kinetics and over time remaining reagent is consumed by this hydrolysis reaction which gainfully allows the reaction to be self-limiting (85,86). For each reaction with NMIA, a similar reaction with just DMSO is also conducted as a negative control for calculating background.

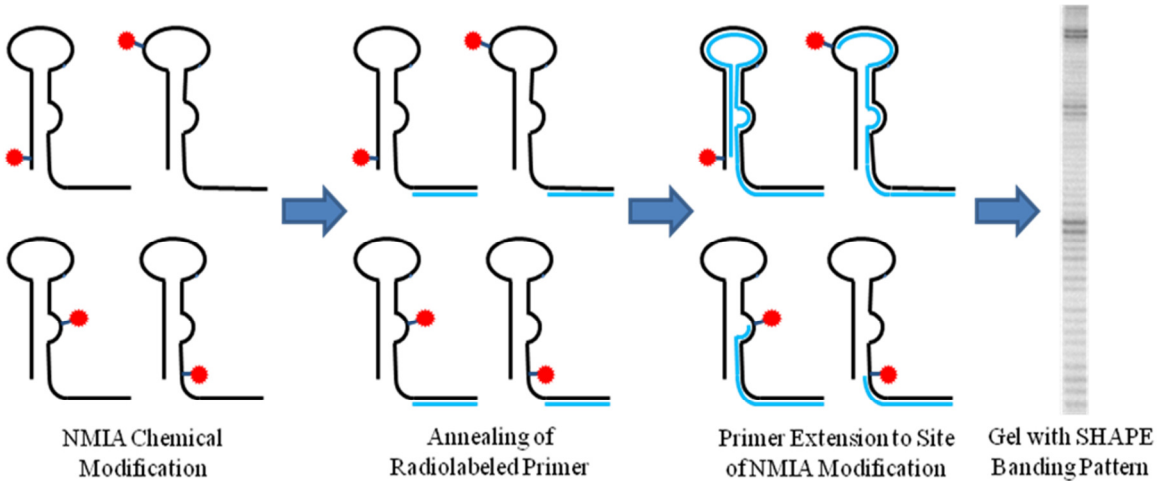


Figure 1.7. SHAPE from beginning to end. The SHAPE footprinting involves chemically modifying the RNA of interest, annealing a radiolaebled DNA primer, and primer extension via reverse transcription to make cDNAs that are then run on an 8% super thin sequencing gel to visualize nucleotide flexibilities. Figure adapted from Merino et al. 2005 (85).

Following RNA chemical modification, the reactivity of the nucleotides to NMIA and quantification of adduct formation is assessed utilizing reverse transcription. A 5'-radiolabeled DNA primer is annealed to the RT binding site incorporated at the 3'end of the RNA and a thermostable, highly efficient reverse transcriptase extends the primers up to the modification sites. Presence of a 2'-O-adduct causes the reverse transcriptase to stop one nucleotide before the modified base, yielding a radiolabeled DNA fragment representing that specific nucleotide (Figure 1.7) (85,86). Upon completion of the extension step, the RNA is degraded by base treatment and the DNA is separated via electrophoresis in a polyacrylamide sequencing gel to resolve nucleotide flexibility data which is represented by varying radiolabeled fragments corresponding to individual RNA bases. In addition, sequencing ladders are generated using one or two dideoxy nucleotides, as a way to infer the sequence and orient oneself within the banding patterns. Quantification of the large amounts of data generated from each SHAPE experiment is completed using a high-throughput MATLAB method, semi-automated footprinting analysis (SAFA), which rectifies data from the thousands of band intensities to numerical values by peak-fitting and quantification algorithms (88,89). Subtraction of DMSO values from the NMIA data followed by normalization then yields actual nucleotide flexibility data that can be manipulated to determine actual RNA structure.

H. Thesis Overview

The research presented in this thesis is focused on determining the structure of telomerase RNAs. In chapter II of this dissertation, we examined the structure of tTR using fast- and slow-reacting SHAPE reagents to identify nucleotides involved in RNA

folding mechanisms. We also examined two structural motifs within the stem IV helix of tTR that have locally flexible conformations and are vital for telomerase activity. In chapter III of this dissertation, we determined the solution structure for three domains of hTR: the pseudoknot, CR4-CR5, and CR7 domains. Analysis of mutants in these domains revealed that the changes in nucleotide sequence cause drastic reorganization of the RNA secondary and tertiary structure thus leading to the development of premature-aging related diseases. In chapter IV, we determined the solution structure of full length hTR and then used the structural information in order to identify the binding sites of aminoglycosides that inhibit telomerase assemblage and activity by binding to hTR.

CHAPTER 2

DIFFERENTIAL SHAPE AND ANALYSIS OF tTR STEM IV MUTANTS

A. Introduction

The telomerase holoenzyme complex was first discovered in the ciliate *Tetrahymena thermophila* (90). Ciliates have two functionally distinct nuclei, a micronucleus that is transcriptionally silent and a macronucleus that contains approximately 20,000 chromosomes, and therefore about 40,000 telomeres insinuating that there is plenty of telomerase present (90). Due to the bountiful amounts of telomerase and the numerous studies that have focused on ciliates, *Tetrahymena thermophila* has emerged as the model organism for studying telomerase. The *Tetrahymena thermophila* telomerase RNP holoenzyme has been characterized by affinity purification, and consists of TERT (telomerase reverse transcriptase), tTR (*Tetrahymena thermophila* telomerase RNA), and four associated proteins, p20, p45, p65, and p75 (91). TERT, the catalytic protein subunit, contains a region of homology with other reverse transcriptases as well as telomerase-specific regions that are required for activity (92). P65, one of the accessory proteins, has been shown to bind with tTR and is thought to contribute to structural reorganization of the RNA subunit required for assemblage with TERT to form the active telomerase complex (93).

The *Tetrahymena thermophila* telomerase RNA (tTR) is 159 nucleotides long and composed of four conserved helices separated by single stranded regions, all of which are functionally significant (94). Stem I is involved in long-range base pairing that bridges the 5' and 3' ends of the RNA and its sequence and structure have been identified as essential for TERT binding (Figure 2.1) (95). The base of Stem II and the single-stranded region immediately preceding the template sequence contain the template boundary element (TBE) as well as a protein-binding site that interacts with the RNA-binding domain of TERT (95). Next is the template sequence, 5'-CAACCCCAA-3' (nucleotides 43-51), that is requisite for the nucleotide addition followed by the template recognition element (TRE) (96). Stems IIIA and IIIb are implicated in the formation of an H-type pseudoknot that is important for telomerase catalytic activity *in vitro* and RNP assembly *in vivo* (97). Lastly, Stem-loop IV is responsible for the correct folding of the Stem III pseudoknot as well as binding P65 and contributing to telomerase enzymatic activity and repeat nucleotide processivity (38,98).

Stem II, the template boundary element, and the 5' single-stranded portion before the template region are essential for defining the 5' template boundary and for tTERT binding (29,99). Of these regions, nucleotides A16-C19 and G37-A40 demonstrate the highest amount of sequence conservation with the remainder of the nucleotides being approximately 70% conserved among ciliates (Figure 2.1) (100). NMR structural analysis of stem II shows that nucleotides C19-U25 and A31-G37 base pair and form a continuous A-form helical stem composed of six Watson Crick base pairs with bulged residues A22 and A34 not bending or disrupting the character of the helix (100). The helix itself is capped by a well-defined five-nucleotide loop structure with U27, A29,

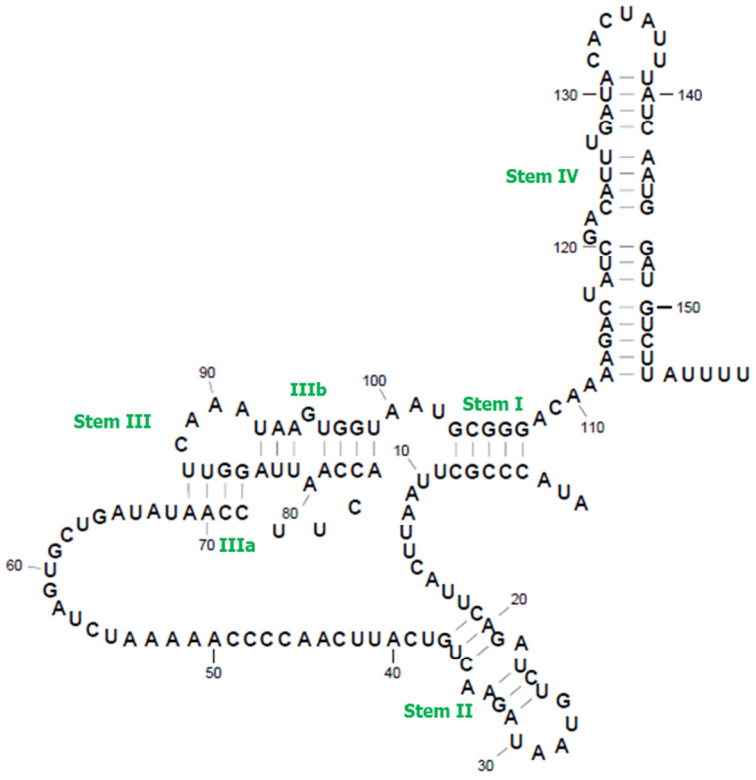


Figure 2.1. Phylogenetic secondary structure of tTR. The phylogenetically predicted secondary structure of *Tetrahymena thermophila* telomerase RNA (tTR). The four important helix regions within the structure are labeled as stems I-IV, with stem III being broken up into two sections.

and U30 being conserved within the *Tetrahymena* species (100). In the loop, nucleotides G26 and U27 stack directly upon the loop-closing base pair U25-A31 with G26 contributing to overall loop stability even though it does not actually interact with any of the loop nucleotides via hydrogen bonding (100). Of the remaining loop nucleotides, both U27 and U30 both adopt C2' endo sugar conformations, A29 adopts a syn conformation with its Watson Crick face solvent exposed, and A28 at the top of the loop points down toward the helical axis (100). Studies focusing on G37-A40 have identified that mutation of any of these nucleotides lead to primer extension beyond the normal 5' template boundary and disruption of tTERT binding (Figure 2.1) (101). These data confirm that the C19-G37 base pair and sequence of G37-A40 are critical for defining the

5’ TBE and tTERT binding and that the nucleotides involved in these two functions are indistinguishable (29).

Sequences of tTR from the template region to stems IIIa and IIIb, which are part of the proposed pseudoknot structure in catalytically active telomerase complex, have been examined via chemoenzymatic footprinting techniques for naked tTR and in complex with tTERT (Figure 2.1) (102,103). DEPC probing of the TRE and DMS probing of the template region both without tTERT showed these regions to be structured and involved in possible base pairing or stacking interactions despite the fact that they have both been predicted to be single stranded (102,103). Under similar conditions, RNase V1 cleavage of the same regions also showed the template and TRE to be structured as well (98). On the other hand, DEPC and DMS probing and RNase VI cleavage indicated that the pseudoknot was unstructured and did not completely form in tTR when tTERT was not present (98,102,103). In stark contrast, DMS modification of intracellular tTR and RNase V1 cleavage in the presence of tTERT revealed a more protected and well-formed pseudoknot structure for stems IIIa and IIIb (98,103). The data derived from these chemoenzymatic footprinting experiments suggests that the pseudoknot undergoes structural reorganization upon tTERT binding.

Stem IV and the helix capping loop are essential structural components in telomerase and are involved in a wide variety of functions from interacting the p65 and tTERT proteins to aiding in the organization of the pseudoknot tertiary structure that is imperative for proper telomerase activity. When considering the entire domain, the apical loop of stem IV has the highest level of sequence conservation with all nucleotides except C134 being completely conserved (104,105). This loop region has been shown to

mediate nucleotide addition processivity (NAP) and repeat addition processivity (RAP) in the telomerase complex (104). The loop closing base pair is a non-canonical interaction between C132 and U138 that is stabilized by a single hydrogen bond that is indispensable for telomerase activity. Mutagenesis of these nucleotides to a Watson Crick base pair completely abrogated telomerase activity, demonstrating that this base pair needs to be a weak interaction (105,106). The base-pairing of these two nucleotides also organizes the stacking of the all the other nucleotides in the loop, with the exception of the flipped-out and unconstrained U137 (106). Nucleotides A133 and U137, above the non-canonical base pair, could potentially form a Watson Crick interaction. However, forcing these two nucleotides to base pair attenuated telomerase activity, showing that flexibility in these loop nucleotides is imperative for proper telomerase activity (105). Characteristics of other loop nucleotides include the assumption of C2' endo sugar puckers by U135, A136, and U137 and mutation of C134 does not affect telomerase activity, consistent with its being the only non-conserved nucleotide in the apical loop (Figure 2.1) (104,105).

The stem portion of this helix also has structural features that greatly affects telomerase activity and plays a role in proper stem IV biological function. First of all, an absolutely conserved GA bulge flanked by absolutely conserved G-C base pairs induces a 43° kink in the helix between the two smaller helical regions creating a large wedge that bends the helix into the major groove (Figure 2.1) (104,105). Whereas nucleotide A122 remains stacked upon C123, nucleotide G121 is bulged out creating the kink in the helix; the local structure at this motif is consistent with both chemical and enzymatic mapping experiments (102). Highlighting the importance of the GA bulge and its flanking nucleotides are studies in which deletion of the GA bulge negatively affected *in vitro*

reconstituted telomerase activity and substitution of the highly conserved G-C base pairs with A-U Watson Crick base pairs reduced telomerase activity 20-fold (95,98,105). Structural implications of deleting the GA bugle would be the creation of a perfectly straight, base-paired helix in the middle of stem IV which would prevent the proposed interaction of the apical loop with the pseudoknot structure (105). The helical bend can be rescued, however, by the binding of p56 to stem IV (107). A second structural feature of stem IV helix is the configurational flexibility noted in nucleotides U126 and U127 and their pairing partner A143 (Figures 2.1 and 2.2). The bulge of either of these nucleotides has been shown to cause the helix to bend slightly (~10%) though, surprisingly, deletion of U127 does not affect either NAP or RAP, though a decrease in overall telomerase activity between tTR and tTERT is observed (104). A third region of interest in stem IV are the two A-U base pairs proximal to the GA bulge that even though they are suggested to, do not actually form base pairs as demonstrated by NMR and increased SHAPE reactivities (105). These A-U interactions are highly conserved within *Tetrahymenae* and substitution of the A-Us with G-Cs diminished telomerase activity to 40% and decreased processivity by 2-fold demonstrating the local structural flexibility they impart (105).

To date, the current secondary structure of tTR has been based on phylogenetic and mutational analyses and supported by NMR data depicting the structure of individual helices (like the data described above). While a good portion of the biochemical data supports the phylogenetically determined secondary structure of tTR, there are discrepancies as well, like the chemoenzymatic footprinting data that suggests that the template and TRE are involved in base pairing or stacking interactions when tTR is in

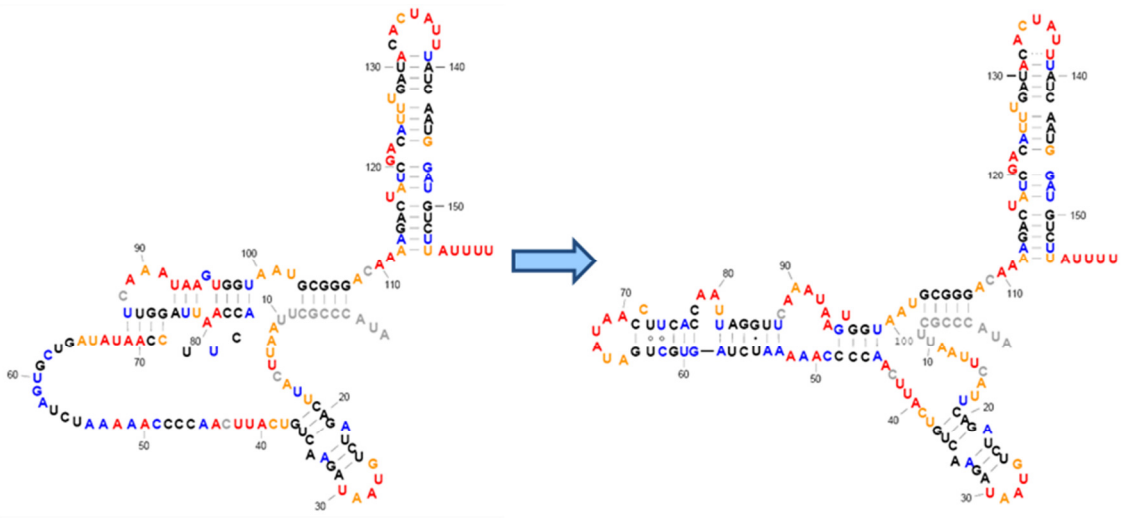


Figure 2.2. tTR secondary structure. Previous studies in our lab have determined the structure of tTR in solution. Analysis of the SHAPE data proved that tTR does not adopt the phylogenetically predicted strutcure when it is in solution. Using RNAstructure and SHAPE reactivity, a new structure was developed that better captured tTR structural data. Adapted from Legassie et al. 2012. (*unpublished manuscript*).

solution (98,102,103). With the goal of gaining a better understanding of tTR structure and to improve upon structural data available for full length tTR, a recent graduate of our lab utilized selective 2' hydroxyl acylation analyzed by primer extension (SHAPE), a high resolution RNA footprinting technique developed at UNC, to determine the secondary structure of tTR in solution and in complex with tTERT. Analysis of SHAPE data corresponding to all of the nucleotides in tTR demonstrated that the template region was not single stranded and that stems IIIa and IIIb did not form a pseudoknot structure in tTR in solution as suggested by the phylogenetically derived secondary structure, therefore validating previous footprinting data (Legassie et al. 2012, *unpublished manuscript*). Disagreement with the phylogenetic structure led to the proposal of an alternative structure that better captured the SHAPE data and is now the accepted structure for tTR in solution in our lab (Figure 2.2).

In this chapter, we utilized SHAPE chemistry in order to further examine tTR secondary structure. In the first experiments, we completed SHAPE using two

derivatives of isatoic anhydride that exhibit differential reaction rates in order to identify nucleotides within tTR that may play a role in governing tTR folding. In the remaining experiments, we focused on examining nucleotides within tTR secondary structure that have been shown to impact telomerase activity. We created one set of mutants to investigate the stem IV bulged Us at U125, U126, and U127 and a second set of mutants to interrogate stem IV apical loop-closing base pair C132-U138 hoping to gain a better understanding of the functional and structural implications of these interactions in tTR.

B. Results

1. Addition of 3' flanking region to tTR for SHAPE experiments

In order to analyze the full-length *Tetrahymena thermophila* telomerase RNA (tTR) by SHAPE tTR was cloned with a 3' flanking region. The 3' flanking region includes an 18 nucleotide reverse transcription primer binding site at the very 3' end followed by a 24 nucleotide linker which allows for the reverse transcriptase to become fully processive before encountering the RNA of interest (85). The 5' flanking region was not added as previous experiments conducted in our lab demonstrated that when the 5' and 3' flanked RNA was reassembled in the telomerase complex that telomerase activity was diminished (Legassie et al. 2012, *unpublished manuscript*). The 43 nucleotide extension forms two stable stem loop structures to avoid folding interactions with tTR (85).

2. IA and 1M7 modification of tTR

Analysis of SHAPE data from experiments that used two different "hit" reagents, the slow-acting reagent IA and the fast-acting reagent IM7, demonstrate that a single RNA can elicit two distinct reactivity profiles (Figures 2.3 and 2.4) (108,109). Of the

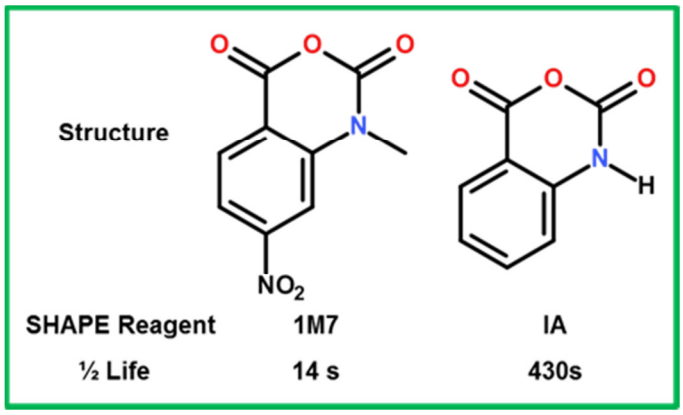


Figure 2.3. SHAPE Reagents: 1M7 and IA. Chemical structure of SHAPE reagents: fast-reacting 1M7 and slow-reacting IA. By utilizing different SHAPE reagents, one can identify nucleotides with varying characteristics.

159 nucleotides in tTR that were examined, approximately 21 nucleotides displayed significant differences in reactivity (Figures 2.4 and 2.5). Nucleotides in green were more reactive when slow-reacting IA was the electrophilic acylating reagent, whereas nucleotides in red were more reactive when fast-acting 1M7 was the reagent used to modify tTR. For 20 of the 23 nucleotides, all of the ones colored red or green in Figure 2.5 with the exception of A91, A93, and A94, differences in SHAPE reactivity were statistically significant with 95% confidence via the T-distribution (data not shown). For nucleotides A91, A93, and A94, there is insufficient evidence to indicate a difference in the SHAPE reactivities. Inspection of the secondary structure of tTR in solution identifies that most of the nucleotides that demonstrate differential SHAPE reactivity are single-stranded or in regions of high flexibility such as bulges or loops (Figure 2.5). Of the other nucleotides, U82, U99, and G103 are base paired and located at the end of helices; U125, U126, and U127 are located in configurationally flexible base pairing with

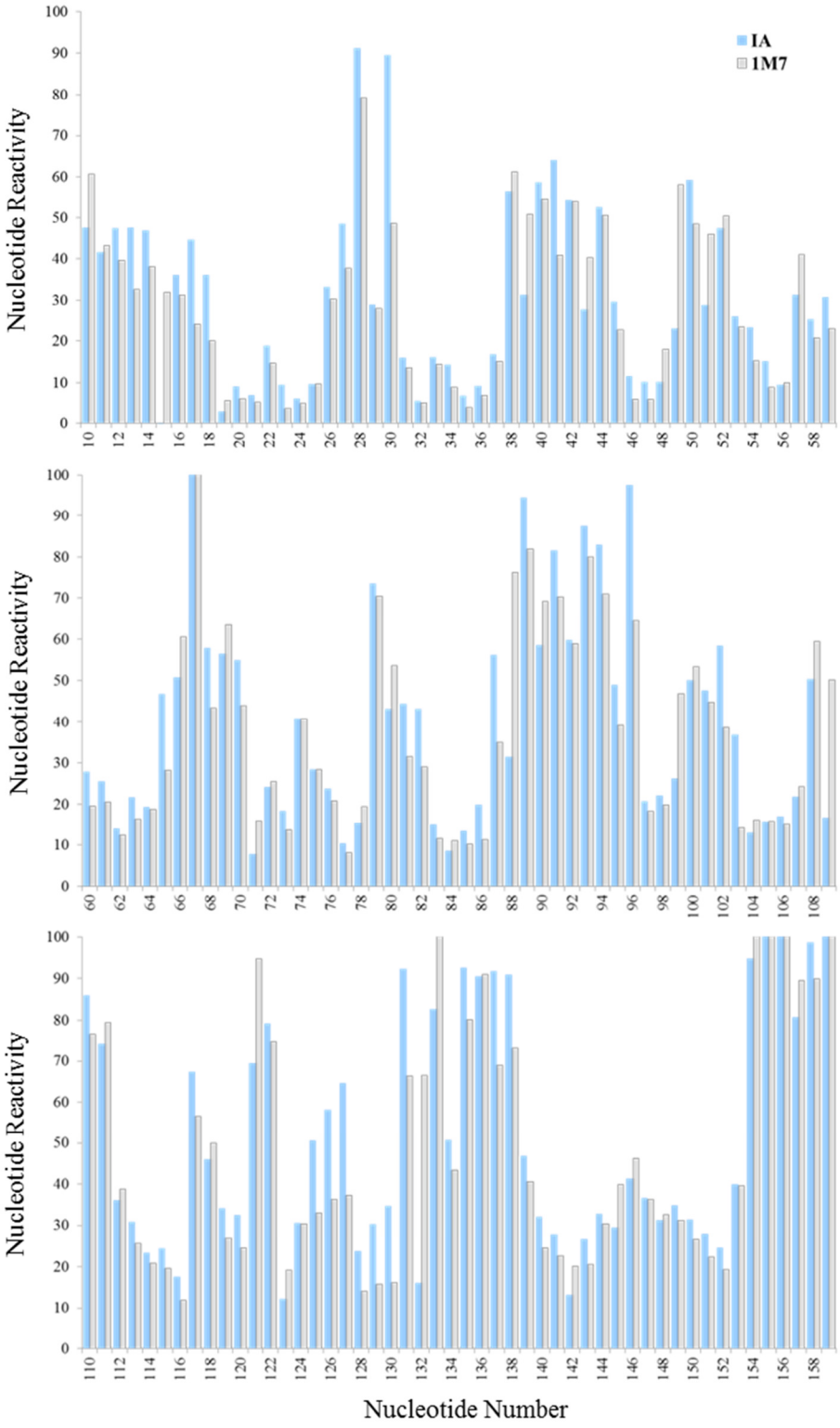


Figure 2.4. SHAPE of tTR with IA and 1M7. Histograms of SHAPE reactivity that corresponds to full-length tTR experiments using two different SHAPE reagents: IA and 1M7. The average band intensities were normalized and plotted according to nucleotide position with error bars (N=7-14).

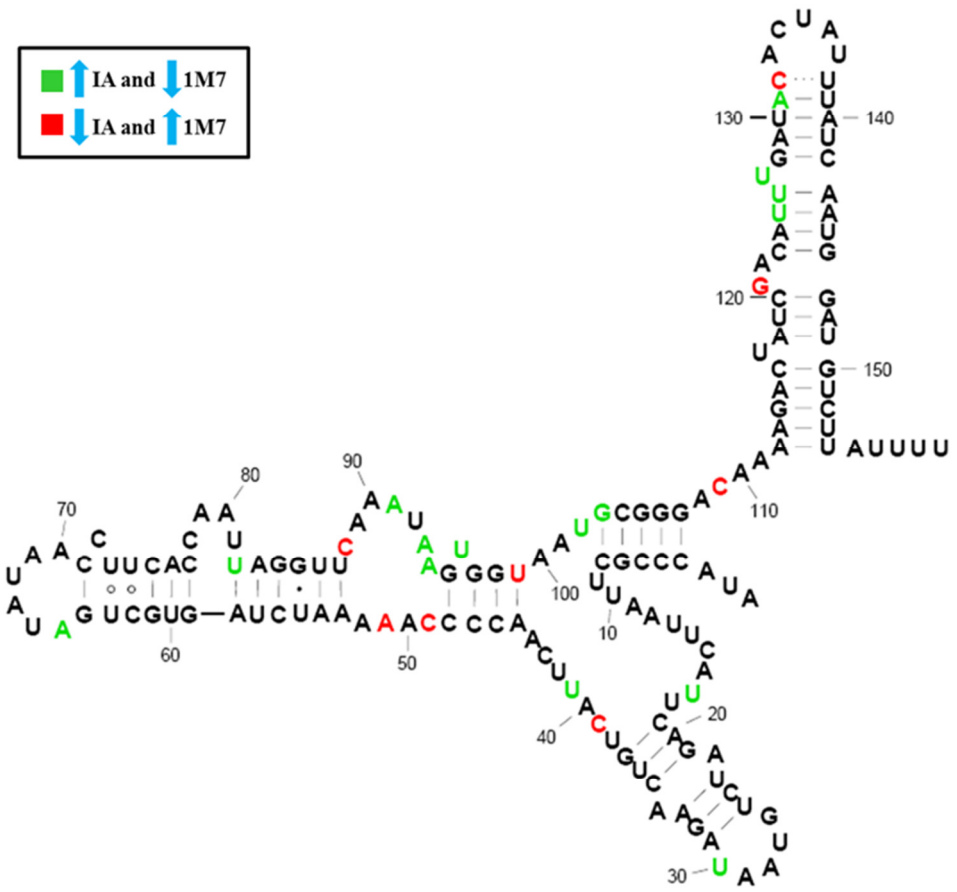


Figure 2.5. Nucleotides in tTR with differential SHAPE reactivity when using IA and 1M7. SHAPE experiments with IA and 1M7 as the SHAPE reagents identifies the colored nucleotides as having differential SHAPE reactivity in the presence of these two reagents. Nucleotides in green are more reactive when tTR is "hit" with IA. Nucleotides in red are more reactive when tTR is "hit" with 1M7.

A143 and A144; A131 is stacked right below a non-canonical base pair; and C132 is involved in a non-canonical base pair, placing them all in locations that would not be surprising to have more flexibility than normal Watson and Crick base pairs.

3. 125-127 U Mutants

Noting the configurationally flexible base pairing near the middle of Stem IV of nucleotides U125, U126, and U127 with A143 and A144 in the solution structure of tTR, we designed mutants to identify if a specific base pairing is preferred (Figure 2.6). In addition, a there was a preference for one side of the helix to be flexible. After purifying

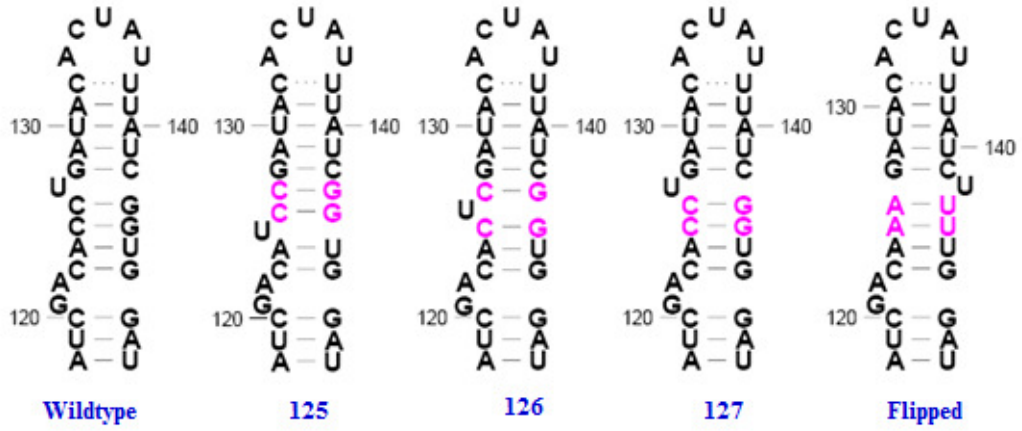


Figure 2.6. tTR stem IV wildtype and U-mutants. Analysis of the tTR solution structure identified that U's 125-127 all appear to be reactive even though only one should base pair at a time. To examine this conformational flexibility, mutants were designed that forced U125, U126, and U127 to bulge. Additionally, a fourth mutant that flipped the orientation of the base pairs was also made and studied.

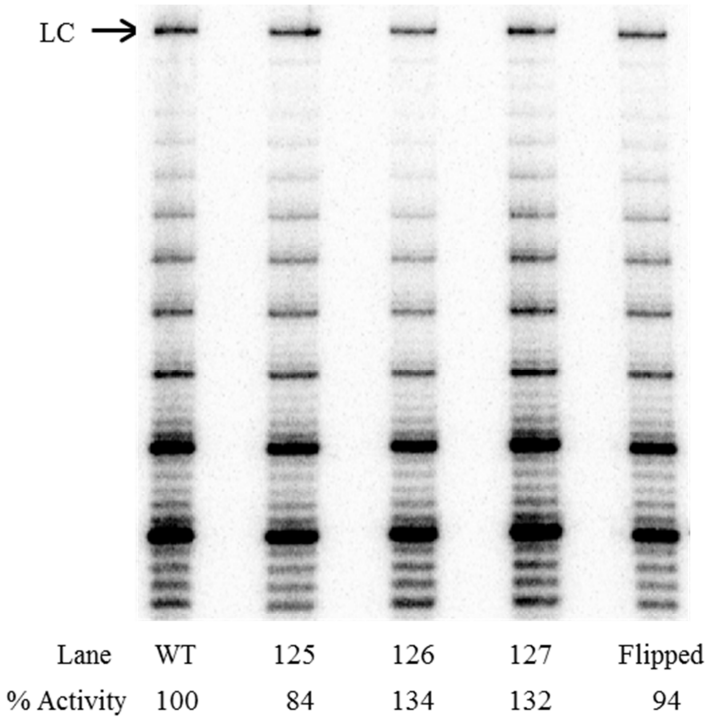


Figure 2.7. Direct telomerase activity assay tTR U mutants. A direct telomerase assay was conducted with purified telomerases incorporating wildtype and U mutant tTRs. The loading control (LC) at the top of each assay was used to normalize telomerase activity for each experiment.

fourth mutant that flipped the U residues and A resides was made in order to determine if telomerase complexes incorporating both wild-type and U-mutant tTRs from rabbit reticulocyte lysates, we performed a direct telomerase assay to assess the degree to which

the mutants affected telomerase activity and processivity. Quantification of telomerase activity demonstrated that the U-mutants did not affect telomerase activity with all samples showing near to or slightly greater than 100% telomerase activity (Figure 2.7). There also appeared to be no change in processivity, with all samples making typical ladder-like products. Direct telomerase assays of two other mutants that created 2 or 3 G-C base pairs in place of nucleotides U125, U126, U127, A143 and A144, therefore removing all flexibility, showed similar results for both telomerase activity and processivity (data not shown).

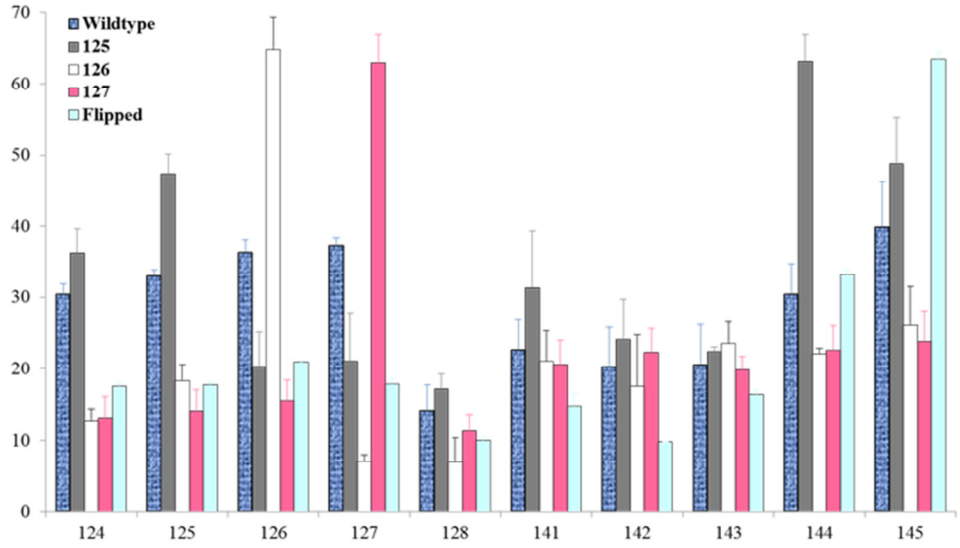


Figure 2.8. SHAPE reactivities of nucleotides neighboring U-mutants. Histogram of SHAPE reactivities corresponding to nucleotides immediately flanking the U-mutants. The average band intensities were normalized and plotted according to nucleotide position with error bars (N=3).

SHAPE analysis of the nucleotides surrounding and including the U and A residues only show slight diversity in nucleotide reactivity as compared to wild-type tTR when U-nucleotides were bulged forcibly. As expected, in the mutant corresponding to each bulged U residue, that U exhibited higher SHAPE reactivity (Figure 2.8). Only when U125 was bulged did a nucleotide flanking the U residues, A124, present higher flexibility; in the wild-type and all the mutants, SHAPE reactivity of G128 was minimal.

On the opposite side of the helix, bulging of U125 did elicit higher than normal flexibility from A144 while bulging of U126 and U127 decreased the flexibility of neighboring U145 in comparison to the wild-type. In the flipped mutant, movement of the A residues to the 5' side of the helix yielded low SHAPE reactivity from A125, A126, and all adjacent nucleotides. When the U residues are moved to the 3' side of Stem IV, U142 and U143 appear to form base pairs with A125 and A126, respectively. While U144 does appear to be slightly reactive to SHAPE modification, definitely more-so than U142 or U143, it seems likely to be making a base pair with A124 and causing U145 to become the bulged nucleotide, as its flexibility increases significantly (Figures 2.8 and 2.9).

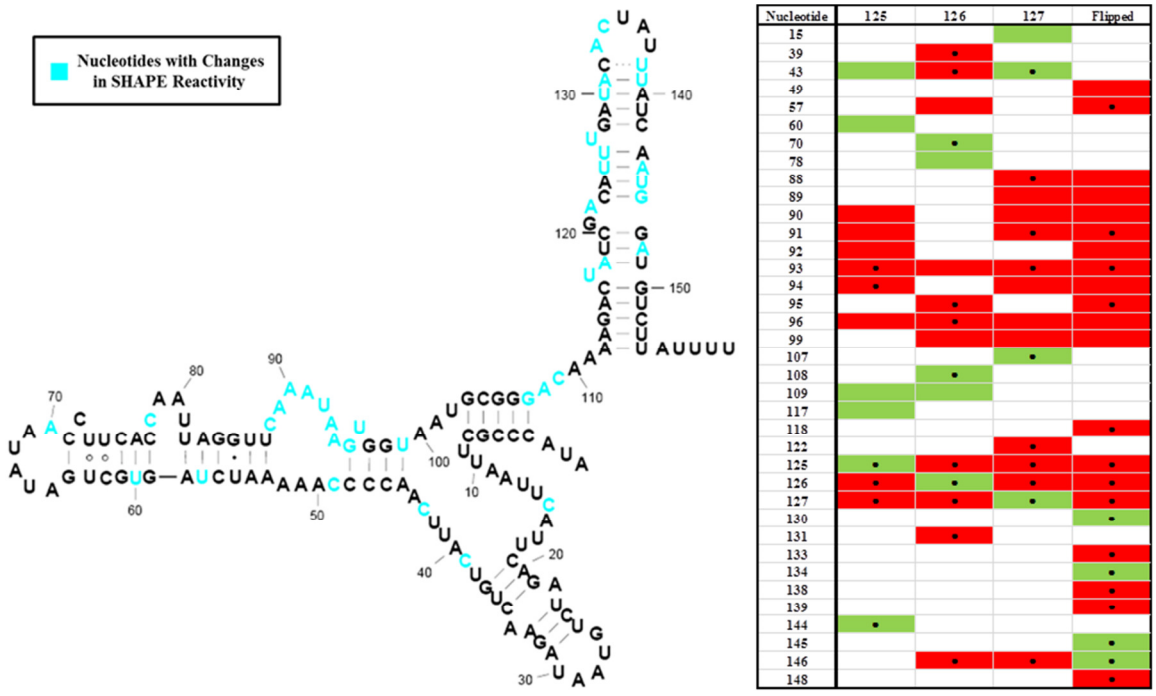


Figure 2.9. U-mutants impact SHAPE reactivity of nucleotides throughout tTR. Mutation of nucleotides in the Stem IV helix of tTR affect nucleotide flexibility throughout the RNA. The secondary structure diagram highlights all of the nucleotides (collectively) that experience large changes in flexibility due to mutations. The chart to the right shows which nucleotides had changes in flexibilities for each of the U-mutants. Red boxes identify nucleotides that become more structured in the mutant, whereas green boxes are more flexible in the mutant as compared to wildtype activity. The ● indicates that the difference in SHAPE reactivity is statistically significant with 95% confidence via the t-test. For nucleotides without the dot, there is insufficient evidence to indicate a difference in the SHAPE reactivities.

In looking at the entire secondary structure of tTR, forcing U's 125-127 to bulge and flipping the configuration of the U and A residues induces changes in nucleotide flexibility throughout tTR (Figure 2.9). Interestingly, the majority of nucleotides whose SHAPE reactivities are influenced are located in single-stranded/highly flexible regions. One region of tTR that seems to undergo restructuring in all of the mutants, though only slightly for the bulged U126, is the newly defined loop that includes nucleotides C88-A94. For all of these nucleotides, the mutants display decreased SHAPE reactivities in comparison to the wild-type. Other nucleotides whose flexibilities appear to change in multiple mutants are C43 from the tTR template sequence and G146 from the middle of the Stem IV helix (33). It is important to point out that many of the same nucleotides that exhibit differences in SHAPE reactivities for the U-mutants also displayed differential SHAPE reactivities when tTR was probed with slow-reacting and fasting-reacting reagents IA and 1M7 respectively.

4. 132-138 Mutants

Previous studies of tTR Stem IV via SHAPE and NMR identified that formation of a non-canonical C·U base pair that closes the apical loop of this helix and that mutation of this loop to an A-U Watson Crick base pair abolished telomerase activity (105). To further examine the significance of this interaction, we designed mutants that formed different canonical and non-canonical base interactions at this same location. Among the mutants designed were C132A and U138G that make Watson Crick base pairs, C132T that makes a U●U base pair, an interaction that has been shown to have a similar bond energy as a C●U base pair, and U138C and the C132A/U138G double

mutant that form C●C and A●G interactions respectively (Figure 2.10) (110). Upon completion of purification of wildtype and 132-138 132-138 mutant telomerases, we

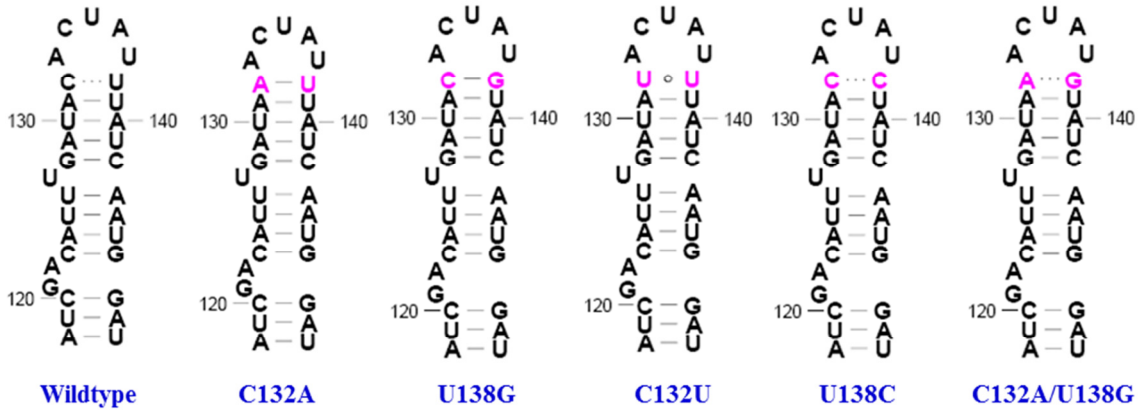


Figure 2.10. tTR stem IV wildtype and 132-138 mutants. Analysis of Stem IV identified that C132 and U138 interact and form a non-canonical base pair that is required for telomerase activity. To further examine the significance of this base interaction, we designed mutants that formed different canonical and non-canonical base interactions at this same location.

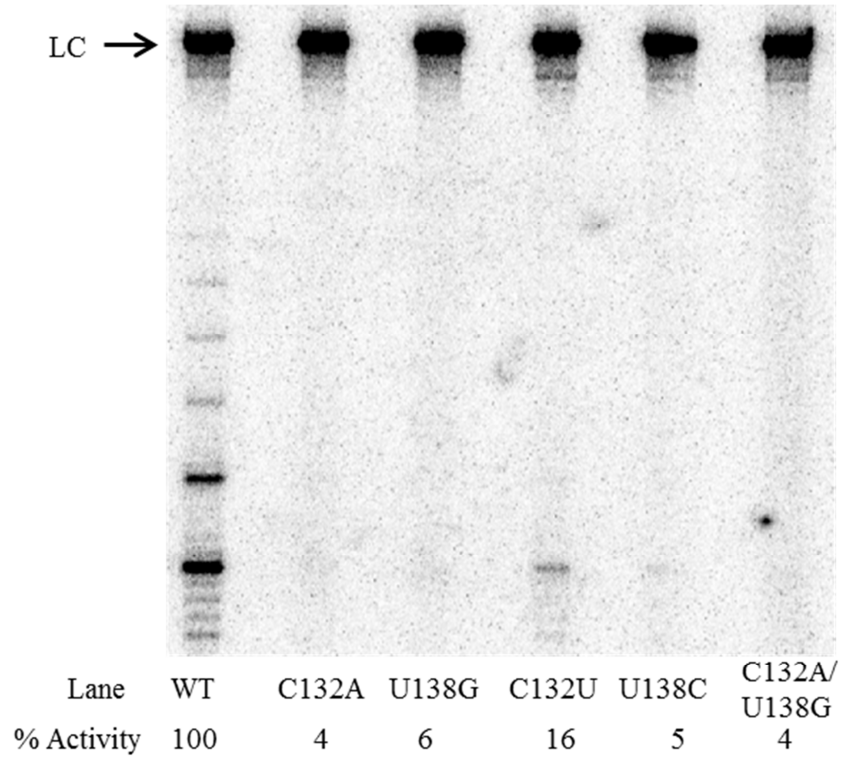


Figure 2.11. Direct telomerase activity assay tTR 132-138 mutants. A direct telomerase assay was conducted with purified telomerases incorporating wildtype and 132-138 mutant tTRs. The loading control (LC) at the top of each assay was used to normalize telomerase activity for each experiment.

performed a direct telomerase assay to assess the effect of the mutants on telomerase activity and processivity. Assessment of telomerase assays showed that all of the mutants abrogated telomerase activity and processivity, though the C132U mutant did show nominal activity at 16% of the wild-type (Figure 2.11). Repetition of the assay a second time generated similar results, confirming that none of these base pairs could sustain telomerase activity like the C●U base pair (data not shown).

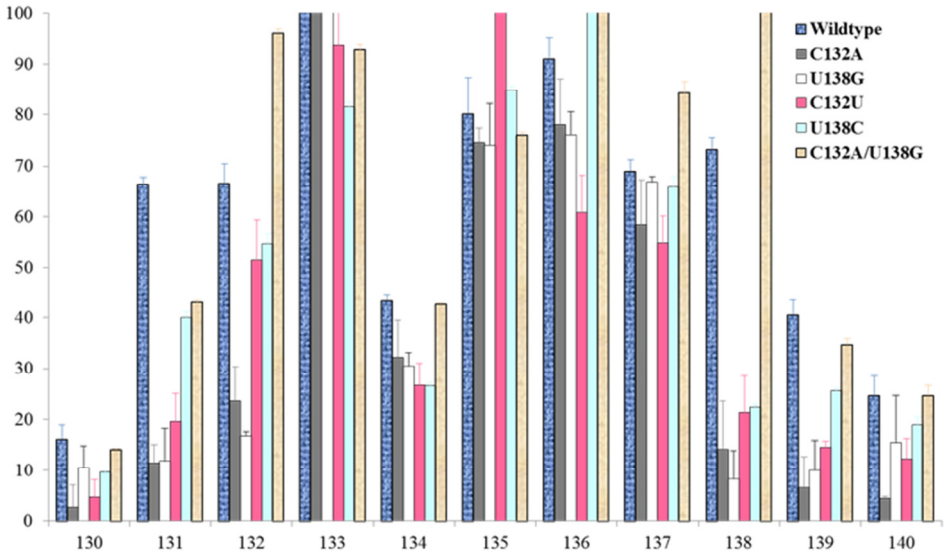
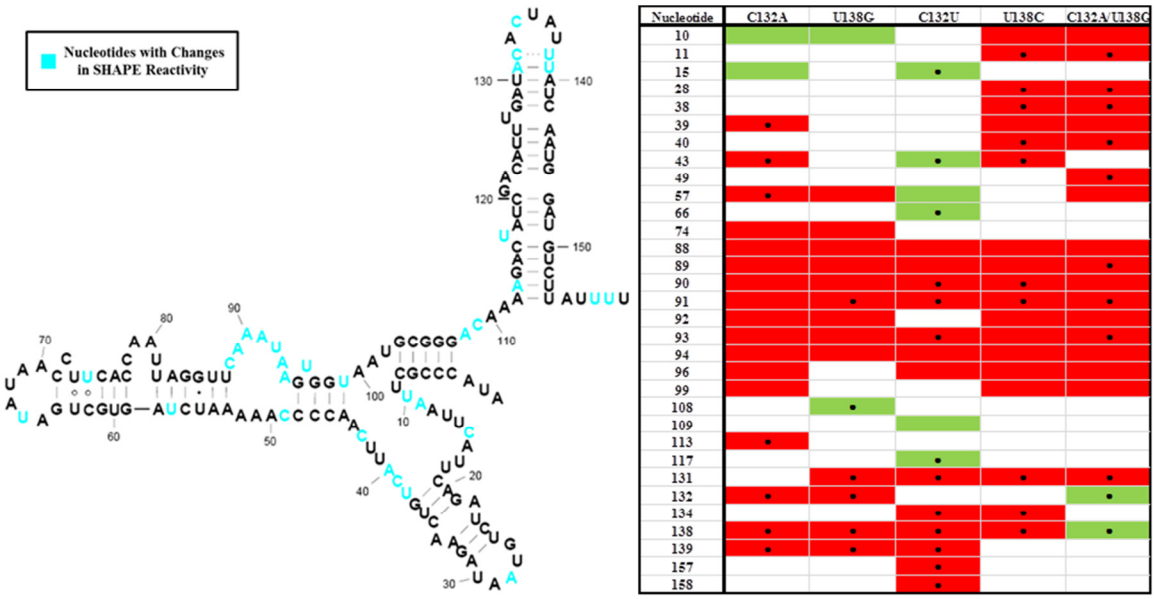


Figure 2.12. SHAPE reactivities of Stem IV apical loop nucleotide in 132-138 mutants. Histogram of SHAPE reactivities corresponding to nucleotides in the apical loop motif and flanking the 132-138 mutants. The average band intensities were normalized and plotted according to nucleotide position with error bars (N=3).

SHAPE analysis of adjacent nucleotides and of the apical Stem IV loop demonstrated that mutating nucleotides C132 and U138 did not have a large effect on the flexibilities of nucleotides in the single-stranded portion of the loop, A133-U137, or the base pair two below the nucleotides of interest, U130-A140. Formation of the Watson Crick base pairs by mutants C132A and U138 considerably reduced nucleotide flexibility for A131, U139, and nucleotides 132 and 138 (Figure 2.12). Interestingly, while nucleotide flexibility at nucleotide 132 remained similar to the wild-type for mutants

C132U and U138C, SHAPE reactivity at 138 was attenuated by the formation of U●U and C●C. In comparison to the C·U base pair, the A●G base pair created by mutant C132A-U138G was significantly more flexible (Figure 2.12). The changes in SHAPE reactivities were statistically significant for all mutants in comparison to wildtype at nucleotide 138 (Figure 2.12).



Nucleotide	C132A	U138G	C132U	U138C	C132A/U138C
10					
11				•	•
15			•		
28				•	•
38				•	•
39	•				
40				•	•
43	•		•	•	
49					•
57	•				
66			•		
74					
88					
89					•
90			•	•	
91		•	•	•	•
92					
93			•		•
94					
96					
99					
108		•			
109					
113	•				
117			•		
131		•	•	•	•
132	•	•			•
134			•	•	
138	•	•	•	•	•
139	•	•	•		
157			•		
158			•		

Figure 2.13. 132-138 mutants impact SHAPE reactivity of nucleotides throughout tTR. Mutation of nucleotides in the Stem IV helix of tTR affects nucleotide flexibility throughout the RNA. The secondary structure diagram highlights all of the nucleotides (collectively) that experience large changes in flexibility due to mutations. The chart to the right shows which nucleotides had changes in flexibilities for each of the 132-138 mutants. Red boxes identify nucleotides that become more structured in the mutant, whereas green boxes are more flexible in the mutant as compared to wildtype activity. The ● indicates that the difference in SHAPE reactivity is statistically significant with 95% confidence via the t-test. For nucleotides without the dot, there is insufficient evidence to indicate a difference in the SHAPE reactivities.

Inspection of the entire secondary structure of tTR demonstrates that mutating nucleotides C132 and U138 induces changes in nucleotide flexibility throughout the RNA (Figure 2.13). Just like in the U-mutants, most of the nucleotides experiencing changes in SHAPE reactivity are located in single-stranded/highly flexible regions. Also similar to the U-mutants is the increased structure seen among nucleotides in the newly defined loop that includes nucleotides C88-A94. Another region displaying increased

structure due to the 132-138 mutants is the single-stranded segment that links stem II to the now base paired template region. Other nucleotides whose flexibilities appear to change in multiple mutants are U10 from the linking region connecting the P1 and P1 helices and U57 in the extended P3 helix (Figure 2.13). Furthermore, like in the U-mutants, many of the same nucleotides that exhibit differential SHAPE reactivities showed similar results when tTR was probed with slow-reacting and fasting-reacting reagents IA and 1M7 respectively.

C. Discussion

The development of selective 2' hydroxyl acylation analyzed by primer extension (SHAPE) has impacted the manner in which the structure of long RNAs are being determined due to its versatility and ability to produce high resolution structural data for every nucleotide within a given RNA from a tiny sample size (85). For example, we have already been able to implement the procedure to determine the structure of the 159 nucleotide tTR in solution and propose a new secondary structure that supports the formation of an elongated stem III helix that incorporates the template region into base pairings and shows that the pseudoknot does not actually form (Legassie et al. 2012, *unpublished manuscript*). Being able to identify nucleotides with interesting characteristics that may play significant roles in RNA folding pathways or as molecular switches by utilizing SHAPE reagents with differential reactivity profiles only increases the wealth of knowledge we can surmise regarding tTR. Structural interrogation studies with 1M7, a fast-reacting derivative of isatoic anhydride with a strongly withdrawing para-nitro substituent that has been shown to increase adduct formation and hydrolysis

rates, have been able to accurately report known structures and detect nucleotides constrained by both base pair and non-canonical tertiary interactions (109). In contrast, investigations with IA, a slow-reacting electrophilic SHAPE reagent, have shown its ability to distinguish between and identify nucleotides that undergo slow conformational dynamics such as nucleotides that assume a C2' endo sugar pucker conformation or are more constrained due its longer reaction time (108). The slow local motions of the C2' endo nucleotide make these nucleotides unreactive to fast reagents like 1M7, but perfect for identification of nucleotides undergoing slow motion as the reagent is not so quickly hydrolyzed and can therefore remain in solution until local motions have been complete.

SHAPE analysis of full-length tTR with both 1M7 and IA identified 23 nucleotides that are demonstrated to have significant differences in SHAPE reactivity between 1M7 and IA (Figures 2.4 and 2.5). Interestingly, 15 of those nucleotides are identified to be more reactive to IA than 1M7 suggesting that the nucleotides undergo slow local conformational changes. NMR structural analysis of tTR helices have located five nucleotides, U27 U30, U135, A136, and U137, that are demonstrated to adopt C2' endo conformations (100,105). While SHAPE analysis does recognize the differences in SHAPE reactivities between 1M7 and IA for nucleotide U27 as statistically significant, comparatively the changes in reactivity are not even among some of the nucleotides with the largest changes (data not shown). Additionally, of the five nucleotides, differential SHAPE reactivities only predicted that U30 would adopt the 2' endo conformation correctly demonstrating that while local dynamics may contribute to selective acylation by slow reagents, that other factors are also involved. Though to actually say this prediction by SHAPE data was correct would be an understatement based on the data

(examined more at the end of the paragraph). In response to the argument that slow-reacting reagents may prefer constrained nucleotides, nucleotides U135, A136, and U137 have been shown to be solvent accessible and highly flexible. Most interesting of all though, is probably the case of U30. Shown to be in a C2' endo conformation by NMR in a constrained conformation with its base twisted and facing the minor groove, however, the SHAPE reactivities for this nucleotide have 1M7 much higher than IA. Current understanding of the fast-reacting SHAPE reagents in this case suggest that the electron poor ring structure of 1M7 makes the carbonyl carbon more reactive and then the electron poor ring would selectively react at certain structural motifs (electron poor base next to reactive hydroxyl). A complete understanding of the chemistry for this situation has yet to be specifically uncovered though, allowing for speculation as to the nature of U30.

A third proposal is that nucleotides that exhibit high reactivities to IA compared to 1M7 identifies nucleotides that can be labeled as 'hot spots" of RNA folding and may be for critical rate-limiting steps in RNA organization. Nucleotides 91-96 are shown to have increased IA reactivity and are located in a region that is known to fold into the pseudoknot structure upon binding of tTERT. This loop region has also been implicated in making interactions with stem IV after its bending and repositioning near the telomerase active site. This allows the proposal of a model where upon binding of p65 to stem IV, the apical loop is repositioned near these looped nucleotides in the extended stem III. The stem III looped nucleotides are undergoing slow local dynamics until they are put into proximity with stem IV which then elicits a change in conformation of the looped nucleotides which allows for a rate-limiting step and then conversion from the

hairpin to the catalytically active tertiary pseudoknot formation along with tTERT binding. Until more is known about the reactivity of the slow-reacting SHAPE reagents, these proposals are mostly speculation. Ultimately, the data we have shows that the highest correlation with reactivity does not correspond with one specific conformation, but with a residue's ability to sample many different conformations.

Both structural NMR studies and mutational and biochemical studies support the ideas that the stem IV helix and loop are essential for a multitude of functions in *Tetrahymena thermophila* telomerase including but not limited to binding p65 and tTERT, promoting the formation of the catalytic pseudoknot, and playing an important role in both NAP and RAP. Therefore, characterizing unique structural features of the RNA will only aid in understanding their role in proper stem IV and loop biological function. Though NMR analysis of the bugled Us in the upper part of stem IV only recognize U126 and U127 as possible bulged nucleotides, SHAPE analysis of full-length tTR in our lab demonstrates that there is configurational flexibility between three U residues, U125, U126, and U127 (105). Interestingly, the Collins lab saw a decrease in overall telomerase activity upon deletion of bulged nucleotide U127, which may or not be attributable to the slight loss in helix bend that was induced from the bulged nucleotide (104,105). In mutants that we made, neither forcing each of the individual U residues to bulge, or flipping the Us to the 3' side of stem IV, nor eliminating the U residues altogether had any effect on telomerase activity (Figure 2.7). Similar results were seen for NAP and RAP in our mutant assays. SHAPE data analysis of those nucleotides and neighboring ones saw no drastic changes in flexibility (Figure 2.8) which leads us to propose that either: 1. local conformational flexibility of the region that includes the U

nucleotides is important, or 2. height of the helix that the U nucleotides cover is significant for telomerase activity. The height of the U nucleotides may position the apical loop just high enough that when the helix is bent over, it places the correct nucleotides in exactly the correct spatial location for them to properly interact in catalysis. Making mutants that have either 0-1 or 3-5 A-U base pairs in this region could easily test the second hypothesis. In the scope of the whole RNA, U mutants mostly appear to affect the structure of nucleotides that are single-stranded. All of the U-mutants appear to decrease SHAPE reactivities of nucleotides in newly formed C88-A94 stem III loop, suggesting that this is a contact point of interaction between stem IV and stem III, though it is also possible that changes throughout tTR structure are due to the dynamic nature of RNA (Figure 2.9).

In both NMR structures of stem IV currently available, a non-canonical base pair between nucleotides C132 and U138 is shown to be the apical loop closing base pair (104,105). The significance of this base pairing interaction is highlighted by mutational evidence that shows that changing the nucleotides to an A-U Watson Crick base pair completely abolished telomerase activity (105,106). In keeping with the initial idea, we created mutants that incorporated other possible base pairing combinations in order to see if other interactions may be able to maintain the interaction caused by the C•U base pair. Much like the A-U Watson Crick base pair, G-C, C•C, and A•G base pairs had negligible amounts of telomerase activity (Figure 2.11). This data agrees with other mutational studies that demonstrate that mutation of U138 to any other nucleotide attenuates telomerase activity (106). The only mutant that retained any amount of telomerase activity was a U•U base pair that maintained 16% telomerase activity; RNA base pairing

experiments have shown that U•U base pairs have similar bond energetics to C•U base pairs, however, in the case of this mutant the amount of telomerase activity was not as similar as the bond energies are to one another (110,111). While local SHAPE reactivities did change in comparison to the wildtype, none of the reactivities were similar enough to maintain the activity elicited by a C•U base pair. Interestingly, the SHAPE reactivity for the U•U base pair was very similar to the flexibility of the C•C interaction, but the C•C was much less active than the U•U (Figure 2.12). These results also lead to the proposal of two hypotheses: 1.the C132-U138 base pair interaction is sequence specific and no other nucleotide combinations will be able to recapitulate the activity of C•U, or 2. the motif is actually two or three base pairs and the conformation they create with C•U stacked on top or below is the defining structural feature. Similar to the U-mutants, when looking at long-range interactions in tTR, the 132-138 mutants also increased structure in the C88-A94 single stranded loop region. These mutants also appear to interact with conserved residues of the TBE, decreasing structural flexibility in that region as well (100,101) (Figure 2.13).

D. Materials and Methods

1. Oligonucleotides and chemicals

Oligonucleotides were purchased from Integrated DNA Technologies (IDT, Coralville, Iowa). The oligonucleotides were then purified via denaturing polyacrylamide gel electrophoresis (PAGE) using the modified “crush and soak” method (112,113). Following electrophoresis in a 12% denaturing gel, oligonucleotides were observed via ultraviolet (UV) light, cut out with a razor blade, crushed through a 3 mL sterile plastic

syringe, and eluted from the gel in 1×TEN buffer (10 mM Tris pH 7.5, 1 mM EDTA, 250 mM NaCl) for one hour at room temperature. After centrifugation, the supernatents were collected and passed through a 0.22 micron MCE membrane syringe filter (FisherBrand) to remove remaining gel particulate. Oligonucleotides were then ethanol precipitated with final concentration 0.6 M ammonium chloride, concentrated, and resuspended in TE (10 mM Tris-Cl pH 8.0, 1 mM EDTA) or DEPC-treated water (Ambion, Inc., Austin, Texas) to a stock concentration of 1 mM. The concentration of further dilutions was determined by UV absorbance at 260 nm using the molar extinction coefficient supplied by the manufacturer.

N-methyl-isatoic anhydride (NMIA) was purchased from Invitrogen Molecular Probes (Eugene, Oregon), isatoic anhydride (IA) was purchased from ACROS Organics (Gael, Belgium), and 1-methyl-7-nitroisatoic anhydride (1M7) was a gift from Dr. Kevin Weeks (UNC Chapel Hill, NC, USA). These compounds were made freshly at 50 mM stocks in DMSO ACS reagent grade (MP Biomedicals, Solon, Ohio).

2. tTR and tTERT expression plasmids

Tetrahymena thermophila TERT was cloned into the BamH1 and Xho1 sites in pET-28a plasmid. The FLAG peptide sequence was added to the N-terminus of tTERT by cloning into the Nco1 and BamH1 sites by Dr. Jason Legassie. The addition of the FLAG-tag removed the N-terminal His- and T7- tags. pTET-telo, a plasmid containing the tTER gene, a promoter for T7 RNA polymerase, and a self-cleaving hammerhead ribozyme that processes the 5′ end of the RNA to generate wild type tTER was a gift from Dr. Art Zaug (103). Site directed mutagenesis, based on the Stratagene Quick

Change site-directed mutagenesis protocol, was used to generate pTET-telo constructs coding for Stem IV mutants.

3. PCR construction of dsDNA tTR construct

In order to make the tTR RNA, dsDNA templates were generated by PCR using primers that overlapped at the 5' end of the Hammerhead ribozyme and the 3' sequences of the tTR gene. The forward primer contains a T7 promoter before the initial Hammerhead ribozyme sequence while the reverse primer had the 3' linker and conRT binding site incorporated after the terminal tTR sequence (85).

Table 2.1: Primers for cloning of tTR

tTR Primer	Sequence (5'-3')
tTR Forward	TCTAATACGACTCACTATAGGG
tTR Reverse	GAACCGGACCGAAGCCCGATTTGGATCCGGCGAA CCGGATC GAAAAATAAGACATCCATTG

Primers in the presence of the pTET-telo expression plasmid underwent the following PCR cycles: 1. 95 °C for 30 seconds enzyme hot-start; 2. 30 cycles of 95 °C for 30 seconds, 55 °C for 30 seconds, 72 °C for 1 minute to yield a full length double stranded product; and 3. 72 °C for 5 minutes to complete any unfinished ends. Upon completion, PCR products were PCI extracted, ethanol precipitated, concentrated, resuspended in TE, and concentration was determined by UV spectroscopy.

4. *In vitro* transcription and purification of tTR

tTRs were made via *in vitro* transcription with the T7 Ampliscribe Transcription Kit (Epicentre Biotechnologies, Madison, WI). The *in vitro* transcription reaction was carried out overnight at 37 °C with one microgram dsDNA template, 1x Ampliscribe T7

Reaction Buffer, 7.5 mM ATP, 7.5 mM CTP, 7.5 mM GTP, 7.5 mM UTP, 10 mM DTT, ScriptGuard RNase Inhibitor, and 2 μL Ampliscribe T7 Enzyme Solution. The next morning, the reaction incubated with 1 μL RNase-Free DNase 1(1MBU/μL) for 15 minutes at 37 °C. Cleavage by the hammerhead ribozyme was initiated by the addition of $MgCl_2$ to a final concentration of 20 mM and incubated at 45 °C for one hour. The reaction was resuspended in an equal volume SHAPE denaturing loading buffer (80% formamide, 0.5×TBE, 4 mM EDTA, 0.01% bromophenol blue, 0.01% xylene cyanol) and purified in an 8% denaturing polyacrylamide gel. The RNA was then recovered by the modified "crush and soak" method as explained above; the only differences being that the RNAs are ethanol precipitated with final concentration of 0.3 M of sodium acetate and final RNA concentrations were determined by nanodropping and RNA molecular mass.

5. Translation of tTERT and telomerase assembly

tTERT was translated and assembled with tTR using a TNT Coupled Reticulocyte Lysate Systems kit (Promega) based on a previously described procedure (114). A typical 50 μL reaction contained 1 μg of pET-28a-tTERT, 70 ng of tTR, 34 pmole of [^{35}S]-methionine (1175 Ci/mmol; Perkin-Elmer) and additional reaction kit components provided by the manufacturer. Reactions were incubated at 30 °C for 90 min and were flash frozen in an ethanol/dry ice bath and stored at –80 °C. Five microliters of each lysate reaction was analyzed by SDS PAGE to make sure full length protein was translated. Samples were denatured in an equal volume of 2× Laemmli's Buffer (100 mM Tris-HCl pH 6.8, 200 mM DTT, 20% glycerol, 4% SDS, and 0.05% bromophenol blue) and heated at 95 °C for 5 min. The denatured samples were run on a 6%

acrylamide/SDS gel (115). The gel was transferred to and dried on Whatman Chromatography paper (GE Healthcare) for 1 hour at 80 °C on a Model 583 Gel Dryer (Biorad) and then exposed on a Phosphor Screen (Amersham Biosciences) overnight. Gels were imaged on a Storm 860 Phosphor Imager (Molecular Dynamics, Sunnyvale, CA) and visualized on ImageQuant TL (Amersham Biosciences).

6. Immunopurification

Flag-tagged *Tetrahymena thermophila* telomerase complexes were immunoprecipitated with anti-FLAG M2 affinity agarose beads (Sigma-Aldrich, St. Louis, MO). Prior to use, beads (50 µL) were washed four times in 400 µL of 1xTris IP buffer (10 mM Tris-Cl, pH 7.5, 100 mM KCl, 1 mM $MgCl_2$, 10% glycerol, 0.1 mM DTT), centrifuging at 1500 x g for 2 minutes between washes. Reticulocyte lysate translation and telomerase assembly reaction (45 µL) was added to 450 µL of 1xTris IP buffer and centrifuged at 16,000 × g for 15 minutes at 4 °C to remove any particulates. The supernatant was added to beads and allowed to nutate at 4 °C for 2 hours. The beads were washed six times in 400 µL 1×Tris IP buffer centrifuging at 1500 × g for 2 min between washes. Beads were washed once with 500 µL TMG (10 mM Tris acetate, pH 8, 1 mM $MgCl_2$, 10% glycerol) and transferred to Protein LoBind tube (Eppendorf) before centrifuging at 1500 × g for 2 minutes. Upon removal of the TMG, BSA (3 µL of 10 mg/ml) was added to the beads to prevent protein from sticking to the tubes. Fifty microliters of FLAG peptide solution [41 µL TMG, 9 µL FLAG peptide (5 mg/ml)] was added to the beads and nutated at 4°C for 1 hour. FLAG peptide (5 mg/ml) was prepared by adding 3×FLAG peptide (Sigma-Aldrich) to 20 mM Tris acetate, pH 7.5. The beads

were centrifuged at 1500 × g for 2 minutes and the eluent was removed, stored in a fresh LoBind tube, and flash-frozen for storage at -80 °C.

7. Direct telomerase activity assay

Telomerase activity was measured by a direct telomerase activity whose protocol was adapted from a previously described assay (113). Each 20 μL reaction contained telomerase reaction buffer (50 mM Tris-Cl pH 8.3, 1.25 mM $MgCl_2$, 5 mM DTT), 2 μM telomeric p5 primer 5'-GTTGGGGTTGGGGTTGG-3', 100 μM dTTP, 10 μM dGTP, and 0.33 μM [α- 32P]-dGTP (3000 Ci/mmol, 10 μCi/μL; Perkin Elmer) and 10 μL FLAG-immunoprecipitated telomerase from RRL reactions with the remaining volume being DEPC-treated water. Primer extension was carried out at 30 °C for 90 minutes. After primer extension, the reaction volume was increased by the addition of water; a [^{32}P]-labeled loading control (114 nucleotide, 5'-end labeled ssDNA oligonucleotide, 1000 cpm per reaction) was also added. Telomerase assay primer extension products were extracted once with phenol/chloroform/isoamyl alcohol (25:24:1) and twice with chloroform/isoamyl alcohol (24:1) and then ethanol precipitated with 2.5 volumes absolute ethanol, 0.6 M ammonium acetate, and 200 μg/mL glycogen. After precipitating at -20 °C for 15 minutes, telomerase assay products were centrifuged at 13,000 x g at 4 °C for 22 minutes, washed with 1 volume of 70% ethanol, and speed vacuumed till dry. Pellets were resuspended in 6 μL SHAPE loading buffer and heated at 95 °C for 5 minutes and resolved on a pre-warmed, 20 cm x 20 cm x 0.4 mm thick 8% denaturing polyacrylamide gel. The gel was run in 0.5× TBE at 600 volts for approximately one hour. The gel was transferred, dried, exposed, and visualized similar to the SDS/PAGE

gel mentioned above. Summing the intensities of all bands in each sample and normalizing the values to the loading control determined overall telomerase activity of the reaction.

8. IA, NMIA, and 1M7 hit reactions

One picomole of RNA (1 μL of 1 μM RNA) in DEPC-treated water was snap-annealed at 95 °C for 3 minutes and then incubated on ice for 5 minutes to allow RNAs to adopt the correct conformation. 5× tTR Hit Buffer (250 mM Hepes Buffer Solution (GIBCO), 10 mM RNase-free $MgCl_2$ (Ambion)) diluted to 1× was combined in the microcentrifuge with the snap-annealed RNA and allowed equilibrate at 37 °C for one minute. The RNA was then treated with either 50 mM IA, NMIA, or 1M7 to a final concentration of 5 mM, or DMSO to a final concentration of 10% as a control, and incubated at 37 °C for 45 minutes, 5 minutes, or 70 seconds respectively for IA, NMIA, and 1M7. The hit reaction was quenched with RNase-free NaCl (Ambion) to a final concentration of 200 mM and precipitated with 200 μg/mL RNase-free glycogen (Ambion) as a counter ion. Following centrifugation, the hit RNA was washed once with 70% ethanol, speed vacuumed till dry, and the resuspended in 10 μL of RNase-free TE buffer pH 8.0 (Ambion).

9. Superscript III primer extension

Table 2.2: SHAPE Primers for mapping tTR

SHAPE Primers	Sequence (5'-3')
c103	GATAGTCTTTTGTCCCGC
conRT	GAACCGGACCGAAGCCCG

Due to the size of tTR, 159 nucleotides, two primers (in separate reactions, not together) were used to in order to map SHAPE chemistry adduct formation. Half a picomole of IA, NMIA, or 1M7 modified RNA (5 μL) was mapped by annealing 1 μL 5'-[^{32}P]-labeled primer via the following thermocycler conditions: 95 °C for 1 min, 65 °C for 6 min, and 35 °C for 10 min. Following a 5 minute incubation on ice, RNAs were supplemented with 5× First Strand Buffer (250 mM Tris-Cl pH 8.3, 375 mM KCl, 15 mM MgCl2) (Invitrogen), 0.1 M DTT to a final concentration of 5 mM, and 10 mM dNTPs to a final concentration of 1.25 mM. These contents were then heated at a temperature of 50 °C on an Isotemp 125D (Fisher Scientific) for one minute. One microliter of Superscript III reverse transcriptase (100 units, Invitrogen) was immediately added to the reaction mixture, mixed by gentle tapping, and allowed to extend for exactly 4 minutes at 50 °C. The reaction was quenched by the addition of 400 mM NaOH and then heated at 95 °C for 5 minutes in order to degrade the RNAs. This was then neutralized by the addition of 400 mM HCl and radiolabeled cDNAs were ethanol precipitated with 3 M sodium acetate and 200 μg/mL glycogen and then resuspended in 6 μL SHAPE loading buffer. Dideoxythymidine sequencing ladders were generated by the addition of 1.25 mM ddTTP (GE Healthcare) to the reverse transcription reaction of 3.5 picomoles of unmodified RNAs.

10. Sequencing gel electrophoresis

The radiolabeled cDNAs created by the reverse transcription of NMIA/1M7/DMSO hit RNAs were resolved on a 35 cm x 43 cm x 0.4 mm 8% denaturing polyacrylamide sequencing gel (29:1 acrylamide; bisacrylamide/7M urea, 90

mM Tris/borate, 2 mM EDTA) run at approximately 2000 volts in 0.5× TBE for about an hour an a half. The gel was transferred, dried, exposed, and visualized similar to the SDS/PAGE gel mentioned above.

11. SAFA data analysis and normalization

Individual band intensities of IA/NMIA/1M7, and DMSO lanes were quantified using the program SAFA (88). SAFA allows the straightening of curved gels (gel recitification) and uses Lorentzian curve integration to determine band intensities with a high degree of accuracy. Full length RNA bands were quantified by employing the analysis method in ImageQuant TL. Radioactivity across the lanes of an individual experiment was then equalized to the level of radioactivity in the NMIA/1M7 lane. Band intensity was calculated by subtracting the value of the DMSO band, control reaction that corresponds to background activity, from the value of the NMIA or 1M7 band giving the corrected band intensity. Corrected band intensities were then ranked in descending order and the top 2% of values was thrown out. The average of the next 8% of values was taken and this value represented the normalization factor. All of the IA/NMIA/1M7 corrected densities were divided by the normalization factor and then multiplied by 100 to give the normalized values. Normalized values from multiple replicates were averaged and plotted in figures with standard error bars shown.

CHAPTER 3

SHAPE ANALYSIS OF THE PSEUDOKNOT, CR4-CR5, AND CR7 DOMAINS OF hTR

A. Introduction

Comparison among telomerase RNAs reveals that different species display great diversity in primary sequence size and little sequence homology (48). However, telomerase RNAs do share phylogenetically conserved secondary structural elements. The four universally conserved regions within the human telomerase RNA (hTR) are the template containing pseudoknot domain (CR1-CR3), the hypervariable region and trans-activation domain (CR4-CR5), the Box H/ACA domain (CR6/CR8), and the Ku binding domain (CR7) (Figure 3.1) (33,48). Functionally speaking, the 5' end of hTR encompassing nucleotides 1-370 (the pseudoknot and trans-activating domains) is essential for telomerase nucleotide addition and processivity, while the 3' end (Box H/ACA and CR7 domains) plays a role in hTR maturation and movement throughout the cell.

The current proposed secondary structure of hTR has been derived from phylogenetic comparison and compensatory mutational analysis (31,48). More recently, studies of isolated domains of hTR have been conducted by NMR and reveal compelling structural data that is updating ideas regarding hTR architecture (40-42,47,49). Leeper et al. showed that the U306•G310 wobble pair in the P6.1 helix stacks above the A-form

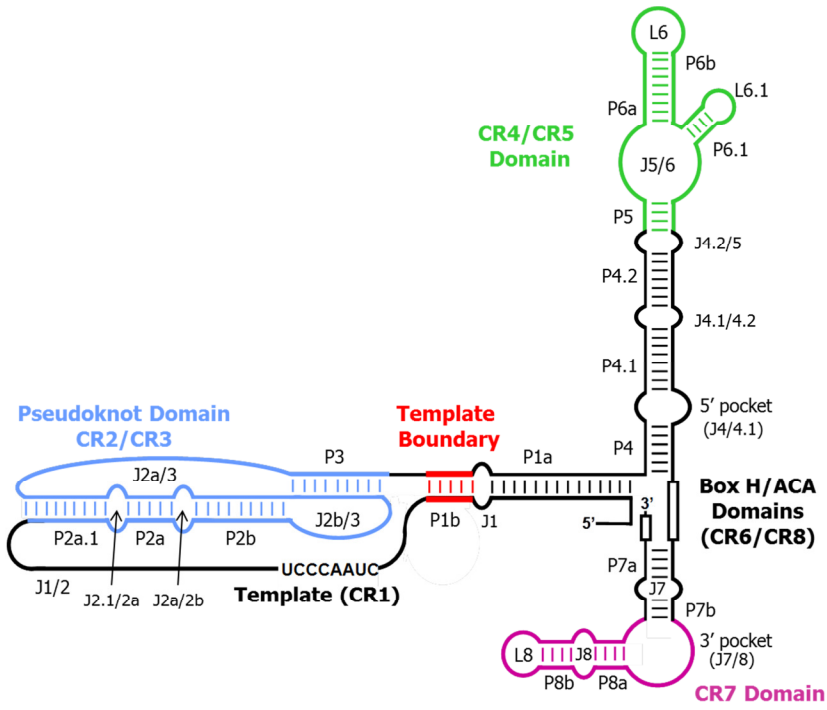


Figure 3.1 Phylogenetic hTR secondary structure. Schematic of hTR secondary structure with names and locations of important conserved regions, domains, helices, and loops. Adapted from Legassie et al. 2006 (33).

four nucleotide stem helix and that the backbone from residues G308 and G309 extend out forming hydrophobic interactions that stabilize the stem loop and make it a fairly rigid structure (49). They later demonstrated that a triple base pairing between nucleotides C267, G268, and C288 near the J6 internal loop of the CR4-CR5 domain of hTR allows for the formation of a solvent accessible surface through the middle of this internal loop that may act as a protein docking site (47). Taking motivation from this study and initiative in search of our own compelling results to better understand hTR structure, we decided to investigate the structure of a minimally conserved pseudoknot construct, the entire CR4-CR5 domain, and the complete CR7 domain using a high resolution, single nucleotide RNA footprinting technique, SHAPE.

Previous studies of hTR secondary structure have identified the pseudoknot domain as being required for catalytic telomerase activity (97). Due to its proximity to

the template region, it has also been proposed to play a vital role in lining up the template region with the telomere (DNA primer) as well as orienting the primer-template towards the active site (37). Though the effects of mutations and deletions to RNA sequence in this region have been extensively investigated, there is little secondary structure information available for the complete pseudoknot region (nucleotides 1-209) due to its size (28,42,116). Within the pseudoknot domain, certain stems and loops (p2b, P3, J2b/3, and the 3' end of J2a/3; nts 90-125 and 170-184) show high levels of conservation (48). Thermodynamic experiments revealed that the wildtype pseudoknot construct composed of these conserved regions was in equilibrium with a hairpin conformation that has been insinuated to be involved in pseudoknot-dependent telomerase activity (42,117). Removal of the bulged nucleotide U177 from the P3 helix alleviated this problem and pushed the structural equilibrium back to favoring the pseudoknot structure (41). Mutations within this region can drastically change the structure of the pseudoknot, by shifting the equilibrium or preventing the formation of key secondary and tertiary interactions, and attenuate telomerase activity (42,59,118). A known two nucleotide substitution, GC 107-108 AG, gives rise to the autosomal dominant form of the premature aging disease dyskeratosis congenita (31).

Like the pseudoknot domain, the CR4-CR5 domain of hTR is required for telomerase catalytic activity and serves as a proposed hTERT binding site (45,119). The CR4-CR5 domain is relatively large, spanning nucleotides 242-328, and has not been thoroughly characterized despite its high level of phylogenetic conservation (32,33). Most studies investigating its structure have been completed by NMR and have focused on isolated helices including P6a, P6b, and P6.1 and have used G-C clamps at the base of

the helices to afford stability or enhance transcription efficiency (47,49). There are limitations with this cut and paste method of RNA structural interrogation especially since studies have suggested that bulges and loops within the elongated stem of the CR4-

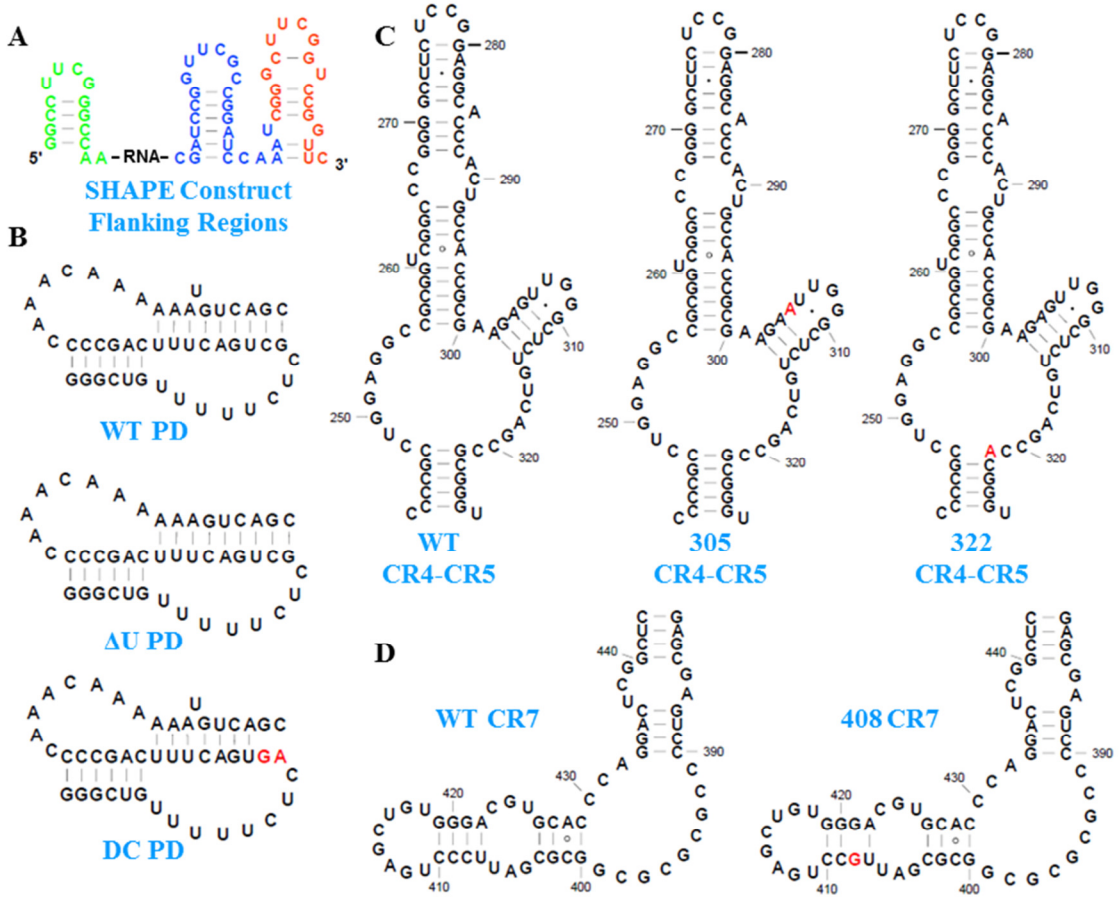


Figure 3.2 SHAPE construct and hTR wildtype and mutant domains. **(A)** In order to complete the SHAPE reaction, 5' and 3' flanking regions were added to the RNA of interest. The green nucleotides are the 5'-linker, the blue nucleotides are the 3'-linker, and the orange nucleotides are the SHAPE primer binding site (85). These are the phylogenetic structures of the three hTR domains examined in this chapter along with the mutants in the same region. **(B)** The abridged pseudoknot domain construct includes highly conserved nucleotides from the core domain. Mutants include ΔU177 and GC 107-108 AG, a two nucleotide substitution that leads to dyskeratosis congenita (DC). **(C)** The CR4-CR5 domain contains nucleotides 243-328. Two mutations that convert G to A at 305 and 322 give rise to a less severe form of DC known as aplastic anemia (AA). **(D)** The CR7 domain is important for hTR maturation and accumulation in cells, and C408G mutant of CR7 which causes DC and abrogates telomerase activity in cells.

CR5 domain may cause kinks or turns in the helix that may produce additional secondary and tertiary RNA interactions between the isolated regions (47). Similar to the

pseudoknot domain, mutations in this domain, specifically G305A and G322A, can give rise to aplastic anemia (AA), a pre-mature aging disease that is closely related to dyskeratosis congenita but much less severe (Figure 3.2). Though functional repercussions of these mutations such as a decrease in telomerase activity have been identified, it is still not clear whether this is due to perturbations of the CR4-CR5 domain structure (60,62,118).

Unlike the pseudoknot and CR4-CR5 domains, the CR7 is dispensable for telomerase activity *in vitro* (52). In cells, though, the CR7 domain is required for proper hTR maturation, localization, and accumulation as directed by the dyskerin and other protein complexes that bind to it (56,57). Even though the CR7 domain is more than 90% structurally conserved in mammals (57), most of the structural data regarding this region in hTR has been inferred from similar sequences in other scaRNAs (54). Like the pseudoknot domain, this region also contains a dyskeratosis congenita mutant, G408C, that has been shown to drastically decrease telomerase accumulation and activity in cells (56,62).

In this chapter, we utilized SHAPE chemistry to assess and determine the secondary structure of three critical domains within hTR: 1) the minimal pseudoknot domain, 2) the CR4-CR5 domain, and 3) the CR7 domain. Additionally, mutants from each of these domains known to give rise to premature aging diseases (dyskeratosis congenita and aplastic anemia) were also investigated to determine how/if changes in the secondary structure could lead to disease progression. Our studies were able to identify with high resolution the structures adopted by individual hTR domains and mutants. Changes in the primary structure gave rise to reorganization of secondary structure in the

case of all the mutants that were tested. Most importantly, the data confirms that RNA structure, particularly the structure of specific domains within hTR, directly correlates with their biological function.

B. Results

1. Addition of 5' and 3' flanking regions to hTR domains for SHAPE

In order to examine the wildtype and mutant hTR domains by SHAPE analysis, 5' and 3' flanking regions (Figure 3.1A) were added to the domain sequences (85). The 5' flanking region contains 14 nucleotide linker (green nucleotides) that allows for complete reverse transcription of the RNA interest all the way to its 5' end so that no data from the initial nucleotides is obscured (85). The 3' flanking region comprises of an 18 nucleotide reverse transcription primer binding site (orange nucleotides) at the very 3' end followed by a 24 nucleotide linker (blue nucleotides) which allows for the reverse transcriptase to become fully processive before encountering the RNA of interest (85). Each of these nucleotide additions to the RNA of interest form hairpin structures and therefore do not interfere with native RNA folding.

2. Wildtype and ΔU177 mutant adopt pseudoknot structures

Analysis of the wildtype and ΔU177 pseudoknot SHAPE data indicates that the two pseudoknots fold similar to one another with exception of area around nucleotide U177 (Figure 3.4A,B). Like the phylogenetically and mutationally derived secondary structure of hTR, nucleotides 93-98 have low flexibilities that correspond to their participation in the forming of the P2b helix of the pseudoknot (Figure 3.4A,B). SHAPE data

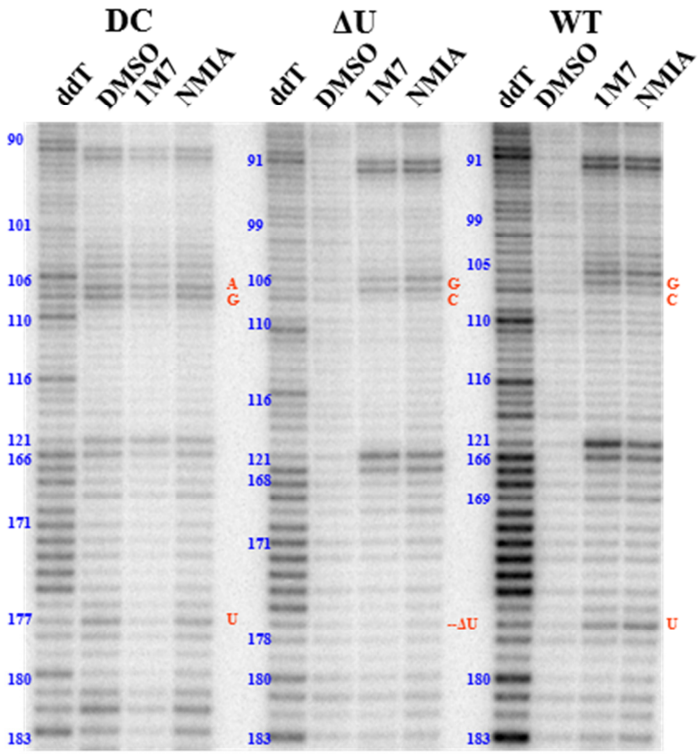


Figure 3.3 SHAPE gel of wildtype and mutant pseudoknots. Flanked pseudoknot RNAs were treated with 5 mM NMIA, 5 mM 1M7, or 10% DMSO (background) and the acylated RNAs were mapped by reverse transcription. The ^{32}P-radiolabeled cDNAs were resolved on a denaturing sequencing gel. Nucleotides "hit" during SHAPE cause reverse transcription to stop one nucleotide before the modification site. A dideoxythymidine (ddT) ladder (serves as a marker for orientation and nucleotide identification) was also run and indicates the location of adenines in the RNA sequence. The gel is numbered according to pseudoknot domain nucleotide number, not the ddT ladder which is off by +1 nucleotide.

demonstrates that U's 99-103 have significantly lower flexibilities than would be expected from single stranded nucleotides involved in a loop structure. This data supports their participation in making the evolutionarily conserved Hoogstein triples with the P3 helix (33,41,44). Nucleotides 104-107 appear to form a shortened loop J2b/3 as reasoned from the high flexibilities and G107 through C120 contribute to the formation of helices P2b and P3 (Figure 3.4A,B). Despite its proposed base pairing with G93 which is also slightly reactive, C121 is highly reactive which is unexpected. The eight nucleotide J2a/3 loop is also downsized possibly only including C166-A169, despite the fact that nucleotides A167 and A168 which would be in the middle of the loop appear to

be quite lowly reactive. Just like U99-U103 in the J2b/3 loop, nucleotides C170-A173 have diminished SHAPE reactivities likely due to their participation in non-canonical Hoogstein triples with helix P2b (Figure 3.4A,B) (41,44). The low reactivity values for A174-A176 and G178-C183 in the wildtype and A174-C183 in the ΔU177 mutant in consistent with these nucleotides forming the P3 helix (Figure 3.4A, B). In the wildtype pseudoknot, nucleotide U177 has very high SHAPE reactivity commensurate with its being a bulged nucleotide (Figure 3.4A). Nucleotides A176 and G178 are also slightly flexible suggesting that the conformation assumed by the U177 makes their bases more accessible to SHAPE modification (40). Altogether, the SHAPE data supports the formation of a pseudoknot structure for both the wildtype and ΔU177 mutant RNAs.

3. DC pseudoknot is actually a hairpin

Unlike the wildtype and ΔU177 constructs which form pseudoknot structures, comparison of the SHAPE profile of the DC construct attests to disruption of the vital pseudoknot structure and supports previous studies implying that the GC 107-108 AG mutation shifts the equilibrium of the RNA toward a tightly zipped hairpin conformation (42). The DC pseudoknot SHAPE structural data shows formation of the P2b helix like in the wildtype pseudoknot, consisting of nucleotides G93-G98 base pairing to C116-C121. From this point on, drastic structural rearrangement in RNA architecture is denoted from the SHAPE reactivity profile (Figure 3.4C). While U's 99-103 maintain decreased reactivities, it is not due to the formation of Hoogstein base triples with helix P3, but instead an extension of the P2b helix which is now 12 base pairs long. This elongated helix also includes U103 and C104, which were originally part of the shortened J2b/3 loop, but have strikingly high reactivities compared to other nucleotides in the stem (Figure 3.34). The four nucleotide J2b/3 loop seen in the wildtype construct is expanded and now makes a five nucleotide hairpin loop, U105- U109, all of which are highly reactive.

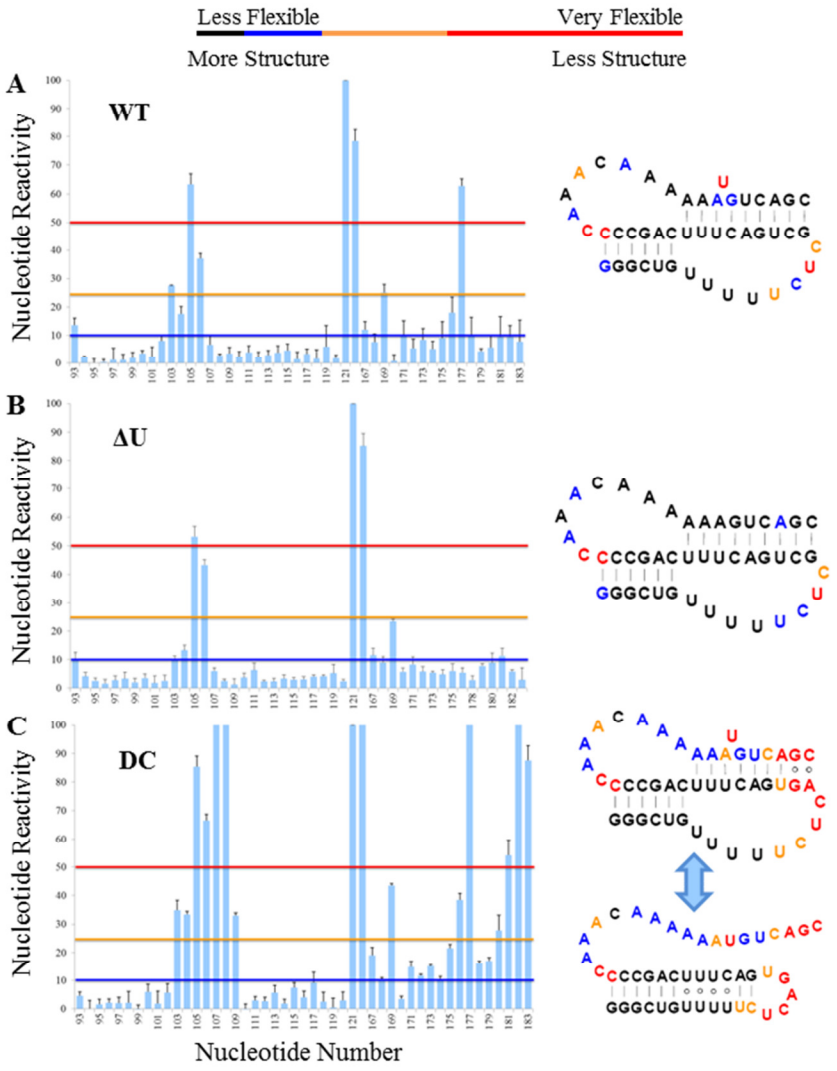


Figure 3.4 SHAPE analysis of the wildtype and mutant pseudoknots. Histograms and color-coded secondary structures of the wildtype, ΔU177, and DC pseudoknots. The average band intensities were normalized and plotted according to nucleotide position with error bars (N=4). Black, blue, orange, and red lines indicate divisions for examining nucleotide reactivity values where 0 ≤ black < 10, 10 ≤ blue < 25, 25 ≤ orange < 50, and red ≥ 50. Lower numbers represent less reactivity and higher numbers represent greater reactivity. The color-coded secondary structures reflect the histogram data for each RNA examined. **(A)** The wildtype pseudoknot. **(B)** The ΔU177 pseudoknot. **(C)** The DC pseudoknot; previous studies suggest that this pseudoknot actually adopts a hairpin conformation(42) and is thus represented in both conformations.

Nucleotides C110-C120 exhibit low flexibilities consistent with base pairings and the formation of the extended stem structure (Figure 3.4C). All the nucleotides from this point on, with the exception of the anomalous C170, display moderate to high flexibilities indicating the entire region is likely single stranded.

4. ΔU177 and DC mutations to pseudoknot diminish telomerase activity

To test the effect of the pseudoknot mutants on telomerase activity, we transfected HEK293T cells with hTERT and hTR plasmids to overexpress telomerase. Once cell extracts were collected, we performed a direct telomerase assay to assess the degree to which mutating the pseudoknot domain impaired telomerase activity and processivity. Though still able to form the characteristic ladder-like products, ΔU177 and DC

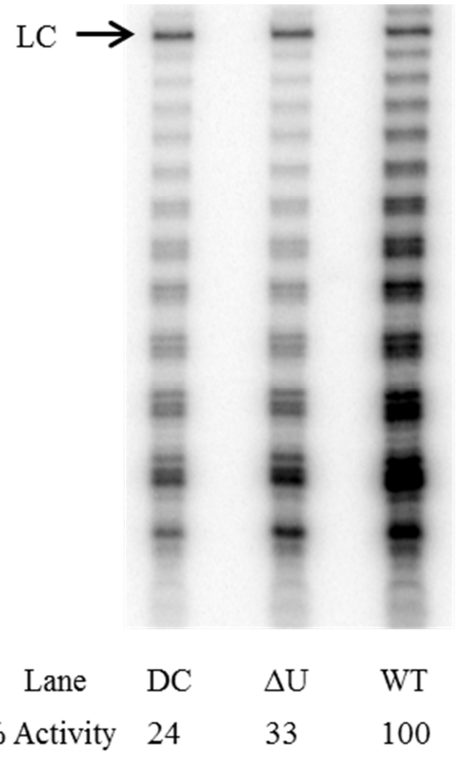


Figure 3.5 Telomerase activity assay of super telomerase extracts with hTRs that incorporate pseudoknot mutants. A direct telomerase assay was conducted with super telomerase extracts incorporating hTRs with mutations to the pseudoknot region similar to the pseudoknots constructs examined by SHAPE above. The loading control (LC) at the top of each assay was used to normalize telomerase activity for each experiment.

telomerases were 33% and 24% of wildtype activity respectively, with no real inhibition of telomerase processive addition noticed (Figure 3.5). This demonstrates that changes made to hTR in the pseudoknot domain not only change the structure of the RNA, but also play a significant role in enzyme activity and possibly even complex formation.

5. Wildtype CR4-CR5 domain

Visual inspection of a secondary structural representation of the wildtype CR4-CR5 generated from the incorporation of SHAPE data into the RNAstructure program confirms that this RNA folds differently from the structure predicted by phylogenetic conservation (31). Nucleotides 242-247 have very low SHAPE reactivities which corresponds well with their base pairing with nucleotides 322-326 to form the P5 helix (Figure 3.6B). Surprisingly, nucleotides 322-326 are very flexible suggesting that they should not be based paired at all but instead forming a single stranded loop region. According to the phylogenetic structure, above helix P5 would be the large J5 loop which also incorporates the P6.1 stem off the right hand side. Instead SHAPE reveals that this loop has completely restructured, now forming a five nucleotide loop from 248-252 and an addition stem with nucleotides 253-255 base-pairing with similarly low-reactive 319-321 (Figure 3.6B). The left hand single stranded region from the J5 loop has completely disappeared with most of these nucleotides now participating in helix formation from base pairings due to their low reactivities, while the right hand single strand region has gone from having seven to four nucleotides, G315-A318; the reactivities of these nucleotide are all really high supporting their being in a single stranded conformation. Nucleotides 302-314 still appear to form the essential P6.1 helix even though all the nucleotides display high SHAPE reactivities (Figure 3.6B).

The P6a and P6b stacked helices that make up the distal end of the CR4-CR5 stem come together similarly to the phylogenetic structure with a few discrepancies (Figure 3.6B). It is suggested that U261 is bulged and C262•A295 form a wobble base pair. Our data, however, dispute that claim and demonstrate that C262, a nucleotide with

higher flexibility, bulges while U261•A295 form a Watson and Crick base pairing, a conclusion that has also been drawn from other studies (35,47). The J6 loop forms with nucleotides 289-291 of the right side displaying moderate flexibility while nucleotides C266 and C267 on the left side remain mostly unreactive suggesting their involvement in some sort of stacking interaction or water-mediated triple base pairing predicted phylogenetically and by NMR (Figure 3.6B) (47). Nucleotide G270 in P6b exhibits moderate SHAPE reactivity leading us to believe that the highly flexible bulged A285 may induce a kink or turn near the top of the helix making it more accessible to SHAPE modification. Though supposedly single stranded and making the loop at the end of the CR4-CR5 domain, U276-G279 displays minimal flexibility (Figure 3.6B) (120).

6. G305A and G322A CR4-CR5 domains

Comparison of the wildtype CR4-CR5 domain with the mutants that give rise to aplastic anemia shows both similiarities and discrepancies in secondary structure. Regions near the top of the domain, specifically the P6a, J6, and P6b structures, form just as they do in the wildtype CR4-CR5 domain and in the proposed hTR structure (31). In the mutants, however, the flexibility noticed with nucleotide G270 no longer appears (Figure 3.6 B,C,D). The G305A nucleotide mutation completely destabilizes and prevents the formation of the P6.1 loop and places nucleotides G309-G319, which were originally part of the P6.1 helix and J6.1/5, into base pairings with C248-C256 forming an extending P5 helix with bulged nucleotides at U312 and G315 (Figure 3.6C). The loop region in the wildtype CR4-CR5 created from C248-A252 no longer exists and is now incorporated in based pairings making the extended P5 helix (Figure 3.6 B,C).

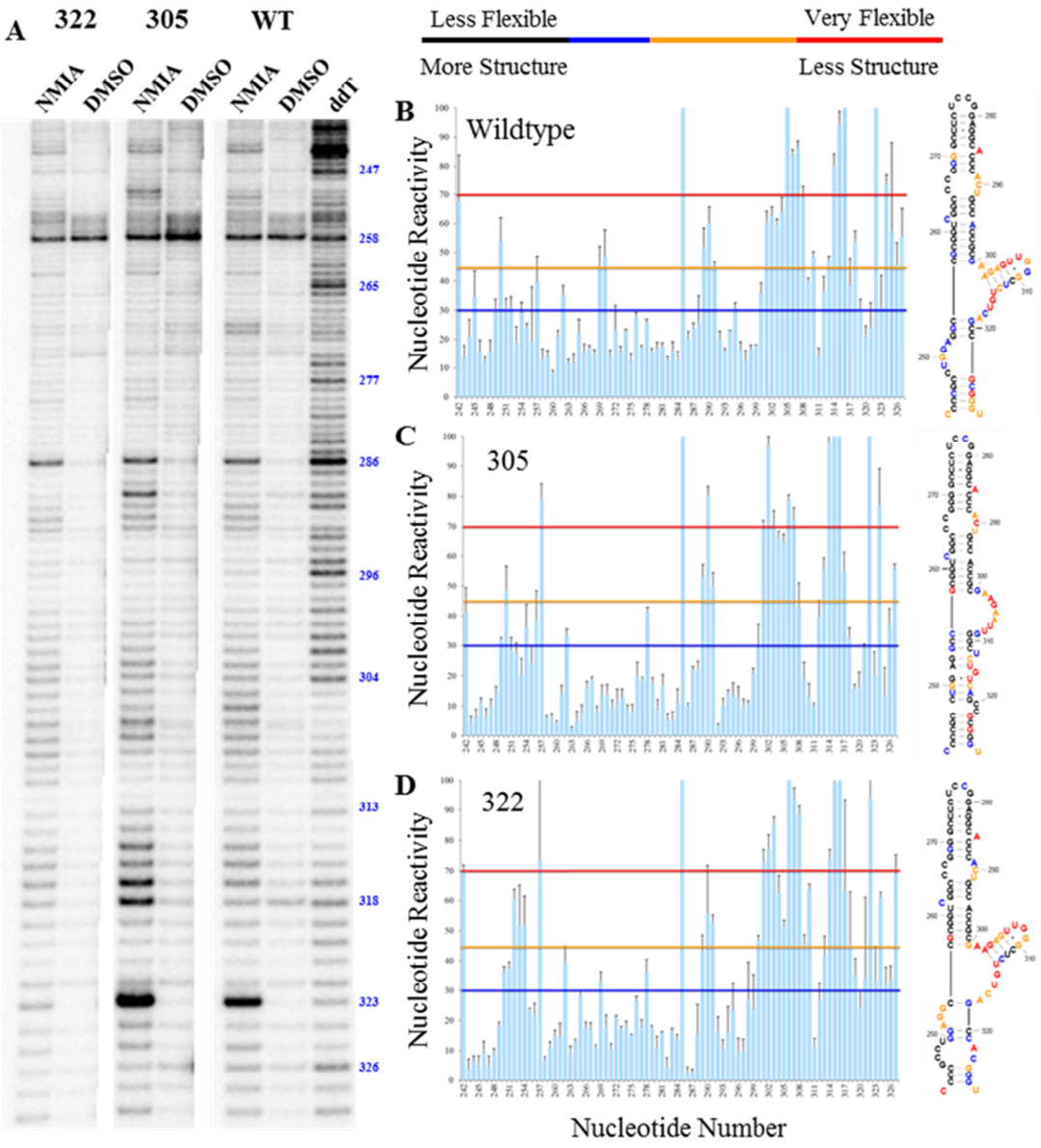


Figure 3.6 SHAPE analysis of wildtype and mutant CR4-CR5 domains. **(A)** Flanked CR4-CR5 domain RNAs were treated with 5 mM NMIA or 10% DMSO (background) and the acylated RNAs were mapped by reverse transcription. The ^{32}P-radiolabeled cDNAs were resolved on a denaturing sequencing gel. Nucleotides "hit" during SHAPE cause reverse transcription to stop one nucleotide before the modification site. A dideoxythymidine (ddT) ladder was also run and indicates the location of adenines in the RNA sequence. The gel is numbered according CR4-CR4 domain nucleotide number, not the ddT ladder which is off by +1 nucleotide. **(B,C,D)** Histograms and color-coded secondary structures of the wildtype, 305, and 322 CR4-CR5 domains. The average band intensities were normalized and plotted according to nucleotide position with error bars (N=4). Black, blue, orange, and red lines indicate divisions for examining nucleotide reactivity values where $0 \leq$ black < 30, $30 \leq$ blue < 45, $45 \leq$ orange < 70, and red ≥ 70 and lower numbers represent less reactivity and higher numbers represent greater reactivity. The color-coded secondary structures reflect the histogram data for each RNA examined. Proposed RNA secondary structures were determined by incorporating SHAPE data into RNAstructure (121). **(B)** The wildtype CR4-CR5 domain. **(C)** The 305 CR4-CR5 domain. **(D)** The 322 CR4-CR5 domain.

Another smaller loop region is formed from C319 and C320, while the original P5 helix from phylogenetic and mutational analysis forms with all of its nucleotides having lesser SHAPE reactivities than in the wildtype, though G322 and G324 are still quite reactive.

Structure of the CR4-CR5 for the G322A mutation is similar to the wildtype CR4-CR5 domain from nucleotides G252-C321 and at the very base of the stem, though the paired nucleotides at the base in this mutant have much lower flexibility as would be expected in a base pairing situation (Figure 3.6 B,D). Directly below this, two new loop regions are formed in the hypervariable region of hTR, a portion of the RNA that is not very well-characterized (48). The first loop includes single stranded nucleotides A252-G254, all of which are highly reactive supporting their single stranded nature (Figure 3.6D). Next, there is a two base pair stem composed of G250-C321 and G251-C320, all nucleotides that were thought to be part of the J5/6 internal loop in the phylogenetically derived structure (31). Under this stem is the second internal loop formed; the right side of the loop includes the mutated A322 and C323, both of which display reactivity, though the mutant nucleotide is significantly more reactive (Figure 3.6D). The sequence of the left side of the internal loop is UCCG, a known stable loop motif thus explaining the lack of SHAPE reactivity seen for these nucleotides (120).

7. Wildtype CR7 domain

Comparatively speaking, the CR7 domain construct looks more structured and clean than the phylogenetically predicted CR7 region and matches well with previous studies that interrogated hTR structure by chemical modification and enzymatic cleavage (31,35,48). The P7a helix is shortened from four base pairs to three due to the extreme

level of SHAPE reactivity denoted for G440, suggesting it is actually a bulged nucleotide, and G441 actually forms the Watson and Crick base pair with C383. The J7 loop region formed between the P7a and P7b completely disappears as nucleotides originally thought the be single stranded show low reactivities and are more likely to form Watson and Crick base pairings G385-C438 and A386-U437 (Figure 3.7B). The 3' pocket maintains the same shape as the proposed structure though the single stranded 3'

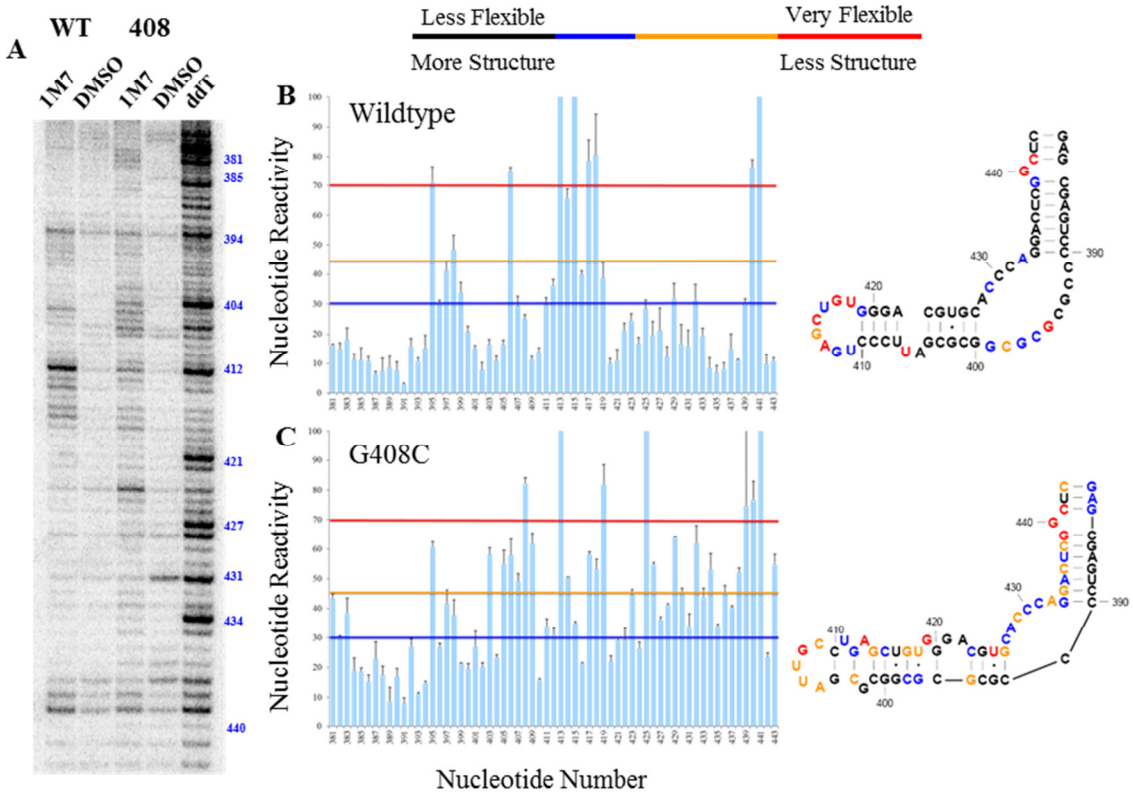


Figure 3.7 SHAPE analysis of wildtype and mutant CR7 domains. **(A)** Flanked CR7 domain RNAs were treated with 5 mM NMIA or 10% DMSO (background) and the acylated RNAs were mapped by reverse transcription. The ^{32}P-radiolabeled cDNAs were resolved on a denaturing sequencing gel. Nucleotides "hit" during SHAPE cause reverse transcription to stop one nucleotide before the modification site. A dideoxythymidine (ddT) ladder was also run and indicates the location of adenines in the RNA sequence. The gel is numbered according to CR7 domain nucleotide number, not the ddT ladder which is off by +1 nucleotide. **(B,C)** Histograms and color-coded secondary structures of the wildtype and 408 CR7 domains. The average band intensities were normalized and plotted according to nucleotide position with error bars (wildtype N=4, 408 N=3). Black, blue, orange, and red lines indicate divisions for examining nucleotide reactivity values where 0 ≤ black < 30, 30 ≤ blue < 45, 45 ≤ orange < 70, and red ≥ 70 and lower numbers represent less reactivity and higher numbers represent greater reactivity. The color-coded secondary structures reflect the histogram data for each RNA examined. Proposed RNA secondary structures were determined by incorporating SHAPE data into RNAStructure (121). **(B)** The wildtype CR7domain. **(C)** The 408 CR7 domain.

portion now encompasses nucleotides A428-A432, two more than predicted (Figure 3.7B). Also, nucleotides within the internal loop are much less flexible than would be expected. This suggests that either nucleotides form base pairs across the loop between two single stranded regions or form stable secondary or tertiary interactions that prevent SHAPE modification. G's 393, 397, and 399 could potentially pair with C's 427, 430, and 431 as all of them have low SHAPE reactivities; a G399-C430 base pair has been predicted in a previous study (Figure 3.7B) (35). The P8 helix still forms two smaller stacked helices, but instead of being separated by the J8 internal loop, they are now separated by a two-nucleotide bulge consisting of A404 and U405 with A 404 showing minimal flexibility and U405 being highly reactive. The end of the CR7 domain forms an eight-nucleotide loop as is expected due to the high conservation of this loop in telomerase RNA across vertebrates (48,52).

8. G408C CR7 domain

Examination of the G408C CR7 mutant reveals that while the P7 helices are very similar to the wildtype CR7 domain, the P8 helices and end loop (L8) have very different RNA structural features (Figure 3.7B,C). The P7 helix forms just like in the wildtype, though the 3' strand of the helix exhibits a low to moderate level of flexibility in each and every nucleotide between G433 and C443 (except G439-C441 which display high flexibility and U442 seems unreactive). This result allows us to propose that perhaps the P7 helix is only forming in a percentage of the folded RNAs and that the rest may have this string of nucleotides present as a single stranded region. Of the nucleotides that made up the 5' side of the 3' pocket only one remains (and even C391 has low

reactivity), with the rest of the nucleotides being pulled into base pairing interactions in the newly formed and extended P8 helix (Figure 3.7C). The newly formed P8 helix is now made up of two different stacked helices (the first stem has 4 base pairs with one G•U wobble and the second stem has 5 base pairs with 2 G•U wobbles) with multiple single and double nucleotide bulges scattered amongst them. Crowning the latter five base pair stem are two isolated base pairs which are further divided by a two-nucleotide internal loop (G402 and A413) and another bulged nucleotide in U411. Nucleotides U411-U418 which were single stranded in the wildtype CR7 and are required to be single stranded for the CR7 domain to bind with telomerase accessory proteins implicit for the proper localization and accumulation in the cell are now involved in base pairings that make the extended yet disjointed P8 helix (Figure 3.7 B,C) (52,58). The complete rearrangement of this structurally conserved helix illustrates how a single substitution could manifest in a premature aging disease such as DC.

C. Discussion

Studies have shown that telomerase RNAs fold similarly despite major differences in their sizes; this indicates that structural arrangements, not necessarily nucleotide sequence, are implicit for proper functionality (48). Regions within the hTR, including the pseudoknot domain, CR4-CR5 domain, and CR7 domain have been shown through phylogenetic analysis, co-variation studies, and sequence alignment to be conserved structures through vertebrates indicating their importance within telomerase for complex formation and biological function (Figure 3.1) (48). Mutational studies that incorporate individual and compensatory mutants further emphasize the importance of

correct domain folding during the development of functional telomerase (57,118,122). Unfortunately, current structural information for these important domains is limited, and the data that is available is either inferred from other RNAs and applied to hTR or analysis of individual loops or helices as larger portions of the RNA are not amenable to techniques utilized for RNA structure determination due to size constraints (47,57,58). For these reasons, SHAPE data that corresponds to the entire hTR domains, and generated in these studies, is a significant contribution to the structural understanding of hTR.

Current understanding in the field identifies the triple helix of the pseudoknot domain as a requisite for telomerase's main function of processive nucleotide addition assigning critical importance to the formation of this tertiary RNA interaction (46,123). However, to this date, the pseudoknot triple helix structure or the P3 helix (an important helix for pseudoknot formation) has not been identified in solution suggesting that it may only form when in complex with hTERT (will be addressed more in Chapter 4) (35). Across vertebrates, the loops and stems (p2b, P3, J2b/3, and the 3' end of J2a/3; nucleotides 90-125 and 170-184) actually involved in creating the pseudoknot reveal high sequence conservation, while the remainder of the domain is only structurally conserved (40,41,48). For this reason, a construct incorporating the conserved elements has been the focus of research trying to understand the role of the pseudoknot triple helix structure (Figure 3.2B) (40,41).

Initial thermodynamic experiments conducted on the conserved pseudoknot identified the RNA to be in equilibrium with an elongated hairpin conformation (42). Removal of nucleotide U177 pushed the equilibrium back towards pseudoknot formation

as determined by isothermal melting studies where the stability of the pseudoknot is higher than that of the hairpin due to tertiary interactions between helical and loop regions (42). Pseudoknot tertiary RNA interactions come together from the formation of two separate helices (p2b and P3) with the single stranded loop regions (J2b/3 and Ja2/3) being in close enough proximity to these stems to form base triples that can stabilize the structure and yield an RNA triple helix. The low flexibilities of nucleotides 93-98/116-121 and 107-115/174-183 (excluding U177) from SHAPE experiments with the wildtype and ΔU177 RNAs and the RNA Structure prediction based on the SHAPE data correlate well with the formation of the two stems (Figure 3.4 B,C) (121).

NMR studies have shown that the first nucleotide in J2b/3 U99 and the last nucleotide in J2a/3 A173 form a Hoogstein base pair between the two stems that prevents coaxial stacking and allows the major groove of P3 and the minor groove of P2b to form a continuous groove (28,40,41). SHAPE data supports the formation of this base pair due to low flexibilities from each of these nucleotides (Figure 3.4 B,C). NMR also demonstrates the making of canonical and non-canonical triple bases from the U-rich J2b/3 laying in the major groove of P3 and the A-rich J2a/3 residing in the minor groove of P2b with both of the loops facing the helix (32,40). Nucleotides U100-U102 make Hoogstein triples with U115-A174, U114-A175, and U113-A176 while non-traditional triples form between A172 and A171 with Watson Crick base pairs C116-G98 and A117-U97 respectively (32,40). These observations are also reflected in the low SHAPE reactivity of U100-U102 and A171-A172 in the wildtype and ΔU177 mutant and the structures generated (Figure 3.4 B,C). Mutational analysis of nucleotides involved in these triples drastically decreases telomerase activity and pseudoknot stability further

supporting the formation and importance of these triples (36,37,41,44,122). Additionally, SHAPE shows that C170 is also substantially less flexible than other loop nucleotides suggesting its involvement in a third possible triple with G118-C96. NMR, however, attributes this to stacking interactions over A171 (41).

Studies place the U177 below the Hoogstein UAU triples and flipped out of the helix though still situated in the minor groove of P3 away from any interactions with the triple helix structure (40). The presence of U177 creates a large roll and tilt between base pairs U113-A176 and C112-G178 and prevents proper base stacking and subsequently makes more space in the major groove of P3 for the J2a/3 loop (40). These conclusions explain the high SHAPE flexibility of U177 observed in the wildtype pseudoknot plus the increased flexibility seen in nearby nucleotides A176 and A178 (Figure 3.4B). Moreover, chemical analysis of the individual nucleotides surrounding U177 imply that the 2'-hydroxyl group of A176 contributes directly to telomerase catalysis (37). Taken together, the data proposes that U177 may act as a hinge to provide additional backbone stability for A176 or may play a role in modulating activity of the catalytic active site during nucleotide addition (37,40). Enzymatic and chemical probing of hTR in complex with hTERT identifies U177 as being protected, reinforcing its importance in catalysis and also suggesting that U177 may partake in hTERT binding (35). This may not be the case though, as other work has shown that neither disruption of helix P3 nor the triple helix abolishes hTERT binding (36,37). This insinuates that another portion of the core domain contains the hTERT binding site or that hTERT binds to a specific structure and not specific nucleotides (28,113,124).

Deletion of nucleotide U177 in the mutant pseudoknot drives down the SHAPE reactivity of U103 from moderate to low indicating its involvement in RNA structure (Figure 3,4B). This is confirmed by NMR studies that demonstrate that the removal of U177 allows the formation of an additional triple between U103 and C112-G178 as the J2b/3 loop now interacts more tightly with the major groove nucleotides of P3 (40). The deletion of U177 also leads to an increase in bending at the helical junction (40). This changes the structure and stacking interactions of the triple helix and pseudoknot interactions with hTERT purportedly contributing to the significant decrease in telomerase activity observed with the ΔU177 mutant (Figure 3.5) (40).

Unlike the wildtype and ΔU177 constructs which form pseudoknot structures, comparison of the SHAPE profile of the DC construct attests to disruption of the vital pseudoknot structure and supports previous studies implying that the GC 107-108 AG mutation shifts the equilibrium of the RNA toward a tightly zipped hairpin conformation (42). Melting analysis of the wildtype pseudoknot demonstrates hierarchical unfolding with tertiary and then secondary structures melting sequentially (125). When the same experiments were completed with the DC pseudoknot, all melting occurred in one step (125). SHAPE data correlates well with these experiments as it points to this RNA actually making one elongated stem loop with 12 base pairs in the expanded P2b stem and a sequence-stabilized YNMG hairpin pentaloop (Figure 3.4C) (42). Interestingly, the three U•U base pairs followed by the C•U base pairs allow the extended P2b to form an A-form helix with the U•U base pairs affording additional stability to the hairpin (42). Molecular modeling total energy simulations show that stem P2b and loop J2b/3 demonstrate high levels of rigidity compared with the J2a/3 loop and P3 which are

flexible and unstable (125). Base stacking implies the formation of a weak P3 helix, though the proposed P3 and J2a/3 regions undergo periodic reshuffling and fluctuations so definite structures cannot be decisively verified (125). SHAPE data reflects these observations as nucleotides A171–C183 all have increased SHAPE reactivities in comparison to the wildtype and ΔU177 pseudoknots supporting a structure where these nucleotides are single stranded (Figure 3.4). The complete unraveling of the triple helix and the subsequent shift in structural equilibrium from the pseudoknot to stabilized hairpin conformation greatly diminished telomerase activity in cell extracts as tertiary RNA interactions have been shown to be required for proper catalytic activity (Figure 3.5) (36,63,118,122).

The CR4-CR5 domain, also referred to as the 5' stem loop or the trans-activating domain, is essential in the telomerase complex for processive nucleotide addition and contains an apparent hTERT binding site. It is widely believed that binding of hTERT to hTR elicits a conformational change that positions the p6.1 helix in close proximity to the pseudoknot domain (46,49). There is high sequence and structural conservation within this domain, however, it is also one of the least characterized regions of hTR (48). Phylogenetic and mutational comparison indicates the P5 is a five base pair stacked helix composed of G-Cs that sits right below the large J5/6 loop region (48). SHAPE data disputes this demonstrating that P5 is actually a longer helix with a mostly single stranded internal loop added on the left hand side of the helix (Figure 3.6B). Footprinting experiments of full length hTR support the formation of the internal loop as U249-A252 were moderately to strongly modified by chemical reagents and cleaved by S1 nuclease denoting their single stranded nature (35). Nucleotides G353-C355, which were once

part of J5/6a, exhibit low SHAPE reactivities and form base pairs with G319-C321 which were single stranded in J6.1/5 adding three base pairs to the length of P5 (Figure 3.6B). As it has not been well characterized, a definite explanation for the extended nature of this helix and the internal loop is speculative, though it is possible that the structure acts as an hTERT/accessory protein binding site or allows the helix to bend for proper positioning its distal end near the active site.

Based on the phylogenetic structure of hTR, mutational studies analyzing the J5/6a and J6.1/5 internal loops have determined that no individual mutation of a single internal loop residue compromises telomerase activity or hTERT binding to the same extent as deletion of the entire domain (118). This implies that the single stranded nucleotides within the internal loop provide a contact service for hTERT or may indirectly promote hTERT binding and that binding is not dependent on specific nucleotides (118). While these conclusions regarding the nucleotides within this region may be true, neither footprinting data nor SHAPE analysis support the formation of this loop region as predicted in the phylogenetic structure instead both incorporate portions of J5/6 into helix P5 (as mentioned above) (35). The P6.1 stem does form as predicted in the phylogenetic structure of hTR, even though SHAPE data presents almost all of the nucleotides between A301-G315 as being moderately to highly flexible (Figure 3.6B). Based on this data, it would be hard not to argue that P6.1 may not actually form in solution and only makes a helix upon hTERT binding which has been shown to protect the loop region of P6.1 from chemical modification and enzymatic cleavage (35,49).

The P6a and P6b helices that make up the distal end of the CR4-CR5 stem come together similarly to the phylogenetic structure with few discrepancies. NMR data shows

that these two stacked helices adopt a canonical A-form helix (47). It has been suggested that U261 is bulged and C262•A295 form a wobble base pair (48). SHAPE data, however, disputes that claim and demonstrates that C262, a nucleotide with higher flexibility, bulges while U261-A295 form a Watson and Crick base pairing, a conclusion that has also been drawn from other studies (Figure 3.6B) (35,47). Truncation mutants that eliminate the J6, which is common to all mammalian telomerases, abolish telomerase activity suggesting this internal loop as the location of an hTERT binding site via the formation of a solvent accessible tunnel (46,47). Chemical probing supports this conclusion as C290-U291 become protected from DMS and CMCT modification when hTR is bound to hTERT (35). SHAPE data shows that nucleotides C266 and C267 have very low flexibilities but still appear to be single stranded while A289-U291 appear moderately flexible (Figure 3.6B). This correlates well with NMR data and phylogenetic data which shows that C266 stacks upon helix P6a and that C267 lies coplanar with G268-C288 and appears to form a water-mediated triple base pair therefore protecting itself from SHAPE modification (47). It has also been suggested that C266 may intermittently form a C•U base with U291 even though the distance between the two nucleotides is greater than what in normally seen for this type of interaction and U291 is moderately reactive to SHAPE (47). SHAPE designates the L6 loop, U276-G279, as having very minimal flexibility despite is single stranded nature (Figure 3.6B). This agrees with the formation of the UCCG tetraloop that has its bases oriented in the anti-conformation with the exception of the 4^{th} nucleotide that is syn making it highly stable (120).

The G305A and G322A mutations that occur within the CR4-CR5 domain and give rise to aplastic anemia elicit changes in CR4-CR5 domain structural organization though in different locations. For each of the mutants, the structure of the CR4-CR5 domain was similar to the wildtype from nucleotides G257-A301. Interestingly though, in both mutants nucleotide G257, which was only slightly flexible in response to SHAPE in the wildtype RNA, are extremely flexible even though they appear to be based paired at the base of the P6a stem (Figure 3.6 B,C,D). Neither phylogenetic, nor mutational, nor functional analyses give any reason to believe changes to this nucleotide in the P6a helix could affect telomerase activity. Perhaps the additional flexibility of this nucleotide could prevent key interactions with hTERT that could lead to the diminished telomerase activity exhibited in these mutants or induce a kink or turn of the helix that inhibits proper helical positioning (62).

SHAPE data demonstrates that G305A mutant destabilizes and prevents the formation of the required P6.1 helix turning all of the nucleotides into an elongated J6a/5 internal loop structure (Figure 3.6C). Interestingly, recent studies have shown that the incorporation of the compensatory mutant C311U into this mutant structure can completely rescue hTERT binding and telomerase activity in this mutant (118). This once again demonstrates how RNA structure, not necessarily nucleotide sequence, governs the RNA role in the telomerase biological function. Other changes in structure compared to the wildtype for this mutant include the loss of the newly formed J5 single stranded internal loop region, as its nucleotides are once again involved in the formation of an even longer extended P5 helix that is now 14 base pairs long and includes multiple bulged and looped nucleotides throughout its helical formation (Figure 3.6 B,C). The

elongated P5 has many nucleotides with moderate to high levels of flexibility suggesting instability in the formation of this region of the trans-activating domain, an area whose purpose is not completely understood. In the case of the G322A mutant, the structure of the P5 helix is perturbed. The 8 base pair helix and J5 loop region from the wildtype domain appear to form three smaller stacked base paired regions separated by large loop regions with nucleotides with high flexibility on both sides of the helix (Figure 3.6 B,D). Nucleotides G246-U249 form an internal loop with a sequence UCCG, a known highly stable tetraloop structure which also forms at the distal end of the CR4-CR5 domain and demonstrates low SHAPE flexibilities due to the conformations of the bases in relation to their sugar puckers (120). Perhaps disruption of the tightly zipped P5 helix prevents association of hTR with hTERT or other accessory proteins required for telomerase activity.

The CR7 domain of hTR is dispensable *in vitro* but required in the cell for proper hTR maturation, localization and processing (52,57). The structure and sequence of this domain resembles that of small Cajal Body specific RNAs (scaRNAs) though it is still unclear whether this domain in hTR actually directs RNA modification in itself or other RNAs (53). It has been shown that hTR associates with the dyskerin complex, the companion complex of scaRNAs, and undergoes processing and maturation in the same cellular location as scaRNAs (57). Comparatively speaking, the CR7 domain shows the most similarity between its SHAPE secondary structure and the phylogenetically and mutationally proposed structure of hTR (Figures 3.1 and 3.7) (48). The major difference being that CR7 domain analyzed by SHAPE appears more structured with fewer and smaller looped regions.

SHAPE shows that P7 forms as a ten base pair helix with a single bulged nucleotide at G440 with very high flexibility; the previously predicted J7 internal loop does not form and instead forms base pairs between the J7a/7b and J7b/7a nucleotides (Figure 3.7B). This data is confirmed for nucleotides G381-C390 by enzymatic cleavage of this helix by RNase V1, an enzyme that cleaves at non-sequence specific base paired nucleotides (35). The bulged G440 increases the flexibility of nearby nucleotides G439 and C441; this could be due to G440 invoking a kink or turn in the helix or preventing normal base stacking interactions due to its flexible nature. Similar to the phylogenetic structure, an asymmetrical internal loop region forms between the P7 and P8 helix (48). J7/8 maintains the same size as predicted phylogenetically while J8/7 is five nucleotides long and still within the range for known comparative species (48). Despite its single stranded appearance, SHAPE data only identifies nucleotides G395-G399 in J7/8 and C429 and A432 as exhibiting flexibility suggesting that there may be some secondary RNA interactions occurring (Figure 3.7B). RNAstructure constrained with SHAPE data predicts the formation of a stem loop akin to the P6.1 from J7/8 with a three base pair stem (C392-C394 base paired with G399-G397) and a two nucleotide loop structure (G395 and C396) (121).

The P8 helices of hTR boast high sequence and structural conservation though until recently, most of the information known about its structure has been extrapolated from scaRNAs (42,48,54). Much like the P7 helix, the P8a and P8b stacked helices appear to form one longer nine base pair helix that is separated in the middle by a two-nucleotide loop of A405 and U406 (Figure 3.7B). SHAPE data does not support the formation of the phylogenetically proposed C401•A428 base pairing, instead having

C401 in a Watson Crick base pair with G426 and showing A428 as a single stranded nucleotide in J7/8 despite its negligible SHAPE reactivity (35,48). L8 forms at the end of P8 as a single stranded eight-nucleotide loop region as predicted with most of the nucleotides displaying high levels of SHAPE reactivity (Figure 3.7B). NMR experiments demonstrate that a U411•G417 wobble pair forms within the loop structure which creates a kink turn motif required for protein binding, though this interaction cannot be substantiated from SHAPE data alone (57,58). The formation of this loop is essential for CR7 domain function as both the CB localization signal and the hTR-specific processing signal require single stranded nature from different nucleotides within the loop (57). The CB localization signal, or CAB box (UGAG), requires U411-G414 be highly flexible and helix formation of P8b with sequence dependence on the last two base pairs of the helix (57). The hTR processing signal requires the U411•G417 wobble pair, a flexible U418, and sequence-dependent base pairs at C410-G419 and C409-G420 at the end of the P8b helix (57). Deletion of the 3' loop or replacement of it with a tetraloop cap reduced hTR accumulation and telomerase activity in the cell (52,126). Mutation of U418 stalled hTR processing after transcription and flipping of the 3' base pairs of P8b inhibited 3' end processing and resulted in mixed cellular localization (57). The SHAPE data in addition to conclusions from previous work highlight the importance of CR7 structure for the generation and maturation of hTR in the cell.

Like the GC 107-108 AG mutant in the pseudoknot domain, the C408G mutant in the CR7 domain leads to the development of the premature aging-related disease dyskeratosis congenita. As with the other mutations that we have examined, this single point mutation leads to extensive secondary structural rearrangement in the CR7 which

contributes to DC disease manifestation and progression. Beginning from the end, nucleotides U425-C443 all exhibited increased SHAPE reactivity when compared to the wildtype CR7 domain (Figure 3.7C). Though the proposed structure of this mutant domain does suggest the formation of helix P7, it is also quite likely that this helix may assume multiple conformations or undergo fluctuations or shuffling similar to what was suggested for the P3 helix of the DC pseudoknot. J7/8 has shrunk to just one nucleotide which despite its single stranded nature has minimal SHAPE reactivity meaning that this nucleotide is probably interacting in some RNA interaction while the J8/7 loop on the other hand has grown to six nucleotides all displaying low to moderate flexibility (Figure 3.7C). The main structural changes occur in the P8 helix and L8 loop. The UGAG CAB box sequence that was single stranded in the wildtype CR7 is now partially based paired with nucleotides C401 and C403. U411 is a bulge at the end of the helix while G417 is base paired with C398, meaning that there is clearly no wobble pair being formed. Additionally, the sequence dependent 3' base pairs (C410-G419 and C409-G420) have also been disrupted and are participating in other structural interactions (Figure 3.7C). The disruption of the CB localization and hTR processing signals definitely contributes to the decreased telomerase activity and improper localization and maturation of hTR (52,62). Like with the G305A AA mutation in the CR4-CR5 domain, incorporation of the compensatory mutation G421C across from C408G overcomes the structural disorganization once again underlining the significance of RNA secondary and tertiary structure on hTR biological function (56).

D. Materials and Methods

1. Oligonucleotides and chemicals

Oligonucleotides were purchased from Integrated DNA Technologies (IDT, Coralville, Iowa). The oligonucleotides were then purified via denaturing polyacrylamide gel electrophoresis (PAGE) using the modified "crush and soak" method (112,113). Following electrophoresis in a 12% denaturing gel, oligonucleotides were observed via ultraviolet (UV) light, cut out with a razor blade, crushed through a 3 mL sterile plastic syringe, and eluted from the gel in 1×TEN buffer (10 mM Tris pH 7.5, 1 mM EDTA, 250 mM NaCl) for one hour at room temperature. After centrifugation, the supernatants were collected and passed through a 0.22 micron MCE membrane syringe filter (FisherBrand) to remove remaining gel particulate. Oligonucleotides were then ethanol precipitated with final concentration 0.6 M ammonium chloride, concentrated, and resuspended in TE (10 mM Tris-Cl pH 8.0, 1 mM EDTA) or DEPC-treated water (Ambion, Inc., Austin, Texas) to a stock concentration of 1 mM. The concentration of further dilutions was determined by UV absorbance at 260 nm using the molar extinction coefficient supplied by the manufacturer.

N-methyl-isatoic anhydride (NMIA) was purchased from Invitrogen Molecular Probes (Eugene, Oregon) and 1-methyl-7-nitroisatoic anhydride (1M7) was a gift from Dr. Kevin Weeks (UNC Chapel Hill, NC, USA). These compounds were made fresh at 50 mM stocks in DMSO ACS reagent grade (MP Biomedicals, Solon, Ohio).

2. hTR and hTERT expression plasmids

The puc19 phTR+HH expression plasmid was a gift from Dr. Jamie Sperger (University of Colorado, Boulder). This plasmid was designed for *in vitro* transcription

of human telomerase RNA and contains a self-cleaving hammerhead ribozyme that generates the natural 5' end inserted in front of the hTR gene. The pVan107 hTERT and the pBS-U1 hTR super telomerase expression plasmids were a gift from Dr. Joachim Lingner (Ecole Polytechnique Federale de Lausanne, Switzerland). These plasmids were designed for in-cell over-expression of human telomerase (127). Site directed mutagenesis, based on the Stratagene Quick Change site-directed mutagenesis protocol, was used to generate puc19 phTR+HH and pBS-U1 hTR constructs coding for pseudoknot and CR4-CR5 domain dyskeratosis congenita and aplastic anemia mutants.

3. Cells and cell culture

The HEK293T (ATCC, Manassas, VA) human embryonic kidney cell line was cultured in Dulbecco's Modified Eagle's Medium (DMEM) (Gibco, Carlsbad, CA) and supplemented with 10% fetal bovine serum (FBS) (Sigma, St. Louis, MO) and 5 mg/mL antibiotic-antimycotic (Gibco).

4. PCR construction of dsDNA pseudoknot constructs

In order to make the pseudoknot construct RNAs, dsDNA templates were generated by PCR using overlapping primers. To minimize the number of primers needed, the primers were designed so that the wildtype (WT) forward and reverse primer could also be used for generating the two mutants. The forward primers contain a T7 promoter and 5' linker before the pseudoknot sequence while the reverse primers have the 3' linker and conRT binding site incorporated after the pseudoknot sequence (85). Complementary primers underwent the following PCR cycles: 1. 95 °C for 10 minutes

Table 3.1: Primers for Making Pseudoknot Constructs

Pseudoknot Primer	Sequence (5'-3')
WT Forward	GCGCTAATACGACTCACTATAGGGGGCCTTCGGGCCA AGGGCTGTTTTTCTCGCTGACTTTCAGCCCCAAAC
WT Reverse	GAACCGGACCGAAGCCCGATTTGGATCCGGCGAACCG GATCGATGCTGACATTTTTTGTTTGGGGCTGAAAGTC
ΔU Reverse	GAACCGGACCGAAGCCCGATTTGGATCCGGCGAACCG GATCGATGCTGACTTTTTTGTTTGGGGCTGAAAGTC
DC Forward	GCGCTAATACGACTCACTATAGGGGGCCTTCGGGCCA AGGGCTGTTTTTCTCAGTGACTTTCAGCCCCAAAC

enzyme hot-start; 2. five cycles of 94 °C for 30 seconds, 55 °C for 30 seconds, 72 °C for 1 minute to yield a full length double stranded product; and 3. two cycles of 94 °C for 1 minute, 55 °C for 30 seconds, 72 °C for 1 minute to complete any unfinished ends. Upon completion, PCR products were PCI extracted, ethanol precipitated, concentrated, resuspended in TE, and the concentration was determined by UV spectroscopy.

5. PCR construction of dsDNA CR4-CR5 domain constructs

The CR4-CR5 domain is much larger than the pseudoknot domain, so primers that included the necessary flanking regions were used to clone these domains directly out of the puc19 phTR+HH expression plasmid (85). Two mutant puc19 phTR+HH plasmids were made via site-directed mutagenesis and incorporated the mutations for the aplastic anemia mutants G305A and G322A. The CR4-CR5 forward and reverse primers were used for the making the wildtype and G305A CR4-CR5 domain templates. Due to the location of the G322A mutation, a different reverse primer was used for making this

Table 3.2: Primers for Cloning CR4-CR5 Domains

CR4-CR5 Domain Primer	Sequence (5'-3')
CR4-CR5 Forward	GCGCTAATACGACTCACTATAGGGGGCCTT CGGGCCAACCCGCCTGGAGGCCGC
CR4-CR5 Reverse	GAACCGGACCGAAGCCCGATTTGGATCCGG CGAACCGGATCGAGACCCGCGGCTGACAG
322 CR4-CR5 Reverse	GAACCGGACCGAAGCCCGATTTGGATCCGG CGAACCGGATCGAGACCCGTGGCTGACAG

template. The PCR conditions used for cloning out the CR4-CR5 domains were: 1. 98 °C for 30 seconds enzyme hot-start; 2. 30 cycles of 98 °C for 30 seconds, 55 °C for 30 seconds, 72 °C for 1 minute to yield a full length double stranded product; and 3. one 72 °C for 5 minute to complete any unfinished ends. Work up of the PCR products occurred in the same manner as with the pseudoknot constructs.

6. PCR construction of dsDNA CR7 domain constructs

The CR7 domain is also larger than the pseudoknot domain, so a combination of two different PCR programs was used to generate these dsDNA templates. The first reaction utilized overlapping primers to make dsDNAs of the wildtype and C408G mutant CR7 domains via the same PCR conditions used to make the pseudoknot

Table 3.3: Primers For Making CR7 Domains

CR7 Domain Primer	Sequence (5'-3')
WT CR7 Forward	GAGCGAGTCCCCGCGCGCGGCGCGATTCCCTGAGCTGTGG
WT CR7 Reverse	GAGCCGAGTCCTGGGTGCACGTCCCACAGCTCAGGGAATCGC
C408G CR7 Forward	GAGCGAGTCCCCGCGCGCGGCGCGATTGCCTGAGCTGTGG
C408G CR7 Reverse	GAGCCGAGTCCTGGGTGCACGTCCCACAGCTCAGGCAATCGC
Add CR7 Forward	GCGCTAATACGACTCACTATAGGGGGCCTTCGGGCCAAGAGCGAGTCCCCGCG
Add CR7 Reverse	GAACCGGACCGAAGCCCGATTTGGATCCGGCGAACCGGATCGAGAGCCGAGTCCTGGGTG

templates. These PCR products were further amplified using the Add CR7 forward and Add CR7 reverse primers to add the necessary flanking regions using the same PCR conditions as used to generate the CR4-CR5 domain templates (85). Once again, work up of the PCR products occurred in the same manner as with the pseudoknot constructs.

7. *In vitro* transcription and purification of domain RNAs

Pseudoknot, CR4-CR5, and CR7 domain RNAs were made via *in vitro* transcription with the T7 Ampliscribe Transcription Kit (Epicentre Biotechnologies, Madison, WI). The *in vitro* transcription reaction was carried out overnight at 37 °C with one microgram dsDNA template, 1x Ampliscribe T7 Reaction Buffer, 7.5 mM ATP, 7.5 mM CTP, 7.5 mM GTP, 7.5 mM UTP, 10 mM DTT, ScriptGuard RNase Inhibitor, and 2 μL Ampliscribe T7 Enzyme Solution. The next morning, the reaction incubated with 1 μL RNase-Free DNase 1(1MBU/μL) for 20 minutes at 37 °C and then resuspended in an equal volume SHAPE denaturing loading buffer (80% formamide, 0.5×TBE, 4 mM EDTA, 0.01% bromophenol blue, 0.01% xylene cyanol) and purified in an 8% denaturing polyacrylamide gel. The RNA was then recovered by the modified "crush and soak" method as explained above; the only differences being that the RNAs are ethanol precipitated with final concentration of 0.3 M of sodium acetate and final RNA concentrations were determined by UV spectroscopy.

8. NMIA and 1M7 hit reactions

One picomole of RNA (1 μL of 1 μM RNA) , DEPC-treated water, and 5× hTR Hit Buffer (250 mM Hepes Buffer Solution(Gibco), 250 mM RNase-free KCl (Ambion), 5 mM RNase-free $MgCl_2$ (Ambion)) diluted to 1× were combined in a microcentrifuge tube and allowed to equilibrate at 37 °C for a couple of minutes to allow RNA folding. The RNA was then treated with either 50 mM NMIA or 50 mM 1M7 to a final concentration of 5 mM, or DMSO to a final concentration of 10% as a control and incubated at 37 °C for 5 minutes or 70 seconds respectively for NMIA and 1M7. The hit

reaction was quenched with RNase-free NaCl (Ambion) to a final concentration of 200 mM and precipitated with 200 μg/mL RNase-free glycogen (Ambion) as a counter ion. Following centrifugation, the hit RNA was washed once with 70% ethanol, speed vacuumed until dry, and the resuspended in 10 μL of RNase-free TE buffer pH 8.0 (Ambion).

9. Superscript III primer extension

Position of NMIA or 1M7 modification were mapped by annealing half a picomole of NMIA or 1M7 modified RNA (5 μL) with 1 μL 5'-[^{32}P]-labeled conRT primer, sequence 5'-GAACCGGACCGAAGCCCG-3', via the following thermocycler conditions: 95 °C for 1 min, 65 °C for 6 min, and 35 °C for 10 min. Following a 5 minute incubation on ice, RNAs were supplemented with 5× First Strand Buffer (250 mM Tris-Cl pH 8.3, 375 mM KCl, 15 mM MgCl2) (Invitrogen), 0.1 M DTT to a final concentration of 5 mM, and 10 mM dNTPs to a final concentration of 1.25 mM. The reactions were heated at 50 °C on an Isotemp 125D (Fisher Scientific) for one minute. One microliter of Superscript II reverse transcriptase (100 units, Invitrogen) was immediately added to the reaction mixture, mixed by gentle tapping, and allowed to extend for exactly 4 minutes at 50 °C. The reaction was quenched by the addition of 400 mM NaOH and then heated at 95 °C for 5 minutes in order to degrade the hit RNAs. This was then neutralized by the addition of 400 mM HCl and radiolabeled cDNAs were ethanol precipitated with 5 M ammonium acetate and 200 μg/mL glycogen and then resuspended in 6 μL SHAPE loading buffer (80% Formamide, 0.5X TBE, 4mM EDTA pH 8.0, 0.01% Bromophenol Blue and Xylene Cyanol). Dideoxythymidine sequencing

ladders were generated by the addition of 1.25 mM ddTTP (GE Healthcare) to the reverse transcription reaction of 3.5 picomoles of unmodified RNAs.

10. Sequencing gel electrophoresis

The radiolabeled cDNAs created by the reverse transcription of NMIA/1M7/DMSO hit RNAs were resolved on a 35 cm x 43 cm x 0.4 mm 8% denaturing polyacrylamide sequencing gel (29:1 acrylamide; bisacrylamide/7M urea, 90 mM Tris/borate, 2 mM EDTA) run at approximately 2000 volts in 0.5× TBE for about an hour an a half. The gel was transferred to and dried on Whatman Chromatography paper (GE Healthcare) for 1 hour at 80 °C on a Model 583 Gel Dryer (Biorad) and then exposed on a Phosphor Screen (Amersham Biosciences) overnight. Gels were imaged on a Storm 860 Phosphor Imager (Molecular Dynamics, Sunnyvale, CA) and visualized on ImageQuant TL (Amersham Biosciences).

11. SAFA data analysis and normalization

Individual band intensities of NMIA, 1M7, and DMSO lanes were quantified using the program SAFA (88). SAFA allows the straightening of curved gels (gel recitification) and uses Lorentzian curve integration to determine band intensities with a high degree of accuracy. Full length RNA bands were quantified by employing the analysis method in ImageQuant TL. Radioactivity across the lanes of an individual experiment was then equalized to the level of radioactivity in the NMIA/1M7 lane. Band intensity was calculated by subtracting the value of the DMSO band, control reaction that corresponds to background activity, from the value of the NMIA or 1M7 band giving the

corrected band intensity. Corrected band intensities were then ranked in descending order and the top 2% of values was thrown out. The average of the next 8% of values was taken and this value represented the normalization factor. All of the NMIA/1M7 corrected densities were divided by the normalization factor and then multiplied by 100 to give the normalized values. Normalized values from multiple replicates were averaged and plotted in figures with standard error bars shown.

12. Super telomerase extracts preparation

Super telomerase cell extracts were prepared using a protocol adapted from previously described method (127). One day before transfection, 1.44 x 10^6 HEK293T cells were plated in the wells of a six-well plate. The next day, HEK293T cells were transiently transfected with ~4.5 μg of plasmid DNA (1.2 μg pVan107 hTERT and 3.33 μg pBS-U1 hTR) and Lipofectamine 2000 (Invitrogen) in 6-well plates according to the manufacturer's protocol. One day post-transfection, the cells were trypsinized and three wells were transferred to T-75 flask to allow for additional growth. Seventy-two hours post-transfection, cells were detached from T-75 flask by trypsinization, washed in phosphate buffered saline (PBS) (Gibco), and spun down to remove all liquid in 1.5 mL tubes. The cells were then lysed in 100 μl of CHAPS lysis buffer (10 mM TrisHCl pH 7.5, 1 mM MgCl2, 1 mM EGTA, 0.5 % CHAPS, 10 % glycerol) supplemented with protease inhibitor cocktail III (Calbiochem) and 5 mM β-mercaptoethanol. Two microliters of RNasin (ProMega) was also added to prevent RNA degradation. After incubation on ice for 1 hour, cell debris was removed by spinning down extracts at 4 °C for 5 minutes at 13000 × g. A Bradford assay was conducted to determine the total

protein concentration and then the supernatant was aliquoted, flash frozen in dry ice, and stored at -80 °C.

13. Direct telomerase assay

Telomerase activity was measured by a direct telomerase activity whose protocol was adapted from a previously described assay (113). Each 20 μL reaction contained 50 mM Tris-HCl, pH 8.0, 50 mM KCl, 1 mM MgCl2, 5 mM β-mercaptoethanol, 1 mM spermidine, 1 μM human telomere primer (5'- TTAGGGTTAGGGTTAGGG-3'), 0.5 mM dATP, 0.5 mM dTTP, 2.9 μM dGTP, 0.33 μM [α- 32P]-dGTP (3000 Ci/mmol, 10 μCi/μL; Perkin Elmer) and 2 μL super telomerase cell extracts with the remaining volume being DEPC-treated water. Primer extension was carried out at 30 °C for 90 minutes. After primer extension, reaction volume was increased by the addition of water; a [^{32}P]-labeled loading control (114 nucleotide, 5'-end labeled ssDNA oligonucleotide, 1000 cpm per reaction) was also added. Telomerase assay primer extension products were extracted once with phenol/chloroform/isoamyl alcohol (25:24:1) and twice with chloroform/isoamyl alcohol (24:1) and then ethanol precipitated with 2.5 volumes absolute ethanol, 0.6 M ammonium acetate, and 75 μg/mL glycogen. After precipitating at -20 °C for 15 minutes, telomerase assay products were centrifuged at 13,000 x g at 4 °C for 22 minutes, washed with 1 volume of 70% ethanol, and speed vacuumed till dry. Pellets were resuspended in 6 μL SHAPE loading buffer and heated at 95 °C for 5 minutes and resolved on a pre-warmed, 20 cm x 20 cm x 0.4 mm thick 10% denaturing polyacrylamide gel. The gel was run in 0.5× TBE at 800 volts for approximately one hour. The gel was transferred, dried, exposed, and visualized similar to the sequencing

gel mentioned above. Summing the intensities of all bands in each sample and normalizing the values to the loading control determined the telomerase activity of the reaction.

CHAPTER 4

SHAPE ANALYSIS OF THE SOLUTION STRUCTURE OF hTR

A. Introduction

The secondary structure of human telomerase RNA (hTR) has been studied by chemical, enzymatic, and spectroscopic methods. The first secondary structure depiction of the hTR was determined by comparative phylogenetic analysis and covariation studies (48). In the years since the development of that initial structure, mutational analysis, biochemical investigations, and the identification of key hTR RNA secondary structural motifs has only improved upon the original secondary structure. Recently, NMR has offered high resolution models of specific hTR helices and RNA motifs that indicate that phylogenetic analysis was quite accurate. Currently, the structures of the minimally conserved pseudoknot, the extended P2 helix, the P6 and P6.1 helices, and the P7b helix have been reported (40,47,49,57,128). On the whole, these structures are consistent with the predicted secondary structure though with some interesting revelations that could not be determined by phylogenetic or biochemical analysis. However, there are limitations to these structures, specifically the fact that they were determined in isolated and very controlled conditions. Additionally, determining the structure of a small portion of an RNA does not necessarily support how it will fold in the context of the full RNA.

Due to its size and the limitations of RNA structure determination techniques

including x-ray crystallography, NMR, and computational prediction programs, high-resolution structural data for full length hTR has been unavailable. As a principal component of telomerase, an enzyme whose current level of understanding is already lacking, the contribution of hTR structure to RNA and telomerase biology is an extremely important research topic. In this chapter, we have utilized a high resolution, single nucleotide RNA footprinting technique called selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE) to characterize the structure of full length hTR (85). Our studies of hTR were able to identify with high resolution the structure adopted by full length hTR in solution without hTERT. While the structure of regions of hTR that are important for RNA maturation, processing, and accumulation appear to be similar to the phylogenetic structure, domains that are important for telomerase catalytic activity undergo drastic reorganization which led to the development of an alternate structure for hTR in solution. Full length hTR structural data was utilized in further studies that identified aminoglycosides as telomerase activity inhibitors due to their ability to bind to the CR4-CR5 domain of hTR and prevent the assemblage of active telomerase complexes.

B. Results

1. Addition of 3' flanking region to hTR for SHAPE

In order to analyze the full-length human telomerase RNA by SHAPE, hTR was cloned with a 3' flanking region. The 3' flanking region includes an 18 nucleotide reverse transcription primer binding site at the very 3' end followed by a 24 nucleotide linker which allows for the reverse transcriptase to become fully processive before encountering

the RNA of interest; this is the same 3' flanking region added to tTR and the hTR domains (85). The 5' flanking region was not added as previous experiments conducted in our lab demonstrated that when the 5' and 3' flanked RNA was reassembled in the telomerase complex telomerase activity was diminished (JDL Dissertation). The 43 nucleotide extension forms two stable stem loop structures to avoid folding interactions with hTR (85). Telomerase reassembled from *in vitro* translated hTERT and the flanked hTRs formed active telomerase complexes.

2. SHAPE of hTR in solution

Flanked hTR was incubated with 1M7 and then the acylated RNA was mapped using radiolabeled primers to create a cDNA library that corresponded the SHAPE hit reactivities for each nucleotide in the RNA. The cDNAs were separated by gel electrophoresis and then data was quantified using a SAFA (Figure 4.1 and 4.2) (88). To evaluate the full length hTR SHAPE data, we superimposed it upon the accepted mutationally and phylogenetically derived secondary structure of hTR (Figure 4.3) (48). The SHAPE profile is consistent with many of the predicted base pairs including the template boundary element, helical regions of the CR4-CR5 and CR7 domains, and the box H/ACA domain (Figure 4.3). In these portions of hTR, RNA stems and base pairs show minimal SHAPE reactivities while internal loops and bulged nucleotides display moderate to high flexibilities consistent with their single-stranded nature. Several regions including the 5' end of hTR, template, pseudoknot-containing core domain, and 5' and 3' pockets do not form as predicted in the phylogenetic structure suggesting that hTR in solution, and not bound in the catalytic telomerase complex, actually adopts an

alternative structure. To identify the alternative structure adopted by hTR in solution, we folded hTR with the program RNAstructure incorporating the SHAPE data as pseudo-energy constraints (121). One of the main limitations of this program is its inability to detect and predict tertiary RNA interactions such as pseudoknots (121). To overcome this obstacle, we also subjected hTR to SHAPE Knots, an RNA prediction program

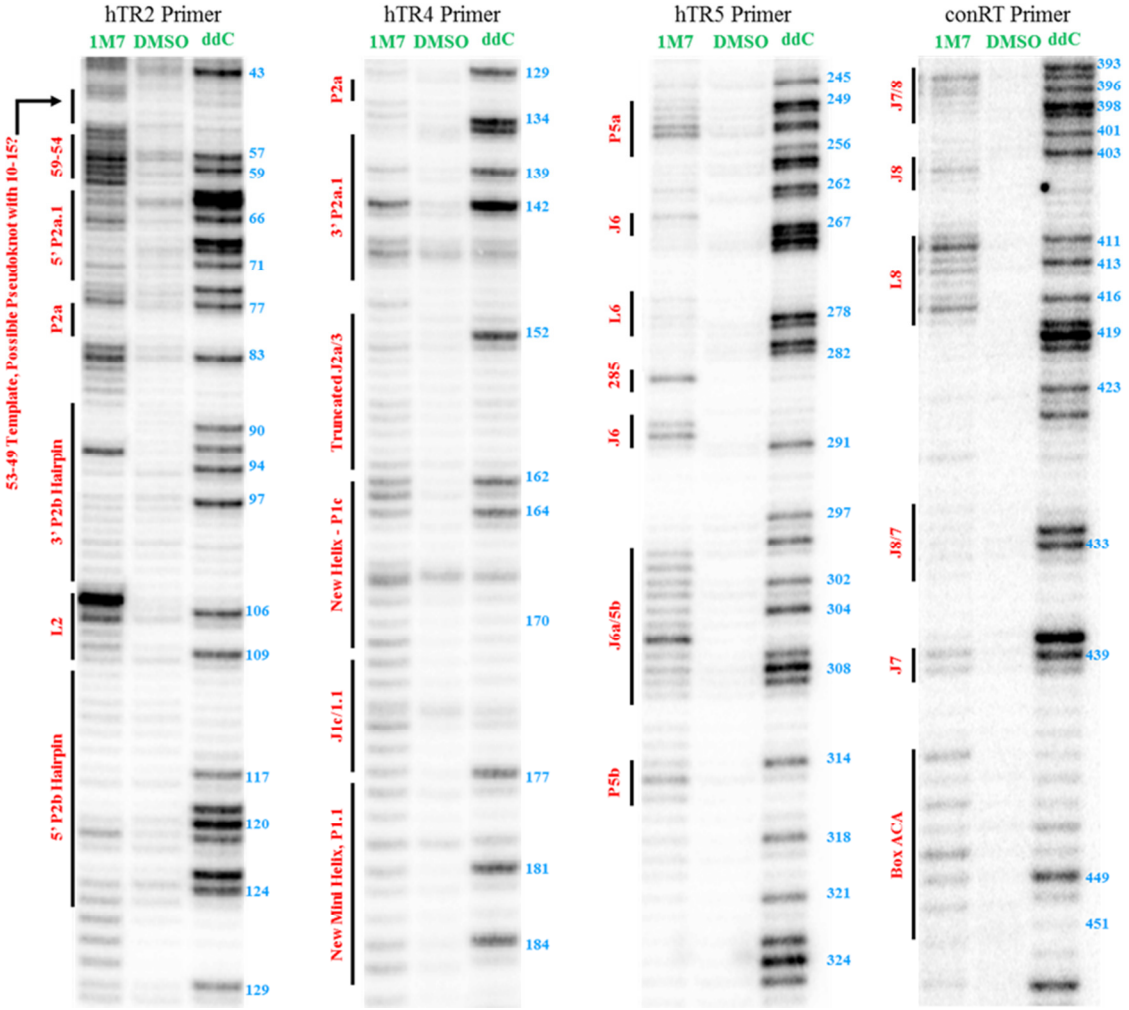


Figure 4.1 SHAPE gel of full length wildype hTR, selected primers. Full length hTR with 3' flanking regions was treated with 5 mM 1M7 or 10% DMSO (background) and the acylated RNAs were mapped by reverse transcription using radiolabed primers. Due to the size of the RNA, 7 different primers (hTR1-6 and conRT) were used in order to map the entire RNA. The ^{32}P-radiolabeled cDNAs were resolved on a denaturing sequencing gel. Nucleotides "hit" during SHAPE cause reverse transcription to stop one nucleotide before the modification site. A dideoxycytidine (ddC) ladder (serves as a marker for orientation and nucleotide identification) was also run and indicates the location of adenines in the RNA sequence. The gels are numbered according to hTR nucleotide number, not the ddC ladder which is off by +1 nucleotide. To the left of the gels, are the identification of structural features of hTR.

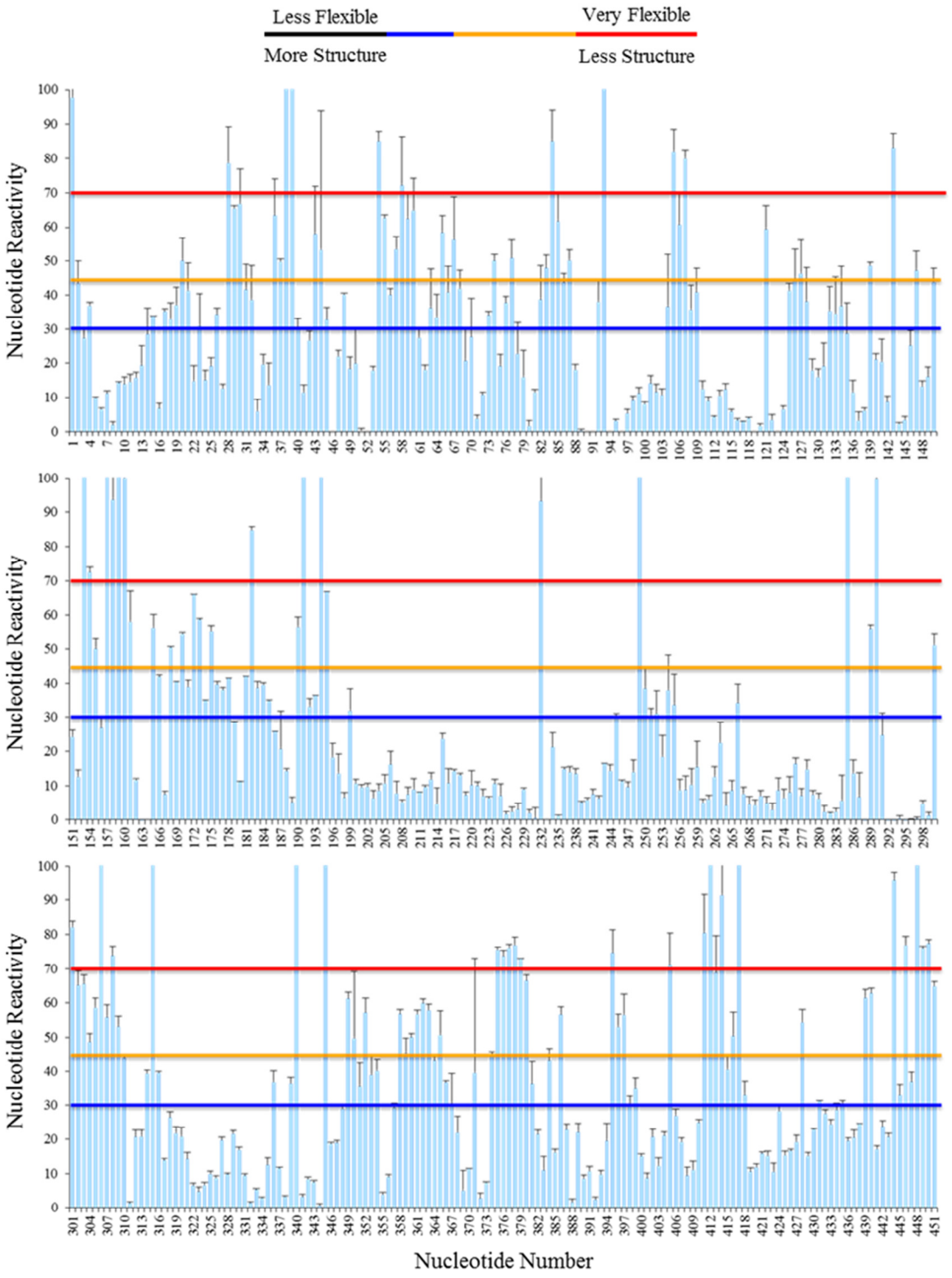


Figure 4.2 SHAPE analysis of full length hTR. Histograms of full length hTR. The average band intensities were normalized and plotted according to nucleotide position with error bars (N=3-9). Black, blue, orange, and red lines indicate divisions for examining nucleotide reactivity values where $0 \leq$ black $<$ 30, $30 \leq$ blue < 45, $45 \leq$ orange < 70, and red ≥ 70. Lower numbers represent less flexibility and higher numbers represent greater flexibility.

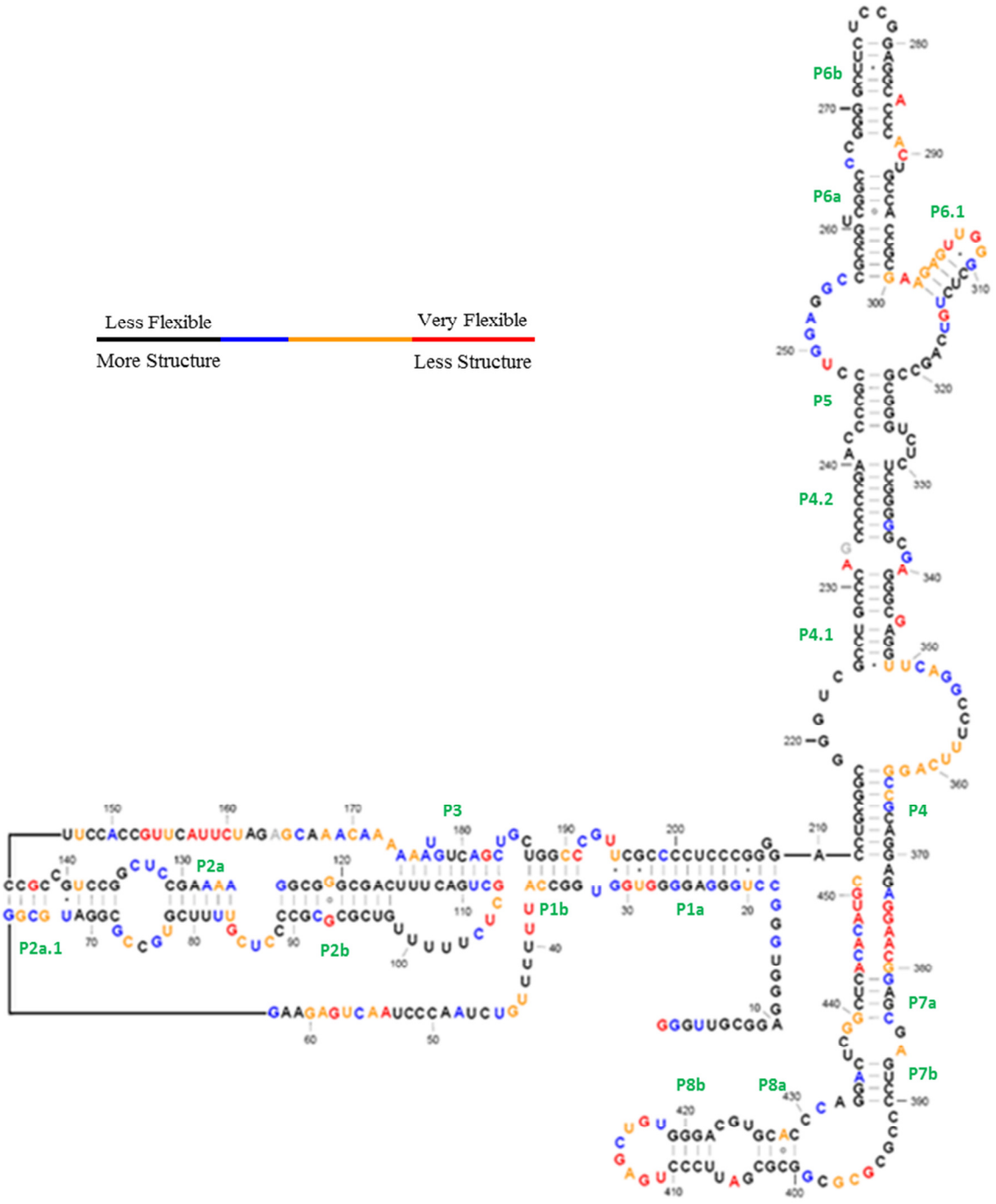


Figure 4.3. Phylogenetic secondary structure of hTR with SHAPE data superimposed. Phylogenetic secondary structure of hTR with nucleotides reflecting full length hTR SHAPE data (48). The color-coded nucleotides reflect the histogram data from Figure 4.2. The important loops and helices have been numbered accordingly.

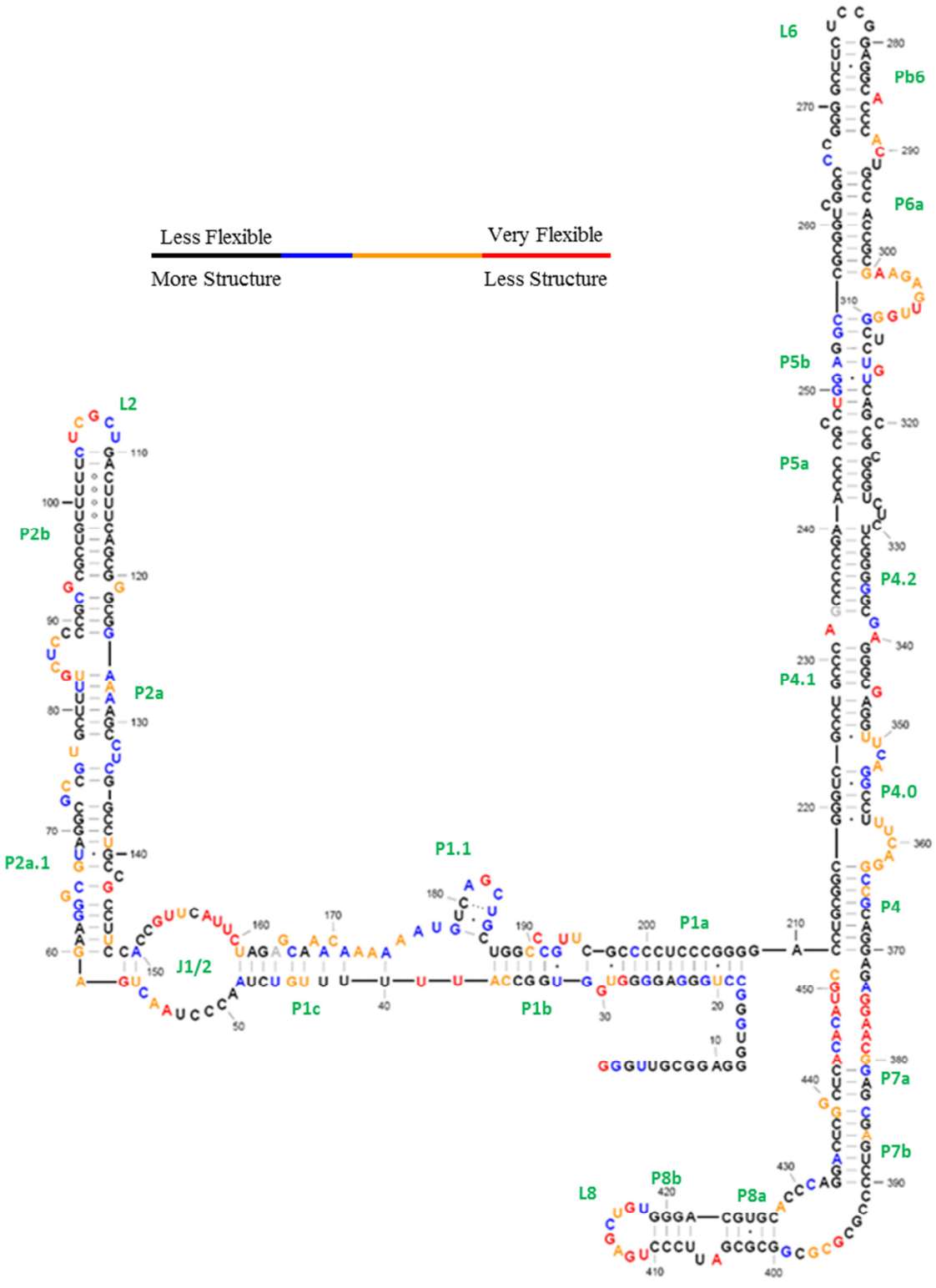


Figure 4.4. Proposed RNA secondary structure of protein-free hTR. The secondary structure was predicted by constraining RNAStructure with full length SHAPE data. The color-coded nucleotides reflect the histogram data from Figure 4.2. The important loops and helices have been labeled to differentiate with the phylogenetic structure. There is no helix 3 (P3).

developed by the Weeks Lab at UNC that demonstrates high sensitivity and can correctly predict pseudoknot structures (Hajdin, Weeks, unpublished). The output from RNAstructure was critically analyzed and further modified to accurately depict all the SHAPE data and rectify shortcomings of RNAstructure (129). Visual inspection of the newly predicted structure for hTR in solution demonstrates a similar RNA architecture to the phylogenetic structure with the most significant differences being that the core domain pseudoknot does not form and actually adopts an extended hairpin formation and that the single stranded template region is involved in helix formation (Figures 4.3 and 4.4). Results from SHAPE Knots also confirm that the core domain pseudoknot does not form (data not shown).

a. 5' hTR and P1

The first 17 nucleotides of hTR have been predicted to be single stranded. Interestingly, SHAPE analysis demonstrates that all of these nucleotides, with the exception of G1, have minimal to low reactivities. This suggests that they are involved in the formation of some sort of RNA secondary or tertiary structural motif or intramolecular interaction with another portion of hTR that is unable to be predicted by RNAstructure or SHAPE Knots (34,130). SHAPE data reveals that the P1 helix forms a well-defined structure. Within the P1a helix, C18-G31 mostly have low SHAPE reactivities which corresponds well with their forming base pairs with C196-G209 (Figure 4.4). Nucleotides U20 and U29 do exhibit moderate SHAPE reactivities though both of these nucleotides are predicted to participate in G•U base pairs which have been shown to induce structural perturbations in helices composed of Watson Crick base pairs

and to be conformationally soft (131). The nucleotides flanking U20 are slightly flexible supporting conclusions that the nature of the G•U base pair makes these bases more accessible to SHAPE modification. U29 also forms a G•U base pair and its flanking nucleotides are both highly flexible with G30 even making a bulge in the helix (Figure 4.4). Surprisingly, the Watson Crick companion of G28, C198, has minimal reactivity. The G31•U195 base pair does not appear to form and the J1 internal loop (internal loop that is entirely within the P1 helix) seems to have shrunk and now only consists as the J1b/1a loop (the single stranded region between helix P1b and helix P1a) that contains highly flexible U194 and U195. The P1b helix, or the template boundary element, forms much as predicted in the phylogenetic structure. The helix has been elongated by the addition of the U32•G193 base pair, both of whose nucleotides exhibit small amounts of flexibility, and C191 bulges with its two neighboring nucleotides also exhibiting moderate levels of SHAPE reactivity (Figure 4.4).

b. Template and extended P1 helix

Phylogenetic analysis of hTR identified the template region, nucleotides U37-G63, as being a long single stranded loop region that connected the template boundary element (P1b) to the first helical region of the proposed pseudoknot (P2a.1). SHAPE data refutes this claim and demonstrates that the template region actually forms an extension of the P1 helix (Figure 4.4). In looking at the augmented P1 helix, there are sequences of hTR that were predicted to be in loop region J2a/3 and the 3' region of the P3 helix that are interacting with the formerly single-stranded template region (Figure 4.4). Following P1b, a new J1 internal loop structure is formed and includes U's 38 and 39, which both have high SHAPE reactivities, and A175-U77, which have moderate

flexibilities, among its single-stranded portions. In the J1c/1b strand of J1, is a small helix with a three base pair stem and three-nucleotide loop that protrudes from the internal loop structure similar to the proposed P6.1 helix in the phylogenetic structure. Included in the three base pair helix of P1.1 are the U179•G185 wobble pair and a non-canonical C180•U184 loop closing base interaction.

After the J1 internal loop, is a run of single base pairs separated by one- or two-nucleotide bulges that form with the beginning of the new P1c helix and include nucleotides U41-G44, which precede the actual template sequence, and G165-A171, which preceded the 3' segment of the P3 helix in the phylogenetically-derived structure (Figure 4.4). Nucleotides in this region display varying flexibilities due to the many bulges that flank the predicted base pairs. The end of the P1c helix is composed of a four base pair helix, all of whose nucleotides have minimal to slight SHAPE reactivities; wrapped up in this helix are the first three nucleotides of the actual hTR template. The final structural feature of the extended P1 helix is the large J1/2 internal loop that connects P1c to P2a.1. Even though it is predicted to be single stranded, the first five nucleotides of the J1c/2a.1 loop, A49-U53 which also make up the last five nucleotides of the hTR template sequence, exude minimal SHAPE reactivities suggesting their involvement in the formation of an RNA secondary structure interaction (Figure 4.4). Nucleotides A55-C57 in this same loop have moderate flexibilities that correspond to their being in a loop structure. On the other side, J2a/1c has two nucleotides with minimal flexibility (C151 and C152) followed by eight nucleotides with relatively high SHAPE reactivities (G153-C160).

c. Extended P2 helix, no P3

The P2 helix is known for contributing stem 1 and loop 1 to the highly conserved pseudoknot tertiary structure that is conserved across telomerase RNAs and is essential for enzyme catalytic activity (48). Examination of SHAPE data for hTR in solution does not support the formation of the pseudoknot, demonstrating that the RNA prefers to adopt the elongated hairpin conformation that the pseudoknot is known to exist in equilibrium with (41). The 5' end of the P2a.1 helix begins with base pair G60-C148 as compared with G63-145C in the phylogenetic structure showing that even this section of the P2 helix is extended. Nucleotides G60-G62 and U147-U148 were originally thought to be parts of the single stranded template region and J2a/3 loop regions respectively. Instead of forming a nine base pair helix with a single bulged C139, P2a.1 now forms a twelve base pair augmented helix that is interrupted by two bulged nucleotides, G65 and C142. While G65 shows moderate SHAPE reactivity, C142 is not reactive demonstrating a structural discrepancy (Figure 4.4). Within P2a.1, nucleotides neighboring the bulges and U68•G140 wobble pair show increased reactivity while the remaining nucleotides are minimally impacted by SHAPE. The J2a internal loop that separates helices P2a.1 and P2a has been condensed and is now present as two separate two-nucleotide bulge regions on either side of the junction between these two more defined helices. Nucleotides that were thought to be contained within the internal loop, C75 and G76 from J2a.1/2a and G135 and C134 from J2a/2a.1, have formed Watson Crick base pairs across the loop due to all having low SHAPE reactivities (Figure 4.4). Internal loop nucleotides G73, C74, U77, C132, and C133 remain single stranded and all exhibit low to moderate reactivity in response to SHAPE modification.

Nucleotides G78-98 and C116-C130 form the P2a helix, J2a/2b, and beginning of P2b exactly as they do in the phylogenetic secondary structure with no changes to base pairings, loop size, or helix length. From this point on, drastic structural rearrangement in RNA architecture is denoted from the SHAPE reactivity profile that matches up exactly with the DC pseudoknot construct that was examined in Chapter 3 (Figure 3.3C) (42). As there is no P3 helix, U's 99-103 cannot form the conserved Hoogstein triples that stabilize the pseudoknot structure, and instead form three U•U base pairs with U's 113-115. After the U•U base pairs, there is one C•U base interaction before two Watson-Crick base pairs finish up the elongated P2b helix which is now 15 base pairs as compared to 9 base pairs when the pseudoknot was the accepted structure for this region. All the nucleotides in the elongated P2b stem loop structure lack SHAPE reactivity supporting helix formation (Figure 4.4). The four nucleotide J2b/3 loop seen in the pseudoknot conformation is expanded and now makes a five nucleotide hairpin loop, U105-U109, all of which are relatively reactive.

d. Hypervariable region and CR4-CR5 domain

Comparison of the hypervariable region and CR4-CR5 domains from the phylogenetic and newly proposed secondary structure for in-solution hTR illustrates that both structures comprise of a similar long helical stem with bulges and internal loop regions that interrupt the extended helix. The alternative structure, however, is much more refined and helical than the phylogenetic structure. Nucleotides G209 and A210 create a single stranded linking region between P1 and the hypervariable region. Based on the SHAPE data, it would be interpreted that these nucleotides actually contribute to

RNA structure; perhaps hTR folding positions these two nucleotides in a solvent inaccessible area that is therefore unable to be impacted by SHAPE modification. The P4 helix forms exactly as it is predicted in the phylogenetic structure, though the 5' pocket which sits atop this helix, has been condensed and divided. Nucleotides G219-C223 pair across the pocket with G353-U357 creating a five base pair helix (P4.0) leaving two segregated loops on either side, all of whose nucleotides exhibit moderate levels of SHAPE reactivity and substantiate their involvement in making internal loops (Figure 4.4). Moving up from the now divided 5' pockets, the P4.1 and P4.2 helices form almost exactly as they do in the phylogenetic structure; the only difference being that the P4.2 helix is one base pair longer as a G233-C338 base pair forms from two of the nucleotides that were proposed to be single-stranded and part of the small J4.2/4.3 internal loop breaking up those helices P4.1 and P4.2. The three nucleotides that remain in that truncated internal loop region all demonstrate flexibility indicating that the dividing loop is actually single stranded (Figure 4.4).

The J4/5 loop is shown to separate the hypervariable region from the CR4-CR5 domain of hTR in the phylogenetic structure. SHAPE data confirms that the two helices are isolated, though instead of an internal loop with two single stranded regions on each side, a single sided loop J5/4.3 that includes nucleotides C328-C330 seems to form despite the fact that these nucleotides exhibit minimal amounts of SHAPE modification (Figure 4.4). Nucleotides A241 and U327, that were both part of the J4/5 loop originally, now form the first base pair in a new and lengthened P5 helix that is subdivided in two sections, P5a and P5b. P5a includes nucleotides that were originally the P5 helix, forming a six base pair helix of mostly Watson Crick base pairs that is broken apart near

the middle by a single bulged nucleotide C323. Like many other instances within hTR, the bulged nucleotide is actually not very flexible suggesting that the nucleotide may stack with neighboring bases or its base may be pointed toward the helix possibly creating a triple base or base-backbone interaction that obscures it from SHAPE reaction (Figure 4.4). Partitioning the P5 helix from creating one continuous helix are a pair of C nucleotides C247 and C320, neither of which are flexible, but both bulge.

The P5b helix assembles from base pairing across the J5/6 helix, and includes all of the nucleotides portrayed to be single stranded in J5/6a strand of the phylogenetic structure, C248-C255, making base pairs with G310-G319, the minimally flexible nucleotides that were part of the P6.1 helix and J6.1/5 strand of the internal loop. P5b contains two bulges, U312 and G314; G314 is highly flexible and causes nucleotides in its immediate vicinity to have increased flexibility, while U312 is not reactive to SHAPE (Figure 4.4). The P6.1 helix, which has been shown under many circumstances to be requisite for telomerase activity and processivity, does not form. Instead, the highly flexible nucleotides A301-G309 that once made up the 5' strand of the P6.1 helix produces a nine-nucleotide loop region. Interestingly, this is the same structure that was shown to form in the G305A aplastic anemia mutant in Chapter 3 (Figure 3.6C). From this point in the structure to UCCG tetraloop that tops off the P6b helix, the alternative structure and the phylogenetic structure match up exactly with one small exception. It has been predicted that a non-canonical C262•A295 base pair exists in P6a with U261 bulging out before it. SHAPE analysis and chemoenzymatic footprinting analyses show that a Watson Crick U261-A295 base pair forms instead with C262 being the bulged

nucleotide (35). Intriguingly, C262 also does not exhibit SHAPE reactivity suggesting that its conformation with respect to the helix may decrease its flexibility.

e. Box H/ACA and CR7 domains

In agreement with the phylogenetic structure, the Box H/ACA domain of hTR is revealed to be single stranded. The only discrepancy being the first two nucleotides of the box H region, A371 and G372, appear to be unreactive to SHAPE modification (Figure 4.4). The CR7 domain is more structured than it appears in the phylogenetic structure and correlates well with the wildtype CR7 domain that analyzed in Chapter 3 (Figure 3.7B). Like the data from the CR7 domain SHAPE, the CR7 domain in the full length hTR begins with a shortened three base pair P7a that is separated from P7b by a single bulged and moderately reactive to SHAPE G440. While G439 is shown to base pair with C383, G440 is also demonstrated to be moderately flexible suggesting the existence of conformational flexibility in that either nucleotide could potentially base pair with C383 or be bulged. The seven base pair P7b helix is followed by the 3' pocket that forms exactly as it did in the isolated CR7 domain with a nine-nucleotide J7b/8a strand and a five-nucleotide J8a/7b strand. The only nucleotides included in the internal loop that are moderately reactive are G395-G397, suggesting that the remaining nucleotides may form base pairs across the loop or adopt another constraining or solvent inaccessible RNA secondary structural configuration. The rest of the domain looks exactly like to the CR7 domain investigated with the P8 helix still forming two smaller stacked helices that are separated by a two-nucleotide bulge consisting of A404 and U405 with A404 being highly reactive and U405 showing minimal flexibility (opposite of the domain SHAPE on

the two-nucleotide bulge) and capped by the highly conserved and reactive eight-nucleotide loop (48,52) (Figure 4.4).

3. Aminoglycosides prevent telomerase assemblage

Using a method called 2-Dimensional Combinatorial Screening (2D-CS) to identify RNA motif-ligand partners, the Disney Lab (Scripps Research Institute, Jupiter, FL) identified aminoglycosides kanamycin-6" Az-FITC (Kan-6") and neamycin-5-O-(AzEt)-FITC (Nea-5-O) as small molecules that would bind to hTR (132). In order to determine if this binding would inhibit telomerase activity, we performed two different

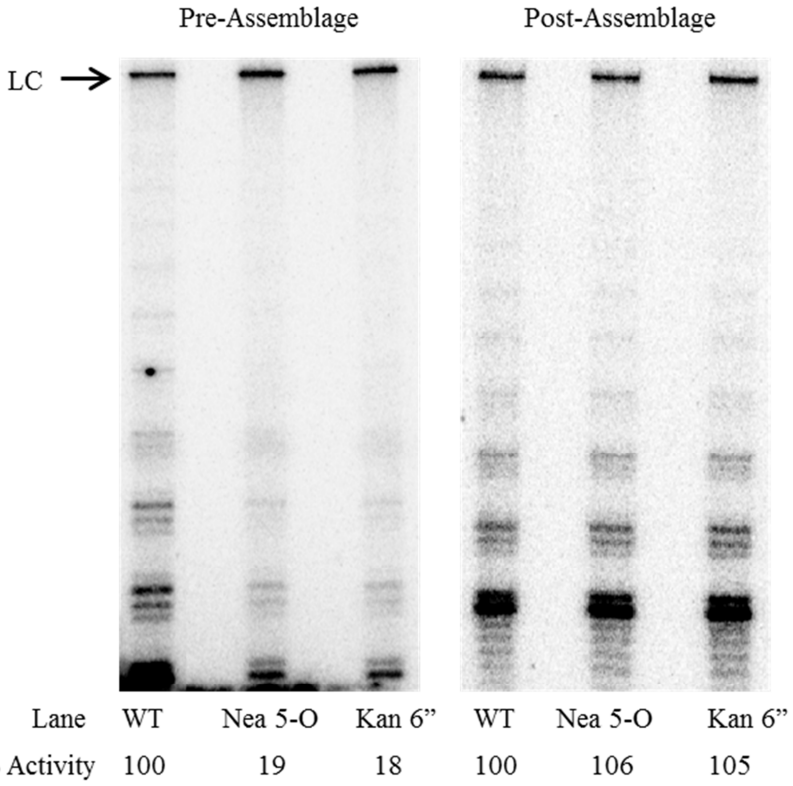


Figure 4.5. Effect of aminoglycosides on telomerase activity. Pre-assemblage and post-assembalge telomerase assays were conducted with *in vitro* assembled human telomerase and aminoglycosides Nea 5-O and Kan 6" to see the affect of the compounds on telomerase activity. In pre-assembalge assays, compounds are added prior to assemblage, and during post-assemblage assays, compounds are added after telomerase complexes had been formed. The loading control (LC) at the top of each assay was used to normalize telomerase activity for each experiment.

direct telomerase assays that were developed in our lab (113,133). In the pre-assemblage assay, Kan 6" and Nea 5-O to a final concentration of 50 μM was incubated with *in vitro* translated hTERT in RRL and *in vitro* transcribed hTR prior to the assemblage reaction. Next, telomerase activity was evaluated via a primer extension assay. The pre-assemblage assay demonstrates that both Kan 6" and Nea 5-O were potent inhibitors of telomerase activity displaying 18% and 19% of wildtype activity respectively (Figure 4.5). To ensure that these compounds do in fact inhibit telomerase by binding hTR and preventing proper association of telomerase, we also performed a post-assemblage assay where Kan 6" and Nea 5-O to a final concentration of 50 μM was added to the primer extension assay with fully assembled telomerase complexes. Assessment of post-assemblage assay showed no real inhibition of either telomerase activity or nucleotide addition processivity confirming that the aminoglycosides inhibit telomerase activity by preventing telomerase assemblage (Figure 4.5).

4. Identifying binding sites of Nea 5-O and Kan 6"

Knowing that Nea 5-O and Kan 6" did in fact inhibit telomerase activity, we next set out to determine the actual binding site of these two compounds within the structure of hTR. With the new alternative structure of hTR in solution to use for comparison, we incubated Nea 5-O and Kan 6" with hTR and then performed SHAPE to identify nucleotides whose SHAPE reactivities fluctuated due to the presence of the aminoglycosides. Taking into account the predicted binding sites suggested by the Disney Lab, we focused on examining the pseudoknot core domain for Nea 5-O and the CR4-CR5 domain for Kan 6". At 20 μM final concentration, large fluctuations in the structure of the pseudoknot

domain appear when Nea 5-O is present (Figure 4.6A). As the majority of the domain experiences changes in SHAPE reactivity, we decided not to pursue further experiments with Nea 5-O at this time.

In contrast, incubation of hTR with the same concentration of Kan 6", indicated that changes in nucleotide flexibility were more localized to specific RNA motifs within the CR4-CR5 domain (Figure 4.6B). At 20 μM, nucleotides within the J5/6 internal loop and L6 end loop both exhibited decreased SHAPE reactivity and therefore appear to be

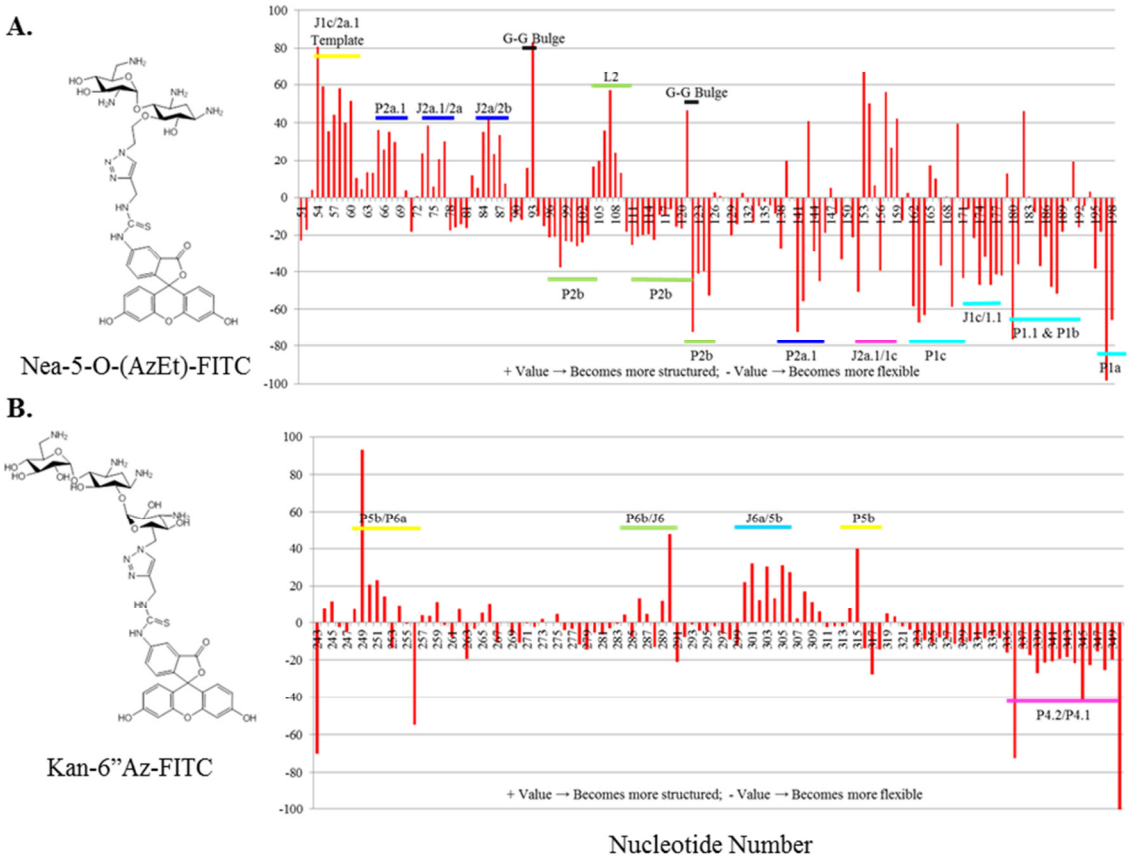


Figure 4.6. Compounds binding to hTR elicit changes in secondary structure. Full length hTR was treated with two compounds that have been suggested to inhibit telomerase activity by binding to hTR and preventing proper assemblage with hTERT and SHAPE was performed. **(A)** Nea-5-O-(AZEt)-FITC is proposed to bind in the core domain and **(B)** Kan-6" Az-FITC is proposed to bind in the CR4-CR5 domain. Histograms represent differences in flexibility between naked hTR and hTR that was incubated with each respective compound at a final concentration of 20 μM. Positive values indicate decreased flexibility while negative values represent increased nucleotide flexibility. Specific regions of hTR that exhibited changes in structural flexibility are denoted.

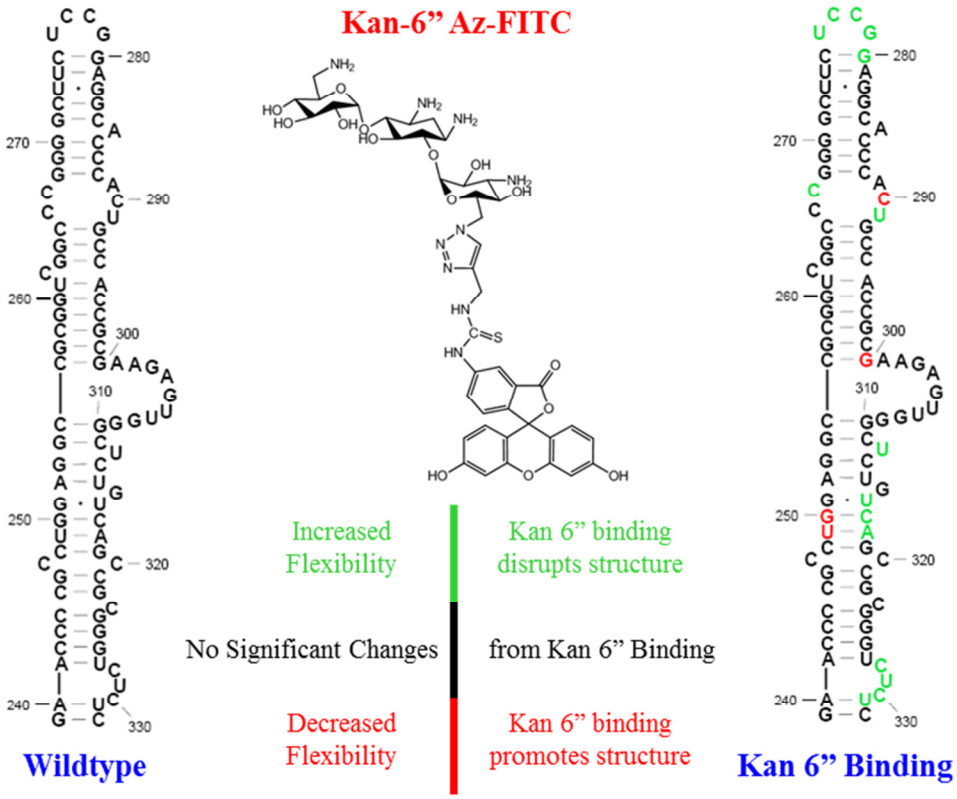


Figure 4.7. Kan 6" binds to the CR4-CR5 domain of hTR. Full length hTR was treated varying concentrations of Kan-6" Az-FITC in order to pinpoint possible binding sites. Comparison with the CR4-CR5 domain data from SHAPE of full length hTR identified nucleotides whose SHAPE reactivities increased (green) and decreased (red) in response to treatment with Kan 6".

becoming more structured, while the P4.1 and P4.2 helix regions appear to become more flexible. To further examine structural perturbations caused by the binding of Kan 6" to the CR4-CR5 domain of hTR, we investigated changes in the CR4-CR5 domain from varying concentrations of Kan 6" from 1-100 μM. After analyzing the data, we identified nucleotides that exhibited changes in flexibility, either increasing or decreasing, over the course of the change in concentration of the compound added. Specifically, single stranded nucleotides in J5a/4.2 strand and the L6 loop demonstrated significant increases in nucleotide flexibility, as both of these loop regions display minimal SHAPE reactivity in the naked RNA (Figure 4.7). Other regions that display changes in nucleotide

flexibility are just below the G251•U316 wobble pair in the P5b helix and in the J6 helix, a suggested hTERT binding site (Figure 4.7) (47).

C. Discussion

Phylogenetic analysis and covariation studies from vertebrates have identified eight conserved regions within the structure of human telomerase RNA with the longest of these being the template, the putative pseudoknot helices P2b and P3, the J5/6 loop region of the CR4-CR5 domain, and the H/ACA box domains (48). Comparatively, the primary sequence of the 5' end, template boundary helix P1, and P2a and loop regions of the core domain are not as conserved (48). Taking this into consideration, it is not altogether surprising that these regions fold differently than is predicted when hTR is in complex with hTERT in the active telomerase complex. As phylogenetic secondary structures are developed from sequence comparison, the development of new, higher resolution techniques for interrogating RNA secondary and tertiary interactions can produce more conclusive structural data that allows for clarifying and updating proposed structures. Considering the phylogenetic data, it is not surprising that SHAPE data supports the predicted structure of hTR in regions that are highly conserved, but appears to deviate from the predicted structure in areas of low conservation in defined solution structures. However, it is also not altogether shocking that analysis of full length hTR may suggest deviations in structure that techniques like NMR or footprinting studies have overlooked from analyzing only smaller and/or constrained portions or unpure samples of RNA as well.

To date, the exact role of the 5' leader sequence of hTR has not been well established despite its presence in the telomerase RNAs of many mammals (134). Deletion of the first 17 nucleotides has been shown to have no effect on telomerase activity which would suggest that these nucleotides are inconsequential (39). Though phylogenetic analysis has suggested that the first 17 nucleotides of hTR are single stranded, several other studies have implicated this region in the formation of intramolecular and intermolecular RNA interactions (34,134). Chemoenyzmatic probing of the 5' G-rich RNA sequence gave mixed results with many nucleotides interacting with both single and double-stranded specific agents proposing that the 5" end of hTR may assume multiple structures in an *in vitro* setting (Figure 4.9) (35). SHAPE data conflicts with this assessment as all of the nucleotides from G2-G17 have minimal to slight reactivities further supporting the notion that these nucleotides are involved in the formation of secondary or tertiary RNA structures (Figure 4.4).

Based on the SHAPE data and what we know about RNA structure, we have proposed two possible structures that could may form and take into account nucleotide flexibility: a G-pseudoknot with the minimally reactive nucleotides from the template region and the formation of a 5' G-quadraplex structure (Figure 4.8). While there is support in the literature and from other studies for the 5' G-quadraplex, the G-pseudoknot is a new proposition. Even though the SHAPE Knots RNA structure prediction program was unable to predict a pseudoknot between these two RNA segments, the decreased flexibilities and primary sequence of the nucleotides A10-U14 in the 5' G-rich strand and A49-U53 in the J1c/2a.1 loop match perfectly (Figure 4.8A). One of the main shortcomings of the G- pseudoknot structure is the unexplained SHAPE reactivity of

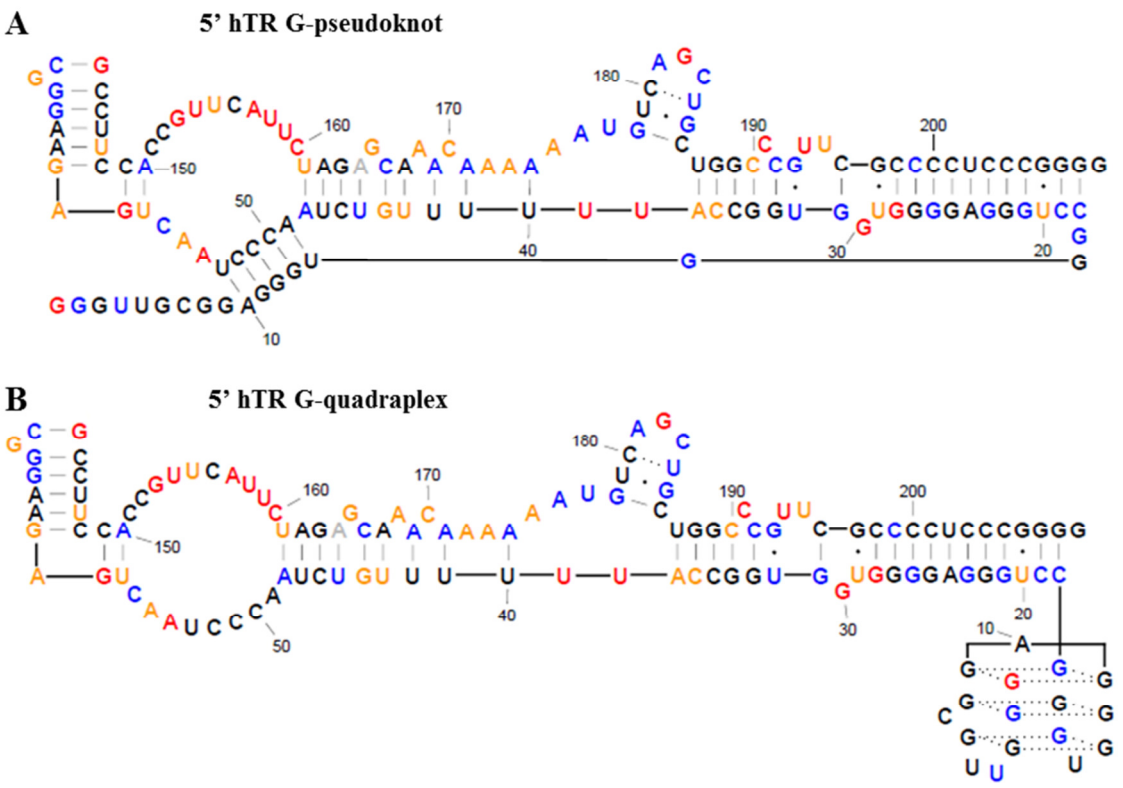


Figure 4.8. Secondary structural motifs of 5' hTR. Proposed secondary structures of 5' nucleotides of hTR. Previous studies have identified the first 17 nucleotides of hTR as single stranded, however, our SHAPE data demonstrates that these nucleotides more than likely participate in secondary or tertiary structure formation due to their low flexibilities. Using what we know about RNA secondary structures and data knowledge of previous work investigating the 5' nucleotides of hTR, we proposed two structures: **(A)** 5' hTR G-pseudoknot and **(B)** 5' G-quadraplex which better explain our hTR SHAPE data.

nucleotides G1-G9 in hTR; it would not be expected for five of the nucleotides to base pair and then literally leave nine others hanging (Figure 4.8A).

Unlike the G-pseudoknot structure, the formation of a G-quadraplex (G4) structure within hTR has been suggested and proven before as G-rich sequences are capable of forming four-stranded helical structures based on the stacking of multiple G• G• G• G•G tetrads (G-tetrads) (135). A 17-nucleotide RNA with the same sequence as the first nucleotides of hTR has been shown by circular dichroism to adopt a G-quadraplex in the presence of K^+ or Li^+ and K^+ cations (34,130). When the same studies were completed with just Li^+ cation, G-quadraplexes did not form as expected (34). This is

because only Na^+/K^+ cations are able to stabilize the G-quadraplex structure by coordinating the carbonyl groups of the guanines at the center of the G-tetrad core (135). Enzymatic probing of the G-quadraplex structure by RNase T1, confirmed protection of guanines in the presence of K^+, but not Li^+ further supporting the formation of the G4 structure (134). This is in support of SHAPE data which demonstrates that nucleotides G2-G17 display minimal amounts of reactivity, suggesting that they are involved in the formation of a structure that protects the nucleotides from nucleotide acylation (Figures 4.2, 4.4, and 4.8). Though G1 does show high reactivity, this could be attributed to its being the first nucleotide and its positioning near the corner/edge of the G4. In a variant of the first 17 nucleotides of hTR, where G2, G8, G12, and G16 were mutated to A, no G-quadraplex formation was observed (34). More recently, studies have demonstrated the hTR interacts with DHX36 (also known as RHAU), a DEXH box RNA helicase that specifically binds to tetramolecular G4-RNA with high affinity (130,136). Replacement of K^+ cations with Li^+ impaired this interaction comparably, and not surprisingly also strongly reduced the thermodynamic stability of the hTR G4 (136). Interestingly, DHX36 associates with the fully assembled telomerase holoenzyme as well, proving that the G4 structure of hTR is maintained even after telomerase assemblage (136). Along with SHAPE, these studies provide compelling evidence that the first 17 nucleotides of hTR, phylogenetically predicted to be single stranded, actually form an intramolecular G-quadraplex and that this conformation is maintained even when hTR is in complex with hTERT.

The phylogenetic representation of hTR identifies P1 as two stacked helices, 14 base paired P1a and 5 base paired P1b, separated by the J1 internal loop that form the

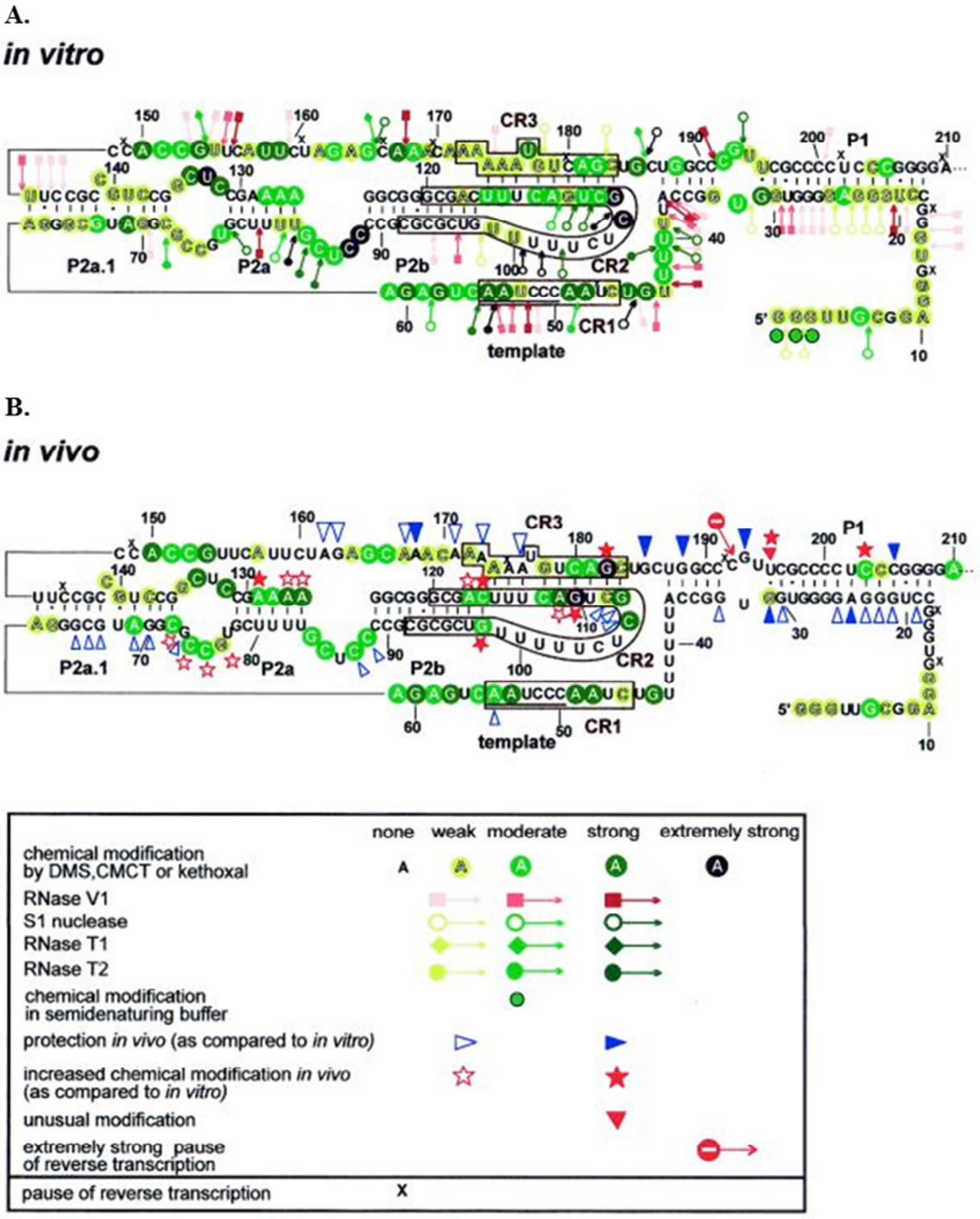


Figure 4.9. Chemoenzymatic footprinting analysis of hTR nucleotides 1-210. The structure of full length hTR was probed by chemical modifications and enzymatic cleavage. This is a representation of the results for the 5' end of hTR including nucleotides 1-210. This figure is taken from Antal et al. 2002 (35).

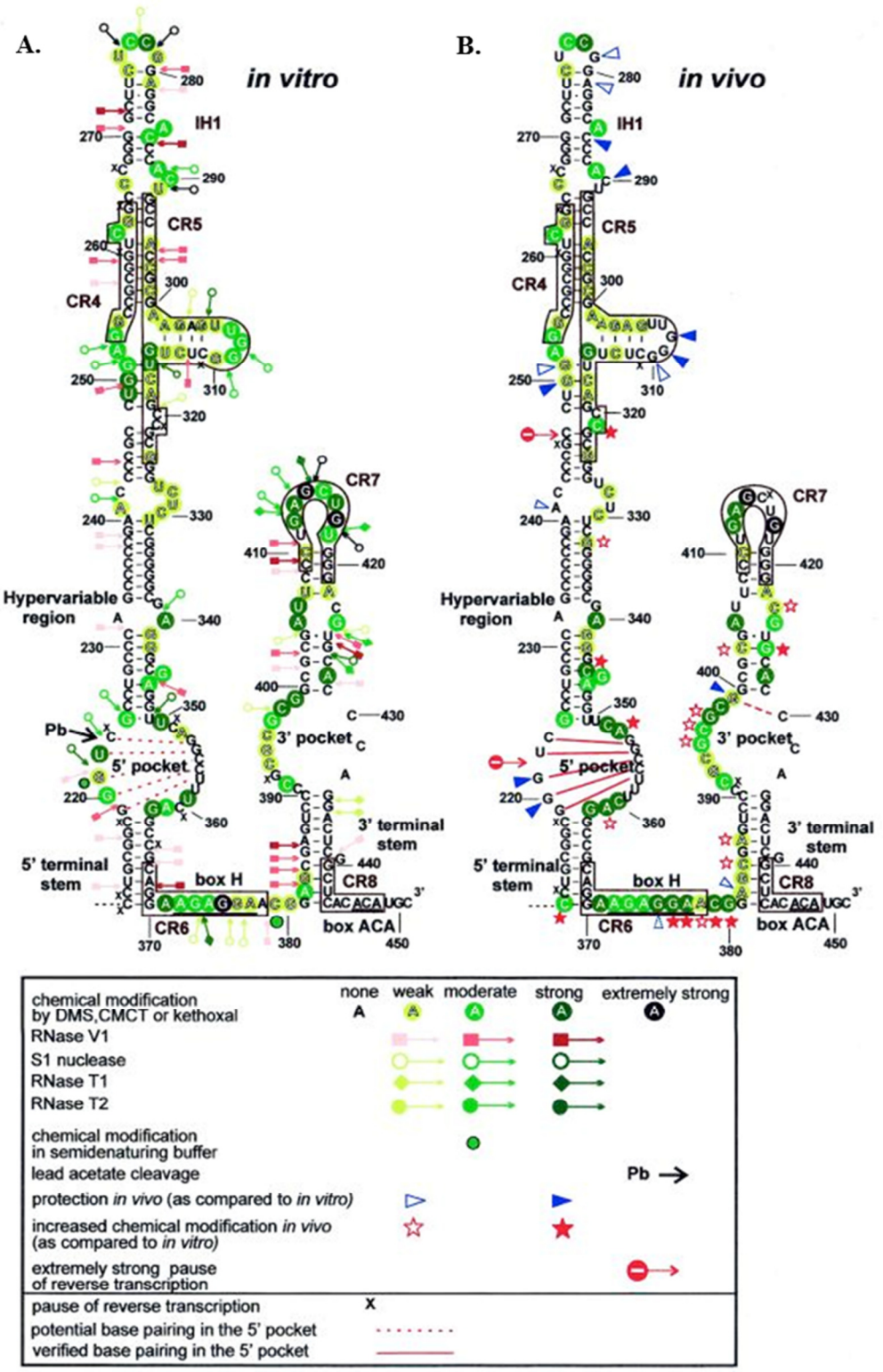


Figure 4.10. Chemoenzymatic footprinting analysis of hTR nucleotides 211-451. The structure of full length hTR was probed by chemical modifications and enzymatic cleavage. This is a representation of the results for the 3' end of hTR including nucleotides 211-451. This figure is taken from Antal et al. 2002 (35).

beginning and terminal helix that separates the hTR 5' region from its 3' portion. This structural feature is known to be conserved in vertebrates, though it is not present in rodents (36). Disruption of P1a has minimal effect on reconstituted telomerase activity, while deletions that disrupt both P1a and P1b exhibit diminished telomerase activity as well as the addition of abnormal product patterns consistent with P1b's role as the human telomerase template boundary element (36,39). Chemoenzymatic footprinting of helix P1 supports the formation of P1b, though it is less consistent with P1a formation due to the susceptibility of C19-G25 to chemical modification and S1 nuclease cleavage (Figure 4.9) (35). Further examination of this footprinting data appears to be in conflict with the authors conclusions, as the nucleotides in question only exhibit weak sensitivity to modification and cleavage and the remainder of the P1 nucleotides G26-G30 are sensitive to RNase V1 suggesting that the conclusions are actually overstating the data (35).

SHAPE data of the P1 helix generally agrees well with the phylogenetic structure for the construction of helix P1a with the exception of the end of the helix where G30 is highly reactive and bulged and G31 base pairs with C196 (Figure 4.4). This is opposite of what is seen in the chemical footprinting that shows that G30 is weakly modified whereas G31 is highly modified (Figure 4.9) (35). SHAPE demonstrates that J1 loop shrinks with only U194 and U195 bulging due to their high reactivities, and moves the terminal G•U wobble pair from the end of P1a to the beginning of P1b, lengthening the template boundary element helix by one base pair (Figure 4.4). The G•U beginning P1b explains the increased flexibility for C192 due to the nature of the wobble pair (131). The bulged C191 increases the flexibility of nearby nucleotides G439 and C441; this

could be due to G440 invoking a kink or turn in the helix or preventing normal base stacking interactions due to its bulged nature.

One of the key features of the telomerase RNAs, is the formation of the pseudoknot tertiary RNA structure that is required for nucleotide addition and processivity. The stems and loops that make up the core pseudoknot structure have high sequence and structural conservation, though RNA sequences that connect these regions are only structurally conserved and even the extent of that conservation is varied among species at best (48). While NMR, FRET, and SHAPE analysis have proven that the hTR pseudoknot structure does in fact form in isolation, mutational studies in hTR and FRET of tTR in complex with tTERT have shown data consistent with the formation of the tertiary pseudoknot structure (37,40,44,137). To date there is no data available to demonstrate that full length hTR forms the pseudoknot in solution (*in vitro*) as neither chemoenzymatic footprinting nor FRET data show conclusive evidence of its formation (35,138). In agreement with these conclusions, SHAPE analysis of full length hTR also does not support the formation of the pseudoknot structure but instead suggests that the core domain forms an elongated helix much like the DC pseudoknot construct (Figure 3.4C). This data suggests that binding of hTERT or other accessory proteins to hTR while forming the active telomerase complex provides key contacts that allow for restructuring and stabilization of the "misfolded" hTR and directs proper formation of the tertiary triple helix from the elongated hairpin conformation. Studies have also shown that the amount of hTR in a cell greatly outnumbers the amount of hTERT (on an order of magnitude in the thousands) due to hTERT transcriptional repression (139). Perhaps hTR is sequestered in the cell in a tightly helical structure to prevent interactions with

other proteins and RNAs or to protect nucleotides that are involved in catalytic activity from being susceptible to RNA modification or degradation.

The adoption of this elongated hairpin structure for the core domain has major structural effects on the template similar to changes seen for the template region in *Tetrahymena thermophila* telomerase (Figure 2.1). Phylogenetic analysis of hTR identified the template region, nucleotides U37-G63, as being a long single stranded loop region that connected the template boundary element (P1b) to the first helical region of the proposed pseudoknot (P2a.1). With the disruption of the pseudoknot, sequences of hTR that were predicted to be in loop region J2a/3 and the 3' region of the P3 helix that are interacting in the pseudoknot, now base pair forming helices with the formerly single-stranded template and U-rich linking regions (Figure 4.4). These interactions form the augmented P1 helix that includes a new J1 internal loop structure with a mini-helix P1.1 in its 3' loop structure, the P1c helix, and a large internal loop J1/2 at the end; flexibilities of nucleotides involved in forming these structures show agreement with previous footprinting studies of hTR (Figure 4.9) (35).

Though they appear to be moderately modified by chemical reagents, nucleotides U42-G45 and A167 are all cleaved by RNase V1, indicating their involvement in the formation of base pairs and these nucleotides form base pairs in the extended P1c of the SHAPE derived structure (Figure 4.9) (35). Nucleotides G178-C180 and U184-C186 of the supposed P3 helix demonstrate only slight DMS, CMCT, and kethoxyl modification while A181 and G182 are much more reactive; this pattern of nucleotide flexibility is indicative of stem loop structure and supports the arrangement assumed by P1.1 (Figure 4.9) (35). Chemoenzymatic analysis of the nucleotides in the J1/2 loop is consistent with

SHAPE data, with A54-C56, G153-U154, A157-U195 all displaying moderate to high chemical and SHAPE reactivity and nucleotides C50-U53 and C156 being protected, though there are some discrepancies among the remaining nucleotides (Figure 4.9) (35). Interestingly, nucleotides C50-U53 of the template region, appear to be protected in both chemoenzymatic and SHAPE footprinting, though in both cases specific nucleotides that base pair with these have not been identified, suggesting that either these nucleotides form a stable RNA motif themselves or participate in a long range RNA interaction such as the G-pseudoknot (Figures 4.4, 4.8, and 4.9). As the chemoenzymatic study did not attempt to propose a new structure for hTR in solution and there is much resemblance between the SHAPE and chemoenzymatic data, it is very likely that the secondary structure developed from SHAPE analysis is a better representation of naked hTR.

P2a.1, P2a, and P2b compose the helical elements of the core domain except for P1 and P3. Since there is no pseudoknot structure in solution for hTR, there is no P3 in the SHAPE-derived secondary structure of hTR (Figure 4.4). While P2a, J2a/b, and P2b are structurally conserved phylogenetically, the P2a.1 helix and J2a.1 loop are mammalian specific extensions of P2a whose phylogenetic conservation is much less stringent than that of P2a and P2b (36,128). Studies have suggested that the P2a.1 helix is longer than predicted in the phylogenetic structure, with mutational analysis predicting base pairing of nucleotides A62-G63 with U146-C145 (36). SHAPE data agrees with this base pairing and actually suggests an even longer P2a.1 helix with nucleotides G60-G63 forming base pairs with C148-C145 making P2a.1 a 14 base pair helix that's divided by two single nucleotides bulges and a two nucleotide bulge, though chemoenzymatic analysis does not concur (Figure 4.4 and 4.9) (35). SHAPE data also suggests a smaller

internal loop J2a.1/2a with base pairs forming across the loop. These data correlate well with NMR data that shows that this asymmetrical internal loop does not induce a bend or significant dynamics between helices P2a.1 and P2a (Figure 4.4) (128). In comparison, increasing the size of the internal loop, as occurs in the C72G mutation that causes aplastic anemia, increases its flexibility and has led to a deficiency in telomerase activity (128).

Comparison of the P2a, J2a/b, and the beginning of the P2b helical portions of the phylogenetic structure with the secondary structure predicted by SHAPE show that these regions form exactly the same. SHAPE data identifies the J2a/b nucleotides as moderately to highly reactive demonstrating their single stranded nature and high flexibility and this data is corroborated by chemical probing (Figures 4.4 and 4.9) (35). The data also supports recent NMR studies that show that the five-nucleotide J2a/b loop adopts an S conformation that induces a bend with almost no twist between helices P2a and P2b that allows for proper positioning of the template in the active site and orienting hTR for optimal telomerase activity and processivity (128). The bend between the helices disrupts the stacking interactions weakening the stabilities of neighboring base pairs U83-A126 and C89-G125 in turn justifying the increased SHAPE reactivities associated with these nucleotides (Figure 4.4) (128). These NMR studies examine solution structures of these P2 helical segments assuming that the tertiary pseudoknot is actually forming, however, SHAPE data accounts that the pseudoknot does not form in solution which begs us to wonder how structurally relevant this model is in the context of the actual *in vitro* structure. The J2a/b loop may not elicit the same conformational bend

between the helices if the protein were present, or it is possible that the sterics of the large protein could potentially clash with these predicted helical arrangements.

As it has been mentioned earlier, SHAPE analysis supports the formation of the elongated hairpin conformation that the pseudoknot domain has been shown to in equilibrium with and that the DC pseudoknot has been shown to assume based on SHAPE, NMR, and thermal denaturation studies (Figure 3.4C) (42,125). Examination of the chemical probing data of the extended P2b helix nucleotides U99-C104 and U115-G110 demonstrate that the J2a/3 loop nucleotides are protected and the 5' P3 helix nucleotides are highly flexible due to their modification suggesting that the elongated helix does not form *in vitro* (Figure 4.9) (35). However, this representation of the chemical modification data neglects to take into consideration that nucleotides in the extended base pairings form Hoogstein base pairs, and the N3 position of bases that participate in Hoogstein interactions do not form hydrogren bonds, leaving them open to chemical modification, meaning that despite its representation, chemical probing of hTR also agrees with SHAPE that the elongated hairpin formation exists in solution (111). Interestingly, recent FRET data of hTR in complex with hTERT demonstrated that after protein binding, interconversion between a pseudoknot and hairpin conformation does not occur (137). This suggests that the lengthened hairpin conformation is only assumed during when protein free, and during catalysis protein binding and association helps the RNA maintain the catalytically relevant pseudoknot structure.

SHAPE analysis of the 3' section of hTR, which includes the hypervariable region, CR4-CR5 domain, Box H/ACA domains, and CR7 domain displays the RNA as much more structured than originally presented in the phylogenetic structure, though the

same general architecture is maintained as this region of the RNA is structurally conserved throughout vertebrates (48). Differences between the SHAPE inferred structure and the commonly accepted structure concentrate on the condensing of internal loop structures due to the creation of Watson Crick base pairs and G•U wobble pairs between the 5' and 3' strands with nucleotides with low SHAPE reactivities (Figure 4.4). Nucleotides G219-C223 form the new P4.0 helix with G353-U357 dividing the 5' pocket into two smaller one-sided loop regions; only nucleotides with moderate to high SHAPE reactivity remaining single stranded (Figure 4.4). These same base pairs were proposed to form in chemoenzymatic evaluation of hTR *in vitro* and are shown to form in similar investigation of hTR *in vivo* (Figure 4.10) (35). Most likely, similar conclusions about the formation of the two separated loop regions were not drawn from previous footprinting studies due to the pauses of reverse transcription occurring at both nucleotides C351 and C361, problems that did not encumber SHAPE footprinting (35). Similar consolidation occurred for both the J4.1/4.2 and J4/5 loops by formation of base pairs G233-C338 and A241-U327 respectively, both of which are supported by chemical modification data (Figure 4.10) (35).

As we discussed in Chapter 3, the CR4-CR5 domain of hTR is indispensable for telomerase catalytic activity and for its proposed hTERT binding site (19,47). Of specific importance is the P6.1 helix whose highly conserved sequence and stem loop structure in vertebrates directly contributes to nucleotide addition and processivity in the active site (49,50). Though an isolated domain of this exact RNA sequence and domain was examined in Chapter 3, the structural data and analysis in the scope of the full length hTR presents varied SHAPE reactivities and a different structure (Figures 3.6B and 4.4). This

demonstrates the principal dilemma presented by RNA structures and models that are developed from interrogation of smaller portions and then pieced together, such as with the recent NMR structures capturing the P2 helix (128); in the scope of a larger RNA, specific helices or looped regions can form different architectures than when they are isolated. In keeping with the theme of internal loop consolidation, the P5 helix in hTR has been elongated due to base pairing of minimally and slightly reactive nucleotides that once were part of J5/6a, J6.1/5, and the 3' strand of helix P6.1. P5 can now be divided into two stacked helices, P5a and P5b, separated by bulges C247 and C320. Ironically, neither of these nucleotides is actually reactive to SHAPE modification suggesting that they may constitute a non-canonical C•C base pair which like the A-U base pair can also make two hydrogen bonds or that RNAstructure forces these two nucleotides to bugle as there are no nearby pairing partners (Figure 4.4) (111,121). As nucleotides from the 3' strand of helix P6.1 are actually part of P5b, in the context of full length hTR, the P6.1 helix does not form and instead comprises a new J6a/5a single stranded loop, all of whose nucleotides are moderately to highly reactive to SHAPE modification (Figure 4.4). Disruption of P6.1 suggests that like the tertiary pseudoknot structure in the core domain, hTERT binding to the CR4-CR5 domain elicits reorganization and stabilization of an essential for catalytic activity "misfolded" hTR structural motif.

Unlike the P6.1 helix of the CR4-CR5 domain, structural analysis of the domain RNAs from Chapter 3 matches perfectly with full length hTR structure for the P6a, P6b, L6 (Figures 3.6B, 4.2, and 4.4), P7a, P7b, P8a, P8b, and L8 (Figure 3.7B, 4.2, and 4.4) helix and loop regions. The H/ACA domains are single stranded (and very reactive to SHAPE modification) exactly as they are predicted in the phylogenetically predicted

structure of hTR (Figure 4.4). Similar results were seen for the box H region in chemical footprinting studies; these studies were unable to determine the structure of box ACA due to primer binding, stressing the importance of the addition of the 3' linker and additional primer binding site incorporated into the 3' end of SHAPE analyzed RNAs (Figure 4.10) (35,85).

After determining the full length structure of hTR in solution, we decided to apply this structural information by investigating the ability of small molecules to bind to hTR and inhibit telomerase activity. As mentioned in the introduction, telomerase is up-regulated in greater than 85% of cancers and confers limitless proliferation upon these cells making them tumorigenic (1,4,6). Therefore, the development of anti-telomerase drugs could take advantage of telomerase as a universal cancer target. Over time, many different types of small molecules that inhibit telomerase function have been discovered including reverse transcriptase inhibitors, G-quadraplex stabilizing compounds, Hsp90 inhibitors, and natural products (Soares, Sekaran, Jarstfer, *unpublished manuscript*) (113,140). However, to date, most telomerase inhibitors that target hTR have been nucleic acid based antagonists (the Geron compund) and catalytic RNAs like ribozymes (141,142). For example, our lab identified DNA oligonucleotides with 2'-O-methyl backbone modifications that were able to prevent telomerase assemblage by inhibiting the binding of the CR4-CR5 domain of hTR to hTERT by 92% (113). Only in a few instances have small molecules been shown to inhibit telomerase activity by binding to hTR (133).

One class of compounds, aminoglycosides, has emerged in the past few years as chemical motifs that are capable of RNA recognition such as being able to discriminate

between A-form and B-form RNA helices and demonstrate preferential binding in or near RNA loop or bulges nucleotides (143). Additionally, NMR experiments have revealed that aminoglycoside binding can elicit local conformational changes in RNA structure (144). Studies of aminoglycoside-quinacridine conjugates of tobramycin have been shown to bind the P6.1 helix through thermal denaturation studies and gel shift assays (143). Similarly, studies in our lab have shown that aminoglycosides that bind to hTR can inhibit telomerase activity in pre-assemblage assays up to 75% (133). Therefore, when the Disney lab identified that aminoglycosides kanamycin-6" Az-FITC (Kan-6") and neomycin-5-O-(AzEt)-FITC (Nea-5-O) would bind to hTR at the CR4-CR5 domain and pseudoknot domains respectively, we decided to investigate their potential to inhibit telomerase activity (132). Pre-assemblage telomerase activity assays revealed marked telomerase inhibition by Nea 5-O and Kan 6"" while the same compounds had no effect on activity in the post-assemblage, confirming that Nea 5-O and Kan 6" prevent proper telomerase complex formation by hTR and hTERT (Figure 4.5).

SHAPE analysis of the pseudoknot domain of hTR in the presence of Nea 5-O revealed major structural vacillations and marked disruption of the elongated hairpin structure (Figured 4.6A). For example, nucleotides within internal and end loops appear to become more structured, while helices become more flexible and likely disrupted. This suggests that either Nea 5-O binds quite promiscuously throughout the elongated helices that comprise the pseudoknot domain or that binding in a single or multiple locations causes structural unraveling and the promotion of new RNA conformations throughout. Based on these results, and our inability to identify one or two specific binding locations within the core domain that may contribute to the prevention of

telomerase assemblage, we decided not to pursue further experiments with Nea 5-O at this time.

In the presence of Kan 6", the overall structure of the CR4-CR5 domain of hTR remained organized and changes in nucleotide flexibilities were localized to three specific RNA motifs: J6/5 and flanking helices, the P6b and J6 helix and internal loop, and stacked helices P4.2-P4.1 (Figure 4.6B). Perturbations in structure were also seen in the 3' strand of the P4.3-4.1 helices; unfortunately we do not have data for the 5' strand of these helices as the hTR5 primer is unable to resolve data in this region. Interestingly, the nucleotides in these helices with the most dramatic changes in flexibility from Kan 6" binding, G336, G339, A340, G345, and U350, all demonstrate SHAPE reactivity when hTR is in solution and, with the exception of G336, are found as bulged or looped nucleotide motifs (Figure 4.6B). All of the binding locations agree well with previous studies that identified the preference of aminoglycosides to bind in or near looped structures (143). After identifying Kan 6" binding sites in the CR4-CR5 domain, compound concentrations were varied over a range to identify nucleotides whose flexibility increased or decreased in response to increasing amounts of Kan 6". Once again, nucleotides that exhibited changes in SHAPE reactivity coincided with nucleotides in bulged or loop regions, specifically L6, J6, U312, and J5a/4.2 (Figure 4.7). Concentration-dependent structural changes were also noticed for U316 and the helix it is stacked on top of (Figure 4.7). U316 is shown to make a wobble pair with G251, and G•U interactions are known to be more flexible than Watson Crick base pairs and induce flexibility in its neighbors, making it not so surprising that it may be a binding site for the aminoglycoside Kan 6" (131). Previous studies have proposed that the J6 may contain an

hTERT binding site due to the formation of a solvent accessible tunnel from the structural orientation of the nucleotides in the this loop region and that deletion of this loop completely abolished the ability of the CR4-CR5 domain to interact with hTERT and to activate telomerase function (46,47). Nucleotides in this internal loop region have been shown to elicit changes in flexibility due to binding identifying it as a Kan 6" binding site. The combination of these conclusions proposes that inhibition of telomerase assemblage by Kan 6" is due to the compound blocking an essential hTERT binding site in the J6 internal loop of the CR4-CR5 domain of hTR.

D. Materials and Methods

1. Oligonucleotides and chemicals

Oligonucleotides were purchased from Integrated DNA Technologies (IDT, Coralville, Iowa). The oligonucleotides were then purified via denaturing polyacrylamide gel electrophoresis (PAGE) using the modified "crush and soak" method (112,113). Following electrophoresis in a 12% denaturing gel, oligonucleotides were observed via ultraviolet (UV) light, cut out with a razor blade, crushed through a 3 mL sterile plastic syringe, and eluted from the gel in 1×TEN buffer (10 mM Tris pH 7.5, 1 mM EDTA, 250 mM NaCl) for one hour at room temperature. After centrifugation, the supernatents were collected and passed through a 0.22 micron MCE membrane syringe filter (FisherBrand) to remove remaining gel particulate. Oligonucleotides were then ethanol precipitated with final concentration 0.6 M ammonium chloride, concentrated, and resuspended in TE (10 mM Tris-Cl pH 8.0, 1 mM EDTA) or DEPC-treated water (Ambion, Inc., Austin, Texas) to a stock concentration of 1 mM. The concentration of further dilutions was

determined by UV absorbance at 260 nm using the molar extinction coefficient supplied by the manufacturer.

1-methyl-7-nitroisatoic anhydride (1M7) was a gift from Dr. Kevin Weeks (UNC Chapel Hill, NC, USA). This compound was made freshly at 50 mM stocks in DMSO ACS reagent grade (MP Biomedicals, Solon, Ohio). Aminoglycosides kanamycin-6" Az-FITC (Kan-6") at 630 μM and neomycin-5-O-(AzEt)-FITC (Nea-5-O) at 565 μM were a gift from Matt Disney (Scripps Research Institute, Jupiter, FL). Kan-6" was further diluted to make a 20 μM stock in DMSO ACS reagent grade.

2. hTR and hTERT expression plasmids

The puc19 phTR+HH expression plasmid was a gift from Dr. Jamie Sperger (University of Colorado, Boulder). This plasmid was designed for *in vitro* transcription of human telomerase RNA and contains a self-cleaving hammerhead ribozyme that generates the natural 5' end inserted in front of the hTR gene. The hTERT plasmid (pCl-neo-hTERT) was a gift from Dr. Lorel Colgin (Children's Medical Research Institute, Westmead, Australia) and CAMBIA (Canberra Australia) (145). The hTERT-containing insert from this plasmid was subcloned into the *Eco*RI and *Sal*I sites of the T7-tag-containing plasmid pET-28c (Novagen) to give the construct pET-28c-hTERT.

3. PCR construction of dsDNA hTR construct

Table 4.1: Primers for cloning of hTR

hTR Primer	Sequence (5'-3')
hTR Forward	GCGCTAATACGACTCACTATAGGGTTGCGGA GGGTGGGCCTG
hTR Reverse	GAACCGGACCGAAGCCCGATTTGGATCCGGC GAACCGGATCGAGCATGTGTGAGCCGAGTC

In order to make the hTR RNA, dsDNA templates were generated by PCR using primers that overlapped at the 5' and 3' sequences of the hTR gene. The forward primers contain a T7 promoter before the initial hTR sequence while the reverse primer had the 3' linker and conRT binding site incorporated after the terminal hTR sequence (85). Primers in the presence of the puc19 phTR+HH expression plasmid underwent the following PCR cycles: 1. 95 °C for 30 SECONDS enzyme hot-start; 2. 30 cycles of 95 °C for 30 seconds, 55 °C for 30 seconds, 72 °C for 1 minute to yield a full-length double-stranded product; and 3. 72 °C for 5 minutes to complete any unfinished ends. Upon completion, PCR products were PCI extracted, ethanol precipitated, concentrated, resuspended in TE, and nanodropped to determine concentration.

4. *In vitro* transcription and purification of flanked hTR

Flanked hTR was made via in vitro transcription with the T7 Megascript Kit (Life Technologies, Invitrogen). The *in vitro* transcription reaction was carried out overnight at 37 °C with one microgram dsDNA template, 1x Reaction Buffer, ATP solution, CTP solution, GTP solution, UTP solution, and 2 μL Enzyme Mix. The next morning, the reaction incubated with 1 μL TURBO DNase 1(2U/μL) for 15 minutes at 37 °C and then resuspended in an equal volume SHAPE denaturing loading buffer (80% formamide, 0.5×TBE, 4 mM EDTA, 0.01% bromophenol blue, 0.01% xylene cyanol) and purified in a 5% denaturing polyacrylamide gel. The RNA was then recovered by the modified "crush and soak" method as explained above; the only differences being that the RNAs are ethanol precipitated with final concentration of 0.3 M of sodium acetate and final RNA concentrations were determined by UV spectroscopy

5. SHAPE 1M7 hit reactions

One picomole of RNA (1 μL of 1 μM RNA), DEPC-treated water, and 5× hTR Hit Buffer (250 mM Hepes Buffer Solution(Gibco), 250 mM RNase-free KCl (Ambion), 5 mM RNase-free $MgCl_2$ (Ambion)) diluted to 1× were combined in a microcentrifuge tube and allowed to equilibrate at 37 °C for a couple of minutes to allow RNA to adopt its correct conformation. The RNA was then treated with either 50 mM 1M7 to a final concentration of 5 mM, or DMSO to a final concentration of 10% as a control and incubated at 37 °C for 70 seconds. The hit reaction was quenched with RNase-free NaCl (Ambion) to a final concentration of 200 mM and precipitated with 200 μg/mL RNase-free glycogen (Ambion) as a counter ion. Following centrifugation, the hit RNA was washed once with 70% ethanol, speed vacuumed till dry, and then resuspended in 10 μL of RNase-free TE buffer pH 8.0 (Ambion).

6. Superscript III primer extension

Table 4.2: SHAPE Primers for mapping hTR

SHAPE Primers	Sequence (5'-3')
hTR1	CTTCCCGCATCCGCGGC
hTR2	CTCGTTTGTTTTTTACAGTC
hTR3	GGGCCCCTGGACGCCGCC
hTR4	GACGGTGGCGCTTCTCAACCC
hTR5	GTCCGGCGTCCTTCTCC
hTR6	GTCCTGAGCCGAGTGTGTACG
conRT	GAACCGGACCGAAGCCCG

Due to the size of hTR, 451 nucleotides, seven primers (in separate reactions, not all together) were used to in order to map SHAPE chemistry adduct formation. Half a picomole of 1M7 modified RNA (5 μL) was mapped by annealing 1 μL of a 5'-[^{32}P]-

labeled primer to hTR via the following thermocycler conditions: 95 °C for 1 min, 65 °C for 6 min, and 35 °C for 10 min. Following a 5 minute incubation on ice, RNAs were supplemented with 5× First Strand Buffer (250 mM Tris-Cl pH 8.3, 375 mM KCl, 15 mM MgCl2) (Invitrogen), 0.1 M DTT to a final concentration of 5 mM, and 10 mM dNTPs to a final concentration of 1.25 mM. These contents were then heated at a temperature of 50 °C on an Isotemp 125D (Fisher Scientific) for one minute. One microliter of Superscript II reverse transcriptase (100 units, Invitrogen) was immediately added to the reaction mixture, mixed by gentle tapping, and allowed to extend for exactly 4 minutes at 50 °C. The reaction was quenched by the addition of 400 mM NaOH and then heated at 95 °C for 5 minutes in order to degrade the hit RNAs. This was then neutralized by the addition of 400 mM HCl and radiolabeled cDNAs were ethanol precipitated with 5 M ammonium acetate and 200 μg/mL glycogen and then resuspended in 6 μL SHAPE loading buffer. Dideoxycytidine (ddC) sequencing ladders were generated by the addition of 1.25 mM ddCTP (GE Healthcare) to the reverse transcription reaction of 3.5 picomoles of unmodified RNAs.

7. Sequencing gel electrophoresis

The radiolabeled cDNAs created by the reverse transcription of NMIA/1M7/DMSO hit RNAs were resolved on a 35 cm x 43 cm x 0.4 mm 8% denaturing polyacrylamide sequencing gel (29:1 acrylamide; bisacrylamide/7M urea, 90 mM Tris/borate, 2 mM EDTA) run at approximately 2000 volts in 0.5× TBE for about an hour and a half. The gel was transferred to and dried on Whatman Chromatography paper (GE Healthcare) for 1 hour at 80 °C on a Model 583 Gel Dryer (Biorad) and then

exposed on a Phosphor Screen (Amersham Biosciences) overnight. Gels were imaged on a Storm 860 Phosphor Imager (Molecular Dynamics, Sunnyvale, CA) and visualized on ImageQuant TL (Amersham Biosciences).

8. SAFA data analysis and normalization

Individual band intensities of 1M7 and DMSO lanes were quantified using the program SAFA (88). SAFA allows the straightening of curved gels (gel recitification) and uses Lorentzian curve integration to determine band intensities with a high degree of accuracy. Full length RNA bands were quantified by employing the analysis method in ImageQuant TL. Radioactivity across the lanes of an individual experiment was then equalized to the level of radioactivity in the 1M7 lane. Band intensity was calculated by subtracting the value of the DMSO band, control reaction that corresponds to background activity, from the value of the 1M7 band giving the corrected band intensity. Corrected band intensities were then ranked in descending order and the top 2% of values were thrown out. The average of the next 8% of values was taken and this value represented the normalization factor. All of the 1M7 corrected densities were divided by the normalization factor and then multiplied by 100 to give the normalized values. Normalized values from multiple replicates were averaged and plotted in figures with standard error bars shown.

9. *In vitro* transcription and purification of hTR

hTR was made by *in vitro* transcription from the phTR+HH plasmid with the T7 Ampliscribe Transcription Kit (Epicentre Biotechnologies, Madison, WI). Fifteen μg of

phTR+HH was linearized in a Fok-1 digestion reaction. Work up of the digestion products included extraction with phenol/chloroform/isoamyl alcohol, ethanol precipitation, and resuspension in a suitable volume of TE (10 mM Tris-HCl, pH 7.5 and 1 mM EDTA). The *in vitro* transcription reaction was carried out overnight at 37 °C with the linearized DNA, 1x Ampliscribe T7 Reaction Buffer, 7.5 mM ATP, 7.5 mM CTP, 7.5 mM GTP, 7.5 mM UTP, 10 mM DTT, ScriptGuard RNase Inhibitor, and 10 μL Ampliscribe T7 Enzyme Solution.

Cleavage by the hammerhead ribozyme was initiated by the addition of $MgCl_2$ to a final concentration of 12 mM and incubated at 45 °C for one hour. The entire reaction was then treated with 5 μL RNase-Free DNase 1 (Epicentre, 1 MBU/μL) for 20 min at 37 °C and then extracted with phenol/chloroform/ isoamyl alcohol and ethanol precipitated in the presence of 0.3 M NaOAc. The RNA was resuspended in denaturing loading buffer (7 M urea/10% glycerol/1×TBE) and purified on a 4% denaturing polyacrylamide gel. The RNA was then recovered by the modified "crush and soak" method as explained above; the only differences being that the RNAs are ethanol precipitated with final concentration of 0.3 M of sodium acetate and final RNA concentrations were determined by nanodropping and RNA molecular mass.

10. Telomerase Assemblage Assay

Recombinant hTERT was transcribed and translated from pET28c-hTERT using the TNT® Coupled Reticulocyte Lysate Systems kit (Promega) following the manufacturer's instructions. Telomerase assemblage assays were initiated by combining 10 μL (~50 fmol) of hTERT from the RRL reaction, 212.5 ng of *in vitro* transcribed hTR,

and the appropriate amount of DMSO or inhibitor to a final volume of 15 μL followed by incubation at 30 °C for 90 min as previously described in detail (113,133). Inhibition studies contained 50 μM concentration of inhibitor. After assemblage, telomerase-catalyzed primer extension was carried out by adding 50 mM Tris-HCl, pH 8.0, 50 mM KCl, 1 mM $MgCl_2$, 5 mM β-mercaptoethanol, 1 mM spermidine, 1 μM human telomere primer (5'-TTAGGGTTAGGGTTAGGG), 0.5 mM dATP, 0.5 mM dTTP, 2.9 μM dGTP, and 0.33 μM [α-^{32}P]-dGTP (3000 Ci/mmol, 10 μCi/μL; Perkin Elmer) to the 15 μL of reconstituted telomerase. Primer extension was carried out at 30 °C for 90 minutes. The inhibitor concentration was maintained during both the assembly reaction and extension reaction.

After primer extension, reaction volume was increased by the addition of water; a [^{32}P]-labeled loading control (114 nucleotide, 5'-end labeled ssDNA oligonucleotide, 1000 cpm per reaction) was also added. Telomerase assay primer extension products were extracted once with phenol/chloroform/isoamyl alcohol (25:24:1) and twice with chloroform/isoamyl alcohol (24:1) and then ethanol precipitated with 2.5 volumes absolute ethanol, 0.6 M ammonium acetate, and 75 μg/mL glycogen. After precipitating at -20 °C for 15 minutes, telomerase assay products were centrifuged at 13,000 x g at 4 °C for 22 minutes, washed with 1 volume of 70% ethanol, and speed vacuumed till dry. Pellets were resuspended in 6 μL SHAPE loading buffer and heated at 95 °C for 6 minutes and resolved on a pre-warmed, 20 cm x 20 cm x 0.4 mm thick 10% denaturing polyacrylamide gel. The gel was run in 0.5× TBE at 800 volts for approximately one hour. The gel was transferred, dried, exposed, and visualized similar to the sequencing gel mentioned above. Summing the intensities of all bands in each sample and

normalizing the values to the loading control determined overall telomerase activity of the reaction.

11. Nea 5-O and Kan-6" binding to hTR studies

To identify possible hTR binding sites of Nea 5-O (in the pseudoknot domain) and Kan-6" (in the CR4-CR5 domain), one picomole of hTR (1 μL of 1 μM hTR) was incubated DEPC-treated water and each of the compounds at varied concentrations at 37 °C for 5 minutes. Following this incubation, each reaction was supplemented with DEPC-treated water, 5× hTR Hit diluted to 1×, 50 mM 1M7 to a final concentration of 5 mM, or DMSO to a final concentration of 10% as a control and incubated at 37 °C for 70 seconds to complete the SHAPE hit protocol mentioned above. The compound concentration was maintained during both the initial incubation and the SHAPE "hit" reaction. The remainder of the SHAPE protocol was completed for these samples and normalized values were compared with normalized values from the SHAPE of the in-solution hTR to determine possible compound binding sites based on changes in nucleotide flexibilities.

REFRENCES

1. Blasco, M.A. (2007) Telomere length, stem cells and aging. *Nat Chem Biol*, **3**, 640-649.

2. Aubert, G. and Lansdorp, P.M. (2008) Telomeres and aging. *Physiol Rev*, **88**, 557-579.

3. Aviv, A. (2004) Telomeres and human aging: facts and fibs. *Sci Aging Knowledge Environ*, **2004**, pe43.

4. Masutomi, K. and Hahn, W.C. (2003) Telomerase and tumorigenesis. *Cancer Lett*, **194**, 163-172.

5. Smith, R.A., Cokkinides, V. and Brawley, O.W. (2009) Cancer screening in the United States, 2009: A review of current American Cancer Society guidelines and issues in cancer screening. *CA Cancer J Clin*, **59**, 27-41.

6. Winer, E., Gralow, J., Diller, L., Karlan, B., Loehrer, P., Pierce, L., Demetri, G., Ganz, P., Kramer, B., Kris, M. *et al.* (2008) Clinical Cancer Advances 2008: Major Research Advances in Cancer Treatment, Prevention, and Screening--A Report From the American Society of Clinical Oncology. *J Clin Oncol*.

7. Gilson, E. and Geli, V. (2007) How telomeres are replicated. *Nat Rev Mol Cell Biol*, **8**, 825-838.

8. Chan, S.R. and Blackburn, E.H. (2004) Telomeres and telomerase. *Philos Trans R Soc Lond B Biol Sci*, **359**, 109-121.

9. Bianchi, A. and Shore, D. (2008) How telomerase reaches its end: mechanism of telomerase regulation by the telomeric complex. *Mol Cell*, **31**, 153-165.

10. Bailey, S.M. and Murnane, J.P. (2006) Telomeres, chromosome instability and cancer. *Nucleic Acids Res*, **34**, 2408-2417.

11. Blackburn, E.H. (2000) The end of the (DNA) line. *Nat Struct Biol*, **7**, 847-850.

12. Pendino, F., Tarkanyi, I., Dudognon, C., Hillion, J., Lanotte, M., Aradi, J. and Segal-Bendirdjian, E. (2006) Telomeres and telomerase: Pharmacological targets for new anticancer strategies? *Curr Cancer Drug Targets*, **6**, 147-180.

13. Blackburn, E.H. (2005) Telomeres and telomerase: their mechanisms of action and the effects of altering their functions. *FEBS Lett*, **579**, 859-862.

14. Tomlinson, R.L., Ziegler, T.D., Supakorndej, T., Terns, R.M. and Terns, M.P. (2006) Cell cycle-regulated trafficking of human telomerase to telomeres. *Mol Biol Cell*, **17**, 955-965.

15. Collins, K. (2006) The biogenesis and regulation of telomerase holoenzymes. *Nat Rev Mol Cell Biol*, **7**, 484-494.

16. Collins, K. (2008) Physiological assembly and activity of human telomerase complexes. *Mech Ageing Dev*, **129**, 91-98.

17. Jady, B.E., Richard, P., Bertrand, E. and Kiss, T. (2006) Cell cycle-dependent recruitment of telomerase RNA and Cajal bodies to human telomeres. *Mol Biol Cell*, **17**, 944-954.

18. Venteicher, A.S., Abreu, E.B., Meng, Z., McCann, K.E., Terns, R.M., Veenstra, T.D., Terns, M.P. and Artandi, S.E. (2009) A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis. *Science*, **323**, 644-648.

19. Sekaran, V.G., Soares, J. and Jarstfer, M.B. (2010) Structures of telomerase subunits provide functional insights. *Biochim Biophys Acta*, **1804**, 1190-1201.

20. Kelleher, C., Teixeira, M.T., Forstemann, K. and Lingner, J. (2002) Telomerase: biochemical considerations for enzyme and substrate. *Trends Biochem Sci*, **27**, 572-579.

21. Autexier, C. and Lue, N.F. (2006) The structure and function of telomerase reverse transcriptase. *Annu Rev Biochem*, **75**, 493-517.

22. Moriarty, T.J., Huard, S., Dupuis, S. and Autexier, C. (2002) Functional multimerization of human telomerase requires an RNA interaction domain in the N terminus of the catalytic subunit. *Mol Cell Biol*, **22**, 1253-1265.

23. Huard, S., Moriarty, T.J. and Autexier, C. (2003) The C terminus of the human telomerase reverse transcriptase is a determinant of enzyme processivity. *Nucleic Acids Res*, **31**, 4059-4070.

24. Cong, Y. and Shay, J.W. (2008) Actions of human telomerase beyond telomeres. *Cell Res*, **18**, 725-732.

25. Bollmann, F.M. (2008) The many faces of telomerase: emerging extratelomeric effects. *Bioessays*, **30**, 728-732.

26. Santos, J.H., Meyer, J.N. and Van Houten, B. (2006) Mitochondrial localization of telomerase as a determinant for hydrogen peroxide-induced mitochondrial DNA damage and apoptosis. *Hum Mol Genet*, **15**, 1757-1768.

27. Bosoy, D., Peng, Y., Mian, I.S. and Lue, N.F. (2003) Conserved N-terminal motifs of telomerase reverse transcriptase required for ribonucleoprotein assembly in vivo. *J Biol Chem*, **278**, 3882-3890.

28. Bachand, F. and Autexier, C. (2001) Functional regions of human telomerase reverse transcriptase and human telomerase RNA required for telomerase activity and RNA-protein interactions. *Mol Cell Biol*, **21**, 1888-1897.

29. Lai, C.K., Mitchell, J.R. and Collins, K. (2001) RNA binding domain of telomerase reverse transcriptase. *Mol Cell Biol*, **21**, 990-1000.

30. Harrington, L. (2003) Biochemical aspects of telomerase function. *Cancer Lett*, **194**, 139-154.

31. Chen, J.L. and Greider, C.W. (2004) Telomerase RNA structure and function: implications for dyskeratosis congenita. *Trends Biochem Sci*, **29**, 183-192.

32. Theimer, C.A. and Feigon, J. (2006) Structure and function of telomerase RNA. *Curr Opin Struct Biol*, **16**, 307-318.

33. Legassie, J.D. and Jarstfer, M.B. (2006) The unmasking of telomerase. *Structure*, **14**, 1603-1609.

34. Li, X., Nishizuka, H., Tsutsumi, K., Imai, Y., Kurihara, Y. and Uesugi, S. (2007) Structure, interactions and effects on activity of the 5'-terminal region of human telomerase RNA. *J Biochem*, **141**, 755-765.

35. Antal, M., Boros, E., Solymosy, F. and Kiss, T. (2002) Analysis of the structure of human telomerase RNA in vivo. *Nucleic Acids Res*, **30**, 912-920.

36. Ly, H., Blackburn, E.H. and Parslow, T.G. (2003) Comprehensive structure-function analysis of the core domain of human telomerase RNA. *Mol Cell Biol*, **23**, 6849-6856.

37. Qiao, F. and Cech, T.R. (2008) Triple-helix structure in telomerase RNA contributes to catalysis. *Nat Struct Mol Biol*, **15**, 634-640.

38. Lai, C.K., Miller, M.C. and Collins, K. (2003) Roles for RNA in telomerase nucleotide and repeat addition processivity. *Mol Cell*, **11**, 1673-1683.

39. Chen, J.L. and Greider, C.W. (2003) Template boundary definition in mammalian telomerase. *Genes Dev*, **17**, 2747-2752.

40. Kim, N.K., Zhang, Q., Zhou, J., Theimer, C.A., Peterson, R.D. and Feigon, J. (2008) Solution structure and dynamics of the wild-type pseudoknot of human telomerase RNA. *J Mol Biol*, **384**, 1249-1261.

41. Theimer, C.A., Blois, C.A. and Feigon, J. (2005) Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Mol Cell*, **17**, 671-682.

42. Theimer, C.A., Finger, L.D., Trantirek, L. and Feigon, J. (2003) Mutations linked to dyskeratosis congenita cause changes in the structural equilibrium in telomerase RNA. *Proc Natl Acad Sci U S A*, **100**, 449-454.

43. Ueda, C.T. and Roberts, R.W. (2004) Analysis of a long-range interaction between conserved domains of human telomerase RNA. *RNA*, **10**, 139-147.

44. Shefer, K., Brown, Y., Gorkovoy, V., Nussbaum, T., Ulyanov, N.B. and Tzfati, Y. (2007) A triple helix within a pseudoknot is a conserved and essential element of telomerase RNA. *Mol Cell Biol*, **27**, 2130-2143.

45. Moriarty, T.J., Marie-Egyptienne, D.T. and Autexier, C. (2004) Functional organization of repeat addition processivity and DNA synthesis determinants in the human telomerase multimer. *Mol Cell Biol*, **24**, 3720-3733.

46. Mitchell, J.R. and Collins, K. (2000) Human telomerase activation requires two independent interactions between telomerase RNA and telomerase reverse transcriptase. *Mol Cell*, **6**, 361-371.

47. Leeper, T.C. and Varani, G. (2005) The structure of an enzyme-activating fragment of human telomerase RNA. *RNA*, **11**, 394-403.

48. Chen, J.L., Blasco, M.A. and Greider, C.W. (2000) Secondary structure of vertebrate telomerase RNA. *Cell*, **100**, 503-514.

49. Leeper, T., Leulliot, N. and Varani, G. (2003) The solution structure of an essential stem-loop of human telomerase RNA. *Nucleic Acids Res*, **31**, 2614-2621.

50. Chen, J.L., Opperman, K.K. and Greider, C.W. (2002) A critical stem-loop structure in the CR4-CR5 domain of mammalian telomerase RNA. *Nucleic Acids Res*, **30**, 592-597.

51. Mitchell, J.R., Wood, E. and Collins, K. (1999) A telomerase component is defective in the human disease dyskeratosis congenita. *Nature*, **402**, 551-555.

52. Fu, D. and Collins, K. (2003) Distinct biogenesis pathways for human telomerase RNA and H/ACA small nucleolar RNAs. *Mol Cell*, **11**, 1361-1372.

53. Kiss, T. (2002) Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell*, **109**, 145-148.

54. Baker, D.L., Seyfried, N.T., Li, H., Orlando, R., Terns, R.M. and Terns, M.P. (2008) Determination of protein-RNA interaction sites in the Cbf5-H/ACA guide RNA complex by mass spectrometric protein footprinting. *Biochemistry*, **47**, 1500-1510.

55. Li, L. and Ye, K. (2006) Crystal structure of an H/ACA box ribonucleoprotein particle. *Nature*, **443**, 302-307.

56. Trahan, C. and Dragon, F. (2009) Dyskeratosis congenita mutations in the H/ACA domain of human telomerase RNA affect its assembly into a pre-RNP. *RNA*, **15**, 235-243.

57. Theimer, C.A., Jady, B.E., Chim, N., Richard, P., Breece, K.E., Kiss, T. and Feigon, J. (2007) Structural and functional characterization of human telomerase RNA processing and cajal body localization signals. *Mol Cell*, **27**, 869-881.

58. Ye, K. (2007) H/ACA guide RNAs, proteins and complexes. *Curr Opin Struct Biol*, **17**, 287-292.

59. Vulliamy, T.J. and Dokal, I. (2008) Dyskeratosis congenita: the diverse clinical presentation of mutations in the telomerase complex. *Biochimie*, **90**, 122-130.

60. Mason, P.J., Wilson, D.B. and Bessler, M. (2005) Dyskeratosis congenita -- a disease of dysfunctional telomere maintenance. *Curr Mol Med*, **5**, 159-170.

61. Marciniak, R.A., Johnson, F.B. and Guarente, L. (2000) Dyskeratosis congenita, telomeres and human ageing. *Trends Genet*, **16**, 193-195.

62. Bessler, M., Du, H.Y., Gu, B. and Mason, P.J. (2007) Dysfunctional telomeres and dyskeratosis congenita. *Haematologica*, **92**, 1009-1012.

63. Vulliamy, T., Marrone, A., Goldman, F., Dearlove, A., Bessler, M., Mason, P.J. and Dokal, I. (2001) The RNA component of telomerase is mutated in autosomal dominant dyskeratosis congenita. *Nature*, **413**, 432-435.

64. Marrone, A. and Dokal, I. (2004) Dyskeratosis congenita: molecular insights into telomerase function, ageing and cancer. *Expert Rev Mol Med*, **6**, 1-23.

65. Walne, A.J. and Dokal, I. (2008) Dyskeratosis Congenita: a historical perspective. *Mech Ageing Dev*, **129**, 48-59.

66. Cerone, M.A., Ward, R.J., Londono-Vallejo, J.A. and Autexier, C. (2005) Telomerase RNA mutated in autosomal dyskeratosis congenita reconstitutes a weakly active telomerase enzyme defective in telomere elongation. *Cell Cycle*, **4**, 585-589.

67. Marrone, A., Walne, A. and Dokal, I. (2005) Dyskeratosis congenita: telomerase, telomeres and anticipation. *Curr Opin Genet Dev*, **15**, 249-257.

68. Chu, C.Y. and Rana, T.M. (2007) Small RNAs: regulators and guardians of the genome. *J Cell Physiol*, **213**, 412-419.

69. Brosius, J. and Tiedge, H. (2004) RNomenclature. *RNA Biol*, **1**, 81-83.

70. Storz, G. (2002) An expanding universe of noncoding RNAs. *Science*, **296**, 1260-1263.

71. Szymanski, M., Barciszewska, M.Z., Erdmann, V.A. and Barciszewski, J. (2005) A new frontier for molecular medicine: noncoding RNAs. *Biochim Biophys Acta*, **1756**, 65-75.

72. Rana, T.M. (2007) Illuminating the silence: understanding the structure and function of small RNAs. *Nat Rev Mol Cell Biol*, **8**, 23-36.

73. Breaker, R.R. (2004) Natural and engineered nucleic acids as tools to explore biology. *Nature*, **432**, 838-845.

74. Storz, G., Altuvia, S. and Wassarman, K.M. (2005) An abundance of RNA regulators. *Annu Rev Biochem*, **74**, 199-217.

75. Bevilacqua, P.C., Cerrone-Szakal, A.L. and Siegfried, N.A. (2007) Insight into the functional versatility of RNA through model-making with applications to data fitting. *Q Rev Biophys*, **40**, 55-85.

76. Laederach, A. (2007) Informatics challenges in structured RNA. *Brief Bioinform*, **8**, 294-303.

77. Doudna, J.A. and Doherty, E.A. (1997) Emerging themes in RNA folding. *Fold Des*, **2**, R65-70.

78. Felden, B. (2007) RNA structure: experimental analysis. *Curr Opin Microbiol*, **10**, 286-291.

79. Holbrook, S.R. (2008) Structural principles from large RNAs. *Annu Rev Biophys*, **37**, 445-464.

80. Gavory, G., Symmons, M.F., Krishnan Ghosh, Y., Klenerman, D. and Balasubramanian, S. (2006) Structural analysis of the catalytic core of human telomerase RNA by FRET and molecular modeling. *Biochemistry*, **45**, 13304-13311.

81. Ke, A. and Doudna, J.A. (2004) Crystallization of RNA and RNA-protein complexes. *Methods*, **34**, 408-414.

82. Tzakos, A.G., Grace, C.R., Lukavsky, P.J. and Riek, R. (2006) NMR techniques for very large proteins and rnas in solution. *Annu Rev Biophys Biomol Struct*, **35**, 319-342.

83. Shen, L.X., Cai, Z. and Tinoco, I., Jr. (1995) RNA structure at high resolution. *FASEB J*, **9**, 1023-1033.

84. Ding, F., Sharma, S., Chalasani, P., Demidov, V.V., Broude, N.E. and Dokholyan, N.V. (2008) Ab initio RNA folding by discrete molecular dynamics: from structure prediction to folding mechanisms. *RNA*, **14**, 1164-1173.

85. Merino, E.J., Wilkinson, K.A., Coughlan, J.L. and Weeks, K.M. (2005) RNA structure analysis at single nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE). *J Am Chem Soc*, **127**, 4223-4231.

86. Wilkinson, K.A., Merino, E.J. and Weeks, K.M. (2006) Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. *Nat Protoc*, **1**, 1610-1616.

87. Wilkinson, K.A., Merino, E.J. and Weeks, K.M. (2005) RNA SHAPE chemistry reveals nonhierarchical interactions dominate equilibrium structural transitions in tRNA(Asp) transcripts. *J Am Chem Soc*, **127**, 4659-4667.

88. Das, R., Laederach, A., Pearlman, S.M., Herschlag, D. and Altman, R.B. (2005) SAFA: semi-automated footprinting analysis software for high-throughput quantification of nucleic acid footprinting experiments. *RNA*, **11**, 344-354.

89. Laederach, A., Das, R., Vicens, Q., Pearlman, S.M., Brenowitz, M., Herschlag, D. and Altman, R.B. (2008) Semiautomated and rapid quantification of nucleic acid footprinting and structure mapping experiments. *Nat Protoc*, **3**, 1395-1401.

90. Greider, C.W. and Blackburn, E.H. (1985) Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell*, **43**, 405-413.

91. Witkin, K.L. and Collins, K. (2004) Holoenzyme proteins required for the physiological assembly and activity of telomerase. *Genes Dev*, **18**, 1107-1118.

92. Lingner, J., Hughes, T.R., Shevchenko, A., Mann, M., Lundblad, V. and Cech, T.R. (1997) Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science*, **276**, 561-567.

93. Prathapam, R., Witkin, K.L., O'Connor, C.M. and Collins, K. (2005) A telomerase holoenzyme protein enhances telomerase RNA assembly with telomerase reverse transcriptase. *Nat Struct Mol Biol*, **12**, 252-257.

94. Romero, D.P. and Blackburn, E.H. (1991) A conserved secondary structure for telomerase RNA. *Cell*, **67**, 343-353.

95. O'Connor, C.M., Lai, C.K. and Collins, K. (2005) Two purified domains of telomerase reverse transcriptase reconstitute sequence-specific interactions with RNA. *J Biol Chem*, **280**, 17533-17539.

96. Miller, M.C. and Collins, K. (2002) Telomerase recognizes its template by using an adjacent RNA motif. *Proc Natl Acad Sci U S A*, **99**, 6585-6590.

97. Gilley, D. and Blackburn, E.H. (1999) The telomerase RNA pseudoknot is critical for the stable assembly of a catalytically active ribonucleoprotein. *Proc Natl Acad Sci U S A*, **96**, 6621-6625.

98. Sperger, J.M. and Cech, T.R. (2001) A stem-loop of Tetrahymena telomerase RNA distant from the template potentiates RNA folding and telomerase activity. *Biochemistry*, **40**, 7005-7016.

99. Licht, J.D. and Collins, K. (1999) Telomerase RNA function in recombinant Tetrahymena telomerase. *Genes Dev*, **13**, 1116-1125.

100. Richards, R.J., Theimer, C.A., Finger, L.D. and Feigon, J. (2006) Structure of the Tetrahymena thermophila telomerase RNA helix II template boundary element. *Nucleic Acids Res*, **34**, 816-825.

101. Lai, C.K., Miller, M.C. and Collins, K. (2002) Template boundary definition in Tetrahymena telomerase. *Genes Dev*, **16**, 415-420.

102. Bhattacharyya, A. and Blackburn, E.H. (1994) Architecture of telomerase RNA. *EMBO J*, **13**, 5721-5731.

103. Zaug, A.J. and Cech, T.R. (1995) Analysis of the structure of Tetrahymena nuclear RNAs in vivo: telomerase RNA, the self-splicing rRNA intron, and U2 snRNA. *RNA*, **1**, 363-374.

104. Richards, R.J., Wu, H., Trantirek, L., O'Connor, C.M., Collins, K. and Feigon, J. (2006) Structural study of elements of Tetrahymena telomerase RNA stem-loop IV domain important for function. *RNA*, **12**, 1475-1485.

105. Chen, Y., Fender, J., Legassie, J.D., Jarstfer, M.B., Bryan, T.M. and Varani, G. (2006) Structure of stem-loop IV of Tetrahymena telomerase RNA. *EMBO J*, **25**, 3156-3166.

106. Robart, A.R., O'Connor, C.M. and Collins, K. Ciliate telomerase RNA loop IV nucleotides promote hierarchical RNP assembly and holoenzyme stability. *RNA*, **16**, 563-571.

107. Berman, A.J., Gooding, A.R. and Cech, T.R. Tetrahymena telomerase protein p65 induces conformational changes throughout telomerase RNA (TER) and rescues telomerase reverse transcriptase and TER assembly mutants. *Mol Cell Biol*, **30**, 4965-4976.

108. Gherghe, C.M., Mortimer, S.A., Krahn, J.M., Thompson, N.L. and Weeks, K.M. (2008) Slow conformational dynamics at C2'-endo nucleotides in RNA. *J Am Chem Soc*, **130**, 8884-8885.

109. Mortimer, S.A. and Weeks, K.M. (2007) A fast-acting reagent for accurate analysis of RNA secondary and tertiary structure by SHAPE chemistry. *J Am Chem Soc*, **129**, 4144-4145.

110. Vendeix, F.A., Munoz, A.M. and Agris, P.F. (2009) Free energy calculation of modified base-pair formation in explicit solvent: A predictive model. *RNA*, **15**, 2278-2287.

111. Leontis, N.B., Stombaugh, J. and Westhof, E. (2002) The non-Watson-Crick base pairs and their associated isostericity matrices. *Nucleic Acids Res*, **30**, 3497-3531.

112. Maxam, A.M. and Gilbert, W. (1977) A new method for sequencing DNA. *Proc Natl Acad Sci U S A*, **74**, 560-564.

113. Keppler, B.R. and Jarstfer, M.B. (2004) Inhibition of telomerase activity by preventing proper assemblage. *Biochemistry*, **43**, 334-343.

114. Bryan, T.M., Goodrich, K.J. and Cech, T.R. (2000) Telomerase RNA bound by protein motifs specific to telomerase reverse transcriptase. *Mol Cell*, **6**, 493-499.

115. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.

116. Autexier, C. and Greider, C.W. (1998) Mutational analysis of the Tetrahymena telomerase RNA: identification of residues affecting telomerase activity in vitro. *Nucleic Acids Res*, **26**, 787-795.

117. Comolli, L.R., Smirnov, I., Xu, L., Blackburn, E.H. and James, T.L. (2002) A molecular switch underlies a human telomerase disease. *Proc Natl Acad Sci U S A*, **99**, 16998-17003.

118. Robart, A.R. and Collins, K. Investigation of human telomerase holoenzyme assembly, activity, and processivity using disease-linked subunit variants. *J Biol Chem*, **285**, 4375-4386.

119. Tesmer, V.M., Ford, L.P., Holt, S.E., Frank, B.C., Yi, X., Aisner, D.L., Ouellette, M., Shay, J.W. and Wright, W.E. (1999) Two inactive fragments of the integral RNA cooperate to assemble active telomerase with the human protein catalytic subunit (hTERT) in vitro. *Mol Cell Biol*, **19**, 6207-6216.

120. Abdelkafi, M., Leulliot, N., Baumruk, V., Bednarova, L., Turpin, P.Y., Namane, A., Gouyette, C., Huynh-Dinh, T. and Ghomi, M. (1998) Structural features of the UCCG and UGCG tetraloops in very short hairpins as evidenced by optical spectroscopy. *Biochemistry*, **37**, 7878-7884.

121. Reuter, J.S. and Mathews, D.H. RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics*, **11**, 129.

122. Chen, J.L. and Greider, C.W. (2005) Functional analysis of the pseudoknot structure in human telomerase RNA. *Proc Natl Acad Sci U S A*, **102**, 8080-8085; discussion 8077-8089.

123. Lin, J., Ly, H., Hussain, A., Abraham, M., Pearl, S., Tzfati, Y., Parslow, T.G. and Blackburn, E.H. (2004) A universal telomerase RNA core structure includes structured motifs required for binding the telomerase reverse transcriptase protein. *Proc Natl Acad Sci U S A*, **101**, 14713-14718.

124. Beattie, T.L., Zhou, W., Robinson, M.O. and Harrington, L. (2000) Polymerization defects within human telomerase are distinct from telomerase RNA and TEP1 binding. *Mol Biol Cell*, **11**, 3329-3340.

125. Yingling, Y.G. and Shapiro, B.A. (2007) The impact of dyskeratosis congenita mutations on the structure and dynamics of the human telomerase RNA pseudoknot domain. *J Biomol Struct Dyn*, **24**, 303-320.

126. Martin-Rivera, L. and Blasco, M.A. (2001) Identification of functional domains and dominant negative mutations in vertebrate telomerase RNA using an in vivo reconstitution system. *J Biol Chem*, **276**, 5856-5865.

127. Cristofari, G., Reichenbach, P., Regamey, P.O., Banfi, D., Chambon, M., Turcatti, G. and Lingner, J. (2007) Low- to high-throughput analysis of telomerase modulators with Telospot. *Nat Methods*, **4**, 851-853.

128. Zhang, Q., Kim, N.K., Peterson, R.D., Wang, Z. and Feigon, J. Structurally conserved five nucleotide bulge determines the overall topology of the core domain of human telomerase RNA. *Proc Natl Acad Sci U S A*, **107**, 18761-18768.

129. Mathews, D.H. (2006) RNA secondary structure analysis using RNAstructure. *Curr Protoc Bioinformatics*, **Chapter 12**, Unit 12 16.

130. Sexton, A.N. and Collins, K. The 5' guanosine tracts of human telomerase RNA are recognized by the G-quadruplex binding domain of the RNA helicase DHX36 and function to increase RNA accumulation. *Mol Cell Biol*, **31**, 736-743.

131. Varani, G. and McClain, W.H. (2000) The G x U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems. *EMBO Rep*, **1**, 18-23.

132. Disney, M.D., Labuda, L.P., Paul, D.J., Poplawski, S.G., Pushechnikov, A., Tran, T., Velagapudi, S.P., Wu, M. and Childs-Disney, J.L. (2008) Two-dimensional combinatorial screening identifies specific aminoglycoside-RNA internal loop partners. *J Am Chem Soc*, **130**, 11185-11194.

133. Dominick, P.K., Keppler, B.R., Legassie, J.D., Moon, I.K. and Jarstfer, M.B. (2004) Nucleic acid-binding ligands identify new mechanisms to inhibit telomerase. *Bioorg Med Chem Lett*, **14**, 3467-3471.

134. Gros, J., Guedin, A., Mergny, J.L. and Lacroix, L. (2008) G-Quadruplex formation interferes with P1 helix formation in the RNA component of telomerase hTERC. *Chembiochem*, **9**, 2075-2079.

135. Phan, A.T. Human telomeric G-quadruplex: structures of DNA and RNA sequences. *FEBS J*, **277**, 1107-1117.

136. Lattmann, S., Stadler, M.B., Vaughn, J.P., Akman, S.A. and Nagamine, Y. The DEAH-box RNA helicase RHAU binds an intramolecular RNA G-quadruplex in TERC and associates with telomerase holoenzyme. *Nucleic Acids Res*, **39**, 9390-9404.

137. Mihalusova, M., Wu, J.Y. and Zhuang, X. Functional importance of telomerase pseudoknot revealed by single-molecule analysis. *Proc Natl Acad Sci U S A*, **108**, 20339-20344.

138. Yeoman, J.A., Orte, A., Ashbridge, B., Klenerman, D. and Balasubramanian, S. RNA conformation in catalytically active human telomerase. *J Am Chem Soc*, **132**, 2852-2853.

139. Yi, X., Shay, J.W. and Wright, W.E. (2001) Quantitation of telomerase components and hTERT mRNA splicing patterns in immortal human cells. *Nucleic Acids Res*, **29**, 4818-4825.

140. Soares, J., Keppler, B.R., Wang, X., Lee, K.H. and Jarstfer, M.B. ortho-Quinone tanshinones directly inhibit telomerase through an oxidative mechanism mediated by hydrogen peroxide. *Bioorg Med Chem Lett*, **21**, 7474-7478.

141. Herbert, B.S., Gellert, G.C., Hochreiter, A., Pongracz, K., Wright, W.E., Zielinska, D., Chin, A.C., Harley, C.B., Shay, J.W. and Gryaznov, S.M. (2005) Lipid modification of GRN163, an N3'-->P5' thio-phosphoramidate oligonucleotide, enhances the potency of telomerase inhibition. *Oncogene*, **24**, 5262-5268.

142. Yeo, M., Rha, S.Y., Jeung, H.C., Hu, S.X., Yang, S.H., Kim, Y.S., An, S.W. and Chung, H.C. (2005) Attenuation of telomerase activity by hammerhead ribozyme targeting human telomerase RNA induces growth retardation and apoptosis in human breast tumor cells. *Int J Cancer*, **114**, 484-489.

143. Kaiser, M., Sainlos, M., Lehn, J.M., Bombard, S. and Teulade-Fichou, M.P. (2006) Aminoglycoside-quinacridine conjugates: towards recognition of the P6.1 element of telomerase RNA. *Chembiochem*, **7**, 321-329.

144. Disney, M.D. Studying modification of aminoglycoside antibiotics by resistance-causing enzymes via microarray. *Methods Mol Biol*, **808**, 303-320.

145. Colgin, L.M., Wilkinson, C., Englezou, A., Kilian, A., Robinson, M.O. and Reddel, R.R. (2000) The hTERTalpha splice variant is a dominant negative inhibitor of telomerase activity. *Neoplasia*, **2**, 426-432.