Biophysical Investigation of G-Quadruplex Recognition by the N-Terminal Construct of RNA Helicase Associated with AU-Rich Element (RHAU)

by

Oksana Marushchak

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

Department of Chemistry

University of Manitoba

Winnipeg, Manitoba, Canada

Copyright © 2014 by Oksana Marushchak
Abstract

G-quadruplexes, characterized by stacked G-tetrad rings held together by Hoogsteen hydrogen bonds, have been visualized in human cells and implicated in transcriptional and translational control, telomere maintenance and disease. RHA Helicase associated with AU-rich element (RHAU), a DEAH-box helicase, is a major G-quadruplex resolvase in human cell lysates. It binds G-quadruplexes through the RHAU specific motif in its N-terminus. In order to investigate the recognition of G-quadruplexes by helicases, the binding between the N-terminal construct of RHAU, RHAU_{53-105}, and the DNA analog of the quadruplex formed by the 5’ terminus of human telomerase RNA component, hTR_{1-20}, was investigated in a comprehensive biophysical approach followed by crystallization screening. RHAU_{53-105}, hTR_{1-20} DNA and their complexes were analysed by gel electrophoresis, UV-visible spectroscopy, spectropolarimetry, dynamic light scattering and small angle X-ray scattering (SAXS). The findings reveal that hTR_{1-20} DNA, separated in two conformations by size exclusion chromatography in the presence of potassium cations, assumes a disk-like parallel G-quadruplex secondary structure in solution. Far-UV circular dichroism spectra and SAXS demonstrate that RHAU_{53-105} assumes an extended ($D_{max} = 7.8$ nm, $r_G = 2.1$ (±0.2) nm) and ordered conformation in solution. The analysis confirms the binding between RHAU_{53-105} and each conformation of the hTR_{1-20} DNA quadruplex. Circular dichroism spectra indicate the retention of quadruplex secondary structure in both RHAU_{53-105}·hTR_{1-20}DNA_{c1} and RHAU_{53-105}·hTR_{1-20}DNA_{c2} complexes. This analysis provides some insight into the interaction between G-quadruplexes and the N-terminal domain of RHAU and identifies 0.2 M sodium formate, 20 % (w/v) polyethylene glycol 3350 and 1.5 M sodium chloride, 10 % (v/v) ethanol as preliminary conditions for crystallization of the complex of RHAU_{53-105} and hTR_{1-20}DNA_{c2}. 
Acknowledgements

I would like to thank Dr. Jorg Stetefeld, Dr. Markus Meier and Dr. Trushar Patel for their supervision, guidance and support. I would also like to extend my gratitude to Dr. Sean McKenna, Dr. Evan Booy and Dr. George Orriss for their assistance.

Many thanks also go to Matthew McDougall and Natalie Krahn for their friendship, advice and support. In addition, I would like to express my appreciation for financial assistance from the Department of Chemistry, Faculty of Science and Faculty of Graduate Studies at the University of Manitoba as well as the Manitoba Health Research Council.
Table of Contents

Abstract ......................................................................................................................... i
Acknowledgements ...................................................................................................... ii
Table of contents ......................................................................................................... iii
List of tables ................................................................................................................ v
List of figures ............................................................................................................... vi
List of copyrighted material for which permission was obtained .................................viii
1. Introduction .............................................................................................................. 1
   1.1 G-quadruplex secondary structure of oligonucleotides .............................................. 1
   1.2 Distribution of G-quadruplex motifs within human genome ...................................... 4
   1.3 G-quadruplex helicases .......................................................................................... 7
   1.4 RHA Helicase Associated with the AU-element is a major G-quadruplex resolvase .... 8
   1.5 Objectives ............................................................................................................ 13
2. Materials and Methods .......................................................................................... 14
   2.1 G-quadruplex preparation ....................................................................................... 14
   2.2 Expression and purification of RHAn53-105 ........................................................... 14
   2.3 Expression and purification of RHAn53-150 ........................................................... 15
   2.4 Preparation of protein-quadruplex complexes ......................................................... 16
   2.5 UV-visible absorption spectroscopy ....................................................................... 16
   2.6 Electrophoretic mobility shift assay ...................................................................... 16
   2.7 Circular dichroism spectropolarimetry ................................................................... 17
2.9 Small angle X-ray scattering ......................................................... 18
2.10 Crystallization trials ................................................................. 19
2.11 Test for expression of RHAU_{53-1008} in *Escherichia Coli* (*E. Coli*) ........................................... 22
2.12 Purification of His_{6}-RHAU_{53-1008} ........................................... 22
2.13 Western blot for detection of RHAU_{53-1008} .................................. 23
2.14 Cloning of RHAU_{53-1008}^{Δ14}.pCEP-Pu.BM40 ........................................ 24

3. Results .............................................................................................. 27
   3.1 DNA analog of the hTR_{1-20} sequence assumes G-quadruplex secondary structure in solution ................................................................. 27
   3.2 RHAU_{53-105} adopts extended conformation with ordered secondary structure in solution ................................................................. 41
   3.3 RHAU_{53-105} binds hTR_{1-20} DNA G-quadruplex in solution ........................................... 46
   3.4 Crystallization trials ..................................................................... 51
   3.5 Expression and purification of the full-length RHAU protein ..................... 56

4. Discussion ........................................................................................... 59

5. Summary and Future Direction ............................................................. 66

6. References .......................................................................................... 68

7. Appendix ............................................................................................. 75
   7.1 List of abbreviations ...................................................................... 75
List of Tables

Table 1. Crystallization solutions screened in an effort to optimize crystals of RHAU$_{53-105}$·hTR$_1$.

20 DNA$_2$ initially obtained in 0.2 M formate salt, 20 % (w/v) PEG 3500..........................21

Table 2. Summary of hydrodynamic data for the protein, quadruplexes and complexes.........50
List of Figures

Figure 1. Schematic representation of a G-quadruplex.................................................................2
Figure 2. Schematic representation of the putative domains of the full-length RHAU protein.................................................................12
Figure 3. Purification of hTR1-20 DNA by size exclusion chromatography.................................31
Figure 4. The analysis of interconversion between the two conformations of hTR1-20 DNA by UV-Vis spectrophotometry and spectropolarimetry.................................................................32
Figure 5. Elution profiles of N-terminal RHAU, quadruplexes and their complexes acquired from the Superdex 200 10/300 GL size exclusion chromatography column........................................33
Figure 6. Electrophoretic Mobility Shift Assay with RHAU53-105 and hTR1-20 DNA quadruplex.................................................................................................................................34
Figure 7. Indication of G-quadruplex structure in UV-visible spectra obtained for hTR1-20 DNAc1 and hTR1-20 DNAc2..................................................................................................................35
Figure 8. Analysis of hTR1-20 DNA G-quadruplexes by spectropolarimetry................................36
Figure 9. Effect of buffer composition on the conformations of hTR1-20 DNAc1 and hTR1-20 DNAc2........................................................................................................................................37
Figure 10. Melting curves of the G-quadruplexes obtained by spectropolarimetry at 263 nm.................................................................................................................................38
Figure 11. Dynamic light scattering profiles of N-terminal RHAU constructs, quadruplexes and their complexes........................................................................................................39
Figure 12. Solution conformation of hTR1-20 DNA G-quadruplexes determined by SAXS.................................................................................................................................40
Figure 13. Expression and purification of RHAU\textsubscript{53-105} .......................................................... 42

Figure 14. Far UV-visible CD spectrum of RHAU\textsubscript{53-105} .......................................................... 43

Figure 15. Effect of buffer composition on the conformation of RHAU\textsubscript{53-105} ........................................ 44

Figure 16. N-terminal domain of RHAU adopts an extended ordered conformation in solution .......................... 45

Figure 17. Conformational change associated with protein binding to DNA G-quadruplex ................................. 48

Figure 18. Solution conformations of G-quadruplex-protein complexes determined by SAXS .................................. 49

Figure 19. Image of crystals of hTR\textsubscript{1-20} DNA\textsubscript{c2} grown by vapour diffusion in 0.4 M potassium sodium tartrate tetrahydrate at 20 °C ................................................................. 53

Figure 20. Image of crystals of hTR\textsubscript{1-20} DNA\textsubscript{c2}-RHAU\textsubscript{53-105} grown by vapour diffusion in 0.2 M sodium formate, 20 % w/v polyethylene glycol 3350 ......................................................... 54

Figure 21. Image of crystals of hTR\textsubscript{1-20} DNA\textsubscript{c2}-RHAU\textsubscript{53-105} grown by vapour diffusion in 1.5 M sodium chloride, 10 % v/v ethanol .......................................................... 55

Figure 22. Expression of RAHU\textsubscript{53-1008} in Escherichia coli cells ......................................................... 57

Figure 23. Purification of His\textsubscript{6}-RHAU\textsubscript{53-1008} expressed in Escherichia coli BL21(DE3)-RIPL cells ................................................................................................................................. 58
List of Copyrighted Material for which Permission was Obtained

Figure 1 - Figure by G-quadruplex.jpg: Julian Huppert Jlh29 derivative work: Iridos (G-quadruplex.jpg) [CC-BY-SA-2.5 (http://creativecommons.org/licenses/by-sa/2.5)], via Wikimedia Commons.

Figure 1 appears on page 2.

Figure 14 – Reference circular dichroism spectra reproduced with permission from Elsevier.


https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=0b0f703a-b81b-460b-b08a-b71b1f020827

Figure 14 appears on page 43.

Figure 16 – Reproduced with permission from Journal of Biological Chemistry.

http://www.jbc.org/site/misc/Copyright_Permission.xhtml


Figure 16 appears on page 45.
1. Introduction

1.1 G-QUADRUPLEX SECONDARY STRUCTURE OF OLIGONUCLEOTIDES

Oligonucleotides possessing consecutive stretches of three or more guanine bases can self-assemble into G-quadruplexes. These thermodynamically stable secondary structures result from stacking of planar guanine tetrads that form when four quanine bases associate in a ring-like structure through Hoogsteen hydrogen bonds (Figure 1). A central monovalent cation, usually potassium, sodium or lithium, is present in a central ion channel, stabilizing the quadruplex. Quadruplex structures with central cation positioned in line with the G-tetrads or in bi-pyramidal arrangement with successive G-tetrads have been reported.\(^1\text{-}^4\) Melting temperatures around 80º C have been reported for G-quadruplexes synthesized \textit{in vitro}.\(^5\text{-}^7\)

Strand orientation, conformation of glycosidic angles of guanines, loop size and the nature of the central monovalent cations affect G-quadruplex conformation. Solution of high resolution deoxyribonucleic acid (DNA; refer to list of abbreviations in appendix) G-quadruplex crystal structures by NMR and X-ray crystallography highlights their structural diversity.\(^1,3,4,8\text{-}15\) Both intermolecular and intramolecular oligonucleotide arrangements can assemble into G-quadruplexes. Intermolecular quadruplexes consist of either two or four guanine-rich strands that assemble together with the guanine bases pointing inside the quadruplex. In contrast, intramolecular quadruplexes are unimolecular structures that form when tandem guanine repeats separated by a few non-guanine nucleotides fold to form guanine rings held together by Hoogsteen-hydrogen bonds. G-quadruplexes can be further classified as parallel, antiparallel or mixed. G-quadruplexes are designated as parallel if all strands involved in the formation of the stacked G-tetrads have the same polarity and antiparallel if they run in opposite directions.
Figure 1. Schematic representation of a G-quadruplex. A. Hoogsteen-hydrogen bonding between four guanines assembled in a ring. A monovalent cation positioned in the central ion channel coordinates with O₆ carbonyl groups. B. A diagram of an intramolecular G-quadruplex. Successive G-tetrads stack on top of each other. Figure by G-quadruplex.jpg: Julian Huppert Jlh29 derivative work: Iridos (G-quadruplex.jpg) [CC-BY-SA-2.5](http://creativecommons.org/licenses/by-sa/2.5), via Wikimedia Commons.
Despite their conformational variability, G-quadruplexes possess high thermodynamic stability due to stacking of the G-tetrads. The crystal structure of a parallel intramolecular quadruplex formed by a 22 nucleotide sequence from human telomeric DNA (d[AGGG(TTAGG)₃]) in the presence of potassium cations shows guanine glycosidic bonds in the anti-conformation and π-π stacking interactions between the three planar G-tetrads.³ Propeller-like loops connect the bottom of one strand with the top of another without blocking the terminal G-tetrads.³ Potassium cations are positioned centrally between the G-tetrad planes in a bi-pyramidal arrangement with the eight O₆ carbonyl groups.³ The same sequence folds into an antiparallel quadruplex in the presence of sodium cations.³ Of the studied sequences, most assume a parallel arrangement in the presence of potassium and antiparallel arrangement in the presence of sodium cations, although high resolution structures of tetramolecular parallel quadruplexes in the presence of sodium cations although exist.¹⁻³ Since the cytoplasmic concentration of potassium in much higher than the concentration of sodium in a cellular environment, it is likely that G-quadruplex structures produced in the presence of potassium are more physiologically relevant.¹⁶ In contrast to DNA quadruplexes, less ribonucleic acid (RNA) G-quadruplex structures are available. A crystal structure of TERRA RNA quadruplex, a recently discovered non-coding transcript of telomeric DNA, in the presence of potassium shows two parallel three-layer quadruplexes held together by stacking interactions between terminal G-tetrads.¹⁷ The structure of RNA G-quadruplex appears to be affected by the ribonucleotide 2’-hydroxyl group that contributes to hydrogen bonding within the quadruplex and determines the preference for a parallel conformation.¹⁷
1.2 DISTRIBUTION OF G-QUADRUPLEX MOTIFS WITHIN HUMAN GENOME

Until recently, G-quadruplexes were assumed to be aberrant secondary structures formed by G-rich nucleotides in vitro. However, within the last few years both DNA and RNA G-quadruplexes were visualized inside human cells and quadruplex forming motifs have been implicated in telomere maintenance, genome expression through translational and transcriptional regulation, immunoglobulin switching, epigenetic instability and disease.\textsuperscript{18-20} In situ analysis of the human genome predicted 376 000 potential G-quadruplex forming motifs.\textsuperscript{21} In order to eliminate random G-rich sequences from the results, the algorithm used in this analysis took into account loop length between guanine runs suitable for G-quadruplex formation. G-quadruplex forming motifs are overrepresented in proto-oncogenes and parts of the genome associated with transcriptional control, growth factors, cell growth, cell signalling and development as well as in 5’-untranslated regions (UTRs), first intron and 3’-UTRs regions of mRNA transcripts.\textsuperscript{22} They are underrepresented in regions of the genome associated with cell adhesion, ubiquitin cycle, nucleosome assembly, G-protein coupled receptors and tumor suppressor genes.\textsuperscript{22} G-quadruplexes have also been implicated in epigenetic control of gene expression. Hypermethylated CG sequences associated with transcriptional silencing are underrepresented within regions of the genome predicted to fold into G-quadruplexes.\textsuperscript{23}

G-quadruplex forming motifs were predicted in over 40% of human gene promoters.\textsuperscript{24} Both c-MYC and c-KIT oncogenes are transcriptionally repressed by G-quadruplex structures within their respective promoter regions.\textsuperscript{6,25} Expression of c-MYC is associated with cellular proliferation, breast, colon and cervical cancers as well as myeloid leukemias. The c-KIT oncogene encodes a tyrosine receptor and is associated with stem cell growth and differentiation.\textsuperscript{6} Human proto-oncogene NRAS and human zinc finger protein Zic-1 are
transcriptionally suppressed by quadruplexes located in the 5'-UTRs of their respective mRNAs.\textsuperscript{26,27} In addition to acting as \textit{cis}-regulatory elements for transcription and translation, G-quadruplexes also regulate mRNA splicing. In \textit{vitro} stabilization of a quadruplex structure within intron 6 of mRNA of the human telomerase reverse transcriptase (hTERT), an enzyme responsible for telomere extension in germ cells, results in production of an inactive splice variant of the enzyme.\textsuperscript{28} Splicing of tumor suppressor gene p53 mRNA is also regulated by a G-quadruplex in its intron 3.\textsuperscript{29} A shorter splice variant produces a protein that regulates p53 activity. Polymorphisms at the G-quadruplex forming sequence in the \textit{TP53} gene, which codes for p53 protein, correlate with early-onset cancers. Moreover, a G-quadruplex motif within the internal ribosome entry site (IRES) on the 5'-alternatively translated region (ATR) of mRNA for human fibroblast factor FGF-2 is required for 5'-cap independent translation of FGF-2.\textsuperscript{30}

G-quadruplexes are also found within telomeres. In humans, 3’ terminal chromosomal regions consist of several kilobasepairs of TTAGGG sequence repeats followed by a single stranded tail a few hundred base pairs in length, which is required for loading of hTERT.\textsuperscript{31} Telomeres protect linear DNA from degradation during replication.\textsuperscript{31} DNA polymerase, which copies single-stranded DNA in the 5’ to 3’ direction, uses an RNA primer to load onto the strand and initiate replication. The leading strand is copied continuously, but the lagging strand is copied in short fragments called Okazaki fragments and the spaces vacated by the RNA primers are later filled by DNA ligase to produce a continuous DNA strand. However, the terminal 5’ sequence is lost during each replication event because DNA ligase can only fill and ligate regions bordered by DNA on each side. Thus, telomeres protect the coding regions of chromosomes from shortening with each replication event. Hence, telomeres determine the number of replications that a somatic cell can undergo before apoptosis.\textsuperscript{32} In addition, telomeres
are thought to protect chromosomal ends from fusion with each other. Telomeres exist in complex with proteins thought to be involved in hTERT recruitment and protection of single-stranded ends. Numerous G-quadruplexes structures separated by flexible linkers form on telomeric strands and their arrangement has been described as “beads-on-a-string”.

G-quadruplexes have been detected inside human cells using a G-quadruplex-specific single chain antibody called BG4. Only about one quarter of the detected spots were located at telomeres, confirming in situ studies of G-quadruplex distribution within the genome. In addition, the lowest number of quadruplexes was detected during the G0/G1 stage of the cell cycle, when no DNA replication occurs and about a five-fold increase in signal was detected during the S phase, when DNA replicates, suggesting that quadruplexes form in G-rich sequences when single-stranded DNA is exposed. In contrast to regulatory quadruplexes, these non-specific secondary structures may cause the DNA and RNA polymerases to stall, effectively inhibiting replication or transcription and leading to disease. Interestingly, however, the stalling by DNA polymerase δ was found to persist in vitro even after the G-quadruplex forming sequences were mutated to disrupt quadruplex formation. Nevertheless, polymorphisms resulting in expansion of G-rich repeats are associated with syndromes characterized by increased chromosomal abnormalities, double-stranded DNA breakages and premature termination of transcription. Neurodegenerative disorders amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) are associated with expansion of GGGGCG repeats in intron 1 of the C9orf72 gene. While healthy individuals carry less than 30 repeats of this sequence, a fraction of affected patients have between 700 and 1600 repeats. Using circular dichroism and nuclear magnetic resonance, this expansion has been shown to form parallel G-quadruplex in vitro. Increased number of the GGGGCG repeat within the C9orf72 is correlated
with accumulation of aborted mRNA transcripts of the C9orf72 gene, which causes nucleolar stress and cytotoxicity.\textsuperscript{40} Aborted transcripts might sequester proteins important for ribosomal assembly or regulation, making them unavailable to other transcripts or some transcripts may be translated into dysfunctional or misfiled proteins that later aggregate in the cytoplasm.\textsuperscript{40} Thus, specialized helicases capable of unwinding G-quadruplexes are required for normal genome function.

1.3 G-QUADRUPLEX HELICASES

Bloom syndrome protein (BLM), Werner protein (WRN) and Fanconi anemia group J protein (FANCJ) are a few of the DNA helicases shown to be able to unwind DNA G-quadruplex structures into single-stranded DNA.\textsuperscript{36-38,42} Deficiency in the BLM protein that unravels G-quadruplex DNA in the 3’ to 5’ direction is associated with Bloom’s syndrome, which is characterized by increased incidence of chromatid breaks and sister chromatid exchange leading to growth retardation, impaired fertility, immunodeficiency and cancers.\textsuperscript{36} The WRN helicase unwinds bimolecular G-quadruplexes formed by sequence d(CGG)$_7$ in the 3’ to 5’ direction and exhibits 3’ to 5’ exonuclease activity.\textsuperscript{37} Unlike BLM and WRN, FANCJ, another DNA G-quadruplex helicase, unwinds the quadruplex in the 5’ to 3’ direction and its deficiency is associated with increased number of deletions within G-rich regions of DNA.\textsuperscript{38,42} Interaction between BLM and FANCJ has been detected, suggesting that they might simultaneously unravel quadruplexes from opposite ends.\textsuperscript{34}

Numerous compounds specific for G-quadruplexes have been identified\textsuperscript{3,20,43} Acridines, tetra-N-pyridyl porphyrins and triazine derivatives bind G-quadruplexes.\textsuperscript{3} Additionally, a dye N-methyl mesoporphyrin IX (NMM) is specific for parallel quadruplexes.\textsuperscript{11} The crystal structure of
a human telomeric quadruplex bound by NMM reveals \(\pi-\pi\) stacking between the terminal G-tetrad of the quadruplex and the compound.\(^{11}\) Unlike in the antiparallel quadruplex, the loops of the parallel quadruplex do not block the terminal G-tetrad faces, allowing them to participate in stacking interactions with ligands. The N-methyl group of NMM was found to fit inside the quadruplex ion channel and interact with the potassium cation.\(^{11}\) Another compound, carboxyPDS, is specific for RNA quadruplexes and together with BG4 antibody has been used to visualize RNA G-quadruplexes in the cytoplasm of human cells.\(^{19}\)

1.4 RNA HELICASE ASSOCIATED WITH \(AU\)-RICH ELEMENT IS A MAJOR G-QUADRUPLEX RESOLVASE

RNA Helicase Associated with \(AU\)-rich element (RHAU) (also known as DHX36 and DNA resolvase 1) is a prominent G-quadruplex resolvase in human cells. RHAU is a DEAH-box helicase composed of 1008 amino acids, weighs 119 kDa and is encoded by the \(Dhx36\) gene. Two splice variants of RHAU were detected in HeLa cell extracts.\(^{44}\) The full length RHAU mainly localizes in the nucleus and RHAU\(^{\Delta14}\), which lacks a 14 amino acid long nuclear localization signal, is only present in the cytoplasm.\(^{44}\) The concentration of RHAU was found to be significantly higher in the nucleus than in the cytoplasm.\(^{44}\) Originally, RHAU was identified as a protein involved in mRNA deadenylation and decay for its ability to co-precipitate with ARE\(^{\text{SPA}}\)-mRNA from HeLa cell extracts.\(^{44}\) It was also found to interact with the exosome and polyadenylate-specific ribonucleases (PARN) in an RNA-independent fashion and with NFAR1 and HuR in a RNA-mediated fashion.\(^{44}\) Since its discovery, RHAU has been implicated in mRNA destabilization, transcriptional and translational regulation and blood cell formation.\(^{44-47}\) RHAU was detected in nucleoplasm and also in nuclear speckles in HeLa cells and was found to
associate with RHA helicases p68 and p72 within nucleolar caps upon transcriptional arrest.\textsuperscript{48} Nuclear speckles are areas of the nucleus enriched in pre-mRNA splicing factors and nucleolar caps contain mostly RNA binding proteins that aggregate in the nucleolus during periods of cellular stress or arrested transcription. RHAU also binds and targets RNA to stress granules, which are cytoplasmic structures that sequester cellular proteins and RNA during periods of stress to the cell.\textsuperscript{49} RHAU functions as a transcriptional inducer for TNAP, which is involved in bone formation and differentiation.\textsuperscript{45} Moreover, RHAU is involved in translational regulation of transcriptional regulator Yin Yang 1 (YY1) and a homeobox transcription factor PITX1 that has been implicated in development and cancer.\textsuperscript{46,50} RHAU was shown to be essential for murine hematopoiesis and its depletion in developing mouse embryos resulted in hemolytic anemia and blood cell differentiation defects.\textsuperscript{47} RHAU was also co-precipitated with neuronal precursor-micro-RNA-134 (miRNA134) and was implicated in miRNA134-dependent gene inhibition and control of growth of dendritic spines, which are dendritic membranous protrusions that receive signals from axons as synaptic junctions.\textsuperscript{51}

RHAU has sustained most attention as a G-quadruplex resolvase. RHAU was found to unwind tetramolecular G-quadruplexes \textit{in vitro}, but was later shown to display higher affinity for tetramolecular RNA quadruplexes.\textsuperscript{52,53} In a subsequent study, RHAU’s ability to unwind intramolecular G-quadruplexes in an ATP-dependent fashion was shown to out-compete its tetramolecular G-quadruplex resolving activity \textit{in vitro}.\textsuperscript{54} RHAU binds and resolves the RNA G-quadruplex in the 5’ terminal region of the RNA component of the human telomerase reverse transcriptase enzyme, allowing proper formation of the P1 helix structure required for correct incorporation of nucleotides into the template used for reverse transcription of telomeric DNA.\textsuperscript{55–57} Human telomerase reverse transcriptase is a ribonucleoprotein active in highly proliferating
germ and cancer cells, but not in somatic cells. It is composed of a catalytic protein subunit (TERT) and a human telomerase reverse transcriptase RNA component (hTR) that acts as a template for telomeric extension. Activation of hTERT in somatic cells is associated with unchecked replication and cancerous tumors. Sexton et al. and Giri et al. separately showed that RHAU binds and unwinds a 5’ terminal quadruplex in hTR in vitro. These findings were confirmed by Lattmann et al., who demonstrated that RHAU targets hTERT in vivo.

Human telomerase RNA is 452 nucleotides long and is transcribed by RNA polymerase II. Eight highly conserved regions (CRs) numbered from 5’ to 3’ have been identified in vertebrate telomerase RNA. Ten conserved helical structures called paired helices (P) make up four domains in the secondary structure of vertebrate telomerase RNA. The conserved domains include the pseudoknot domain, the CR4-CR5 domain, the BOX H/ACA domain and the CR7 domain. CR4-CR5 is required for hTERT binding to hTR and formation of an active hTERT holoenzyme. CR1, located 45 nucleotides downstream of the 5’ end of hTR, encodes the template sequence 5’-CUAACCCU-3’ required for extension of telomeric DNA. In humans, incorporation of correct nucleotides into the template sequence during reverse transcription is accomplished by constraining the sequence through hTR-hTERT interaction and RNA-RNA base-pairing upstream of the sequence. Helix P1 that forms upstream of the 5’ region of the template acts as an anchor that tethers the template region through a short linker. Disruption of P1 helix formation through mutation results in synthesis of telomeric repeats of incorrect length and incorporation of extra nucleotides into the template. The P1 helix consists of two regions, P1a and P1b, separated by a loop. The P1b structure is essential for correct template incorporation. The 5’ terminal region of hTR can form several quadruplexes that might incorporate the nucleotides required to form the P1 helix. Nucleotides 1 to 17 correspond to the
minimal length of hTR capable of G-quadruplex formation in vitro.\textsuperscript{57} These quadruplexes may serve to protect the single stranded 5’ terminus of hTR from degradation during maturation of hTERT, but they must be unwound to allow P1 helix formation and correct template demarcation.\textsuperscript{56,57} Bisquinolinium compounds Phen-DC\textsubscript{3} and Phen-DC\textsubscript{6} and a 2,6-pyridine-dicarboxamide derivative 360A are G-quadruplex-specific ligands that have been shown to prevent P1 helix formation in hTR by binding to and stabilizing the 5’-terminal G-quadruplex structures in hTR.\textsuperscript{60,61} These compounds had an inhibitory effect on P1 helix formation only if the oligonucleotide sequence was able to form G-quadruplex and the effect was absent if the sequence was mutated to disrupt G-quadruplex formation.\textsuperscript{61} Since the RHAU binds and unwinds both DNA and RNA synthetic G-quadruplexes formed by the 5’ terminus of hTR and knockdown of endogenous RHAU in HEK293T cells results in observed reduction in average telomere length.\textsuperscript{57} Additionally, deletion studies have revealed that RHAU-specific motif (RSM) within the N-terminus of RHAU is required for quadruplex recognition (Figure 2).\textsuperscript{56} The RSM region of RHAU is conserved among vertebrates, suggesting its evolutionary importance and RHAU’s physiological significance.\textsuperscript{56} The helicase domain is not required for quadruplex binding, which is independent of ATP-hydrolysis.\textsuperscript{56}
RHAU is classified as a RNA helicase belonging to the DEAH family. DHA and RNA helicases are multifunctional ATP-dependent enzymes involved in nucleic acid remodelling and nucleoprotein dissociation. Most helicases belong to the phylogenetically conserved superfamily 2 (SF2), which is subdivided into five families based on conserved structure and sequence. Most of the helicases studied belong to the DEAD and DExH/DEAH families that share conserved structural elements. SF2 helicases share seven conserved motifs encompassed within two domains that fold into a claw-like structure with an ATP-binding and hydrolysis site in the cleft between the two domains. The DEAD (Asp-Glu-Ala-Asp) and DExH/DEAH (Asp-Glu-x-His/Asp-Glu-Ala-His) families are classified based on the amino acids in the ATP-binding site. Additional N-terminal and C-terminal domains have protein-specific functions. Both DEAD-box and DExH/DEAH-box helicases unwind duplex DNA in the 3’ to 5’ direction, but unlike non-processive DEAD-box helicases that unwind only small helices, DExH/DEAH-box helicases translocate along the nucleotide strand and unwind helices up to several dozen base pairs in
length. In addition to local structural rearrangement, helicases can remodel protein-RNA and protein-DNA complexes by displacing bound proteins. The high resolution structure of the yeast RNA helicase Prp43p, which shares considerable sequence similarity with RHAU, has become available recently and provides some insight into the mode of action of RNA helicases. Prp43p lacks an extended N-terminal domain present in RHAU, but the structure suggests that a conformational change induced by ATP hydrolysis is required to uncover the RNA binding site. The N-terminal domain of RHAU is followed by a flexible linker that may be involved in the conformational rearrangement of the protein required to deliver the bound quadruplex to the helicase domain. RHAU is an important helicase to be studied in the context of its interaction with G-quadruplexes.

1.5 OBJECTIVES

Here, we investigated the binding of the N-terminal construct of RHAU (residues 53-105) with DNA analog of the quadruplex-forming nucleotides 1-20 from the human telomerase RNA component. We used an integrated biophysical approach including circular dichroism, dynamic light scattering and small angle X-ray scattering to study the interaction between RHAU$_{53-105}$ and hTR$_{1-20}$ DNA. In addition, several potential conditions for crystallization of the complex between RHAU$_{53-105}$ and hTR$_{1-20}$ DNA$_{c2}$ were identified. Although structures of complexes between G-quadruplexes and small molecule ligands are available, there is a lack of information regarding the recognition of G-quadruplexes by helicases. This is an important path for investigation due to widespread involvement of G-quadruplexes in DNA and RNA metabolism and in the correct functioning of hTERT in particular.
2. Materials and Methods

2.1 G-QUADRUPLEX PREPARATION

Synthetic hTR1-20 DNA sequence 5’-GGGUUGCGAGGGUGGGCCU-3’ was purchased from Alpha DNA, Canada and was dissolved in 20 mM HEPES, pH 7.5, 100 mM KCl at a concentration of 10 mM. The solution was heated to 95°C and was allowed to passively cool to room temperature in a water bath. The solution was stirred continuously. The folded quadruplex was purified using the HiLoad Superdex 75 26/60 size exclusion chromatography (SEC) column and was eluted in 20 mM HEPES, pH 7.5, 100 mM KCl buffer. The quadruplex eluted in two peaks. The first peak to elute was identified as hTR1-20 DNA C1 and the second as hTR1-20 DNA C2. The extinction coefficient ($\varepsilon_{260\text{ nm}}=189 000 \text{ M}^{-1}\text{ cm}^{-1}$) used to calculate hTR1-20 DNA concentration was calculated using IDT SciTools® (Primer-Quest Program, Integrated DNA Technologies).

2.2 EXPRESSION AND PURIFICATION of RHAU53-105

RHAU53-105 was expressed and purified according to the protocol described previously. The gene coding for RHAU53-105 with an N-terminal hexahistidine (His6) tag followed by a thrombin cleavage site was cloned into a pET15b vector carrying an ampicillin resistance gene. RHAU53-105 was expressed in Escherichia coli (E. coli) strain BL2(DE3) in luria broth (LB) in the presence of ampicillin. A 25 ml pre-culture grown overnight was used to start a 1.8 L culture grown at 37 °C with constant shaking at 200 rpm. Expression was induced with addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG) to final concentration of 5 mM once OD600 reached 0.6. Cells were grown for three hours and the pellet was collected by centrifugation at
6000 rpm for 20 minutes at 4 °C. The pellet was frozen at -80 °C. The cells were thawed on ice and lysed in 20 mM sodium phosphate, 2 mM imidazole, 6 M guanidine hydrochloride, pH 7.0 in the presence of proteases using homogenization followed by sonication. The protein was purified by affinity chromatography using Talon\textsuperscript{®} cobalt resin. The column was equilibrated using two column volumes of 20 mM sodium phosphate, 2 mM imidazole, pH 7.0 followed by two volumes of the lysing buffer. Following the loading of the cell lysate, the column was washed using the equilibration buffer and the protein was refolded and eluted in 150 mM imidazole, pH 7.0, I=300 mM (NaCl). The purified protein was concentrated to 1 mg/ml and dialysed overnight into 10 mM HEPES, pH 7.5, I=154 mM (NaCl). Thrombin was added to the dialysis membrane to cleave the His\textsubscript{6} tag. Following dialysis and His\textsubscript{6} tag cleavage, thrombin was removed using a HiTrap benzamidine FF column (GE Healthcare). The extinction coefficient (\(\varepsilon_{280\text{ nm}}=6990\text{ M}^{-1}\text{ cm}^{-1}\)) for RHAU\textsubscript{53-105} was calculated using the ExPASy ProtParam and the concentration of the protein was determined spectrophotometrically.\textsuperscript{72}

2.3 EXPRESSION AND PURIFICATION of RHAU\textsubscript{53-150}

Originally, expression and purification of RHAU\textsubscript{53-209} was attempted by a fellow laboratory member. His\textsubscript{6}-tagged RHAU\textsubscript{53-209} was cloned into the pET15b vector and expressed in E. coli strain BL21(DE3) cells using the same protocol as for RHAU\textsubscript{53-105}. Purification and His\textsubscript{6} tag cleavage were also carried out as described for RHAU\textsubscript{53-105}. The molecular weight of the resulting protein, estimated using SDS-PAGE, was less than expected for RHAU\textsubscript{53-209}, suggesting that the protein degraded during purification. Mass spectrometric analysis revealed the resulting protein was RHAU\textsubscript{53-150}.
2.4 PREPARATION OF PROTEIN-QUADRUPLEX COMPLEXES

RHAU\textsubscript{53-105} and hTR\textsubscript{1-20} DNA were mixed in 1:1 molar ratio and diluted in 20 mM HEPES, pH 7.5, 100 mM KCl. Each component was added at a concentration of 10 mM. The solution was left at room temperature for 20 minutes and was then concentrated to 5 mg/ml. The concentration of the complex was measured spectrophotometrically and calculated using the absorption coefficient for the DNA component. The complexes were purified on a Superdex 200 10/300 GL SEC column and eluted using 20 mM HEPES, pH 7.5, 100 mM KCl. The extinction coefficient for the nucleotide component (\(\epsilon_{260\text{ nm}}=189\ 000\ M^{-1}\text{cm}^{-1}\)) was used to calculate the concentration of the complexes.

2.5 UV-VISIBLE ABSORPTION SPECTROSCOPY

Spectra were obtained on the Evolution 260 Bio spectrophotometer (Thermo Scientific) using a temperature controlled 1.0 mm sample cell and 1.0 mm water-filled reference cell. Sample and buffer were each measured in triplicate and the average buffer spectrum was subtracted from the average sample spectrum. The percentage of hypochromicity at 297 nm was calculated by dividing the difference between \(\epsilon_{297}\) for folded DNA and \(\epsilon_{297}\) for unfolded DNA by \(\epsilon_{297}\) for unfolded DNA.

2.6 ELECTROPHORETIC MOBILITY SHIFT ASSAY

RHAU\textsubscript{53-105} and either hTR\textsubscript{1-20} DNA\textsubscript{c1} or hTR\textsubscript{1-20} DNA\textsubscript{c2} were mixed in molar ratios ranging from 0:1 to 6:1 in 20 mM HEPES, pH 7.5, 100 mM KCl, 20% glycerol and 140 mM glycine. Final nucleic acid concentration was 200 nM in each solution. The samples were well-mixed, centrifuged at 13 000 rpm for 45 seconds and loaded onto 12% polyacrylamide TBE gels.
The gels were resolved by electrophoresis at 75V and 4°C for 2.5 hours. The gels were stained for 10 minutes using either SYBR Green II or N-methylmesoporphyrin IX and imaged using the FluorChem Q imaging system (Cell Biosciences).

2.7 CIRCULAR DICHOISM SPECTROPOLARIMETRY

Spectra were collected on the J-810 spectropolarimeter (Jasco Inc.) calibrated using (2.583 mM (1S)-(+) -camphor-10-sulfonic acid monohydrate; Alfa Aesar). 1.0 mm cell and integration times of 8 seconds were used for 350 to 220 nm, 1.0 mm cell and integration time of 32 seconds for 220 to 200 nm and 0.1 mm cell and integration time of 32 seconds for 200 to 180 nm. Sample concentration was 30 to 50 µM in the 1.0 mm cell and 200 µM in 0.1 mm cell. Sample and buffer were measured separately in triplicate and the average buffer spectrum was subtracted from the average sample spectrum. Nucleic acids were measured in 20 mM HEPES, pH 7.5, 100 mM KCl or in 20 mM sodium phosphate, pH 7.5, 100 mM KF. RHAU53-105 was measured in 10 mM HEPES, pH 7.5, I =154 mM (NaCl). Nucleic acid spectra were normalized by number of bases per unit volume and protein spectra were normalized by number of peptide bonds per unit volume. Melting curves for nucleic acids were obtained by recording ellipticity at 263 nm over a range of temperatures in 20 mM HEPES, pH 7.5, 100 mM KCl. The protein-quadruplex complexes were measured in 20 mM HEPES, pH 7.5, I=100 mM KCl and the spectra were normalized by number of nucleotide bases per unit volume.

2.8 DYNAMIC LIGHT SCATTERING

Data were collected using the Nano-S Dynamic Light Scattering system (Malvern Instruments Ltd., Malvern, UK) equipped with a 633 nm laser and using a 173° scattering angle.
All solutions were filtered using a 0.1 µm centrifugal filter (Millipore, USA) and allowed to equilibrate to 20°C for 20 minutes prior to data collection. Samples and buffer were measured separately in a 45 µL quartz cuvette at a range of concentrations. RHAU was measured in 10 mM HEPES, pH 7.5, I=154 mM and nucleic acids and quadruplexes in 20 mM HEPES, pH 7.5, 100 mM KCl. Measurements were collected in automatic mode and four measurements were obtained for each concentration.

2.9 SMALL ANGLE X-RAY SCATTERING

Data collection and data analysis were performed by Dr. Trushar Patel. Data was collected with a Rigaku 3-pinhole camera (S-MAX3000) on a 200-mm multi-wire 2D detector calibrated with gold particles (NIST, Standard Reference Material 8012, National Institute of Standards and Technology, Maryland, USA). The instrument was equipped with a Rigaku MicroMax+200 microfocus sealed tube and Confocal Max-Flux (CMF) optics. Cu-Kα radiation at 1.54 Å was generated and used to illuminate the sample. Buffer and sample at a range of concentrations were each exposed for 2 to 4 hours and the raw data was reduced using Rigaku’s SAXGUI data processing software. The momentum transfer, s, was defined as follows: 

\[ s = \frac{4\pi \sin \theta}{\lambda} \]

where \( \theta \) is the scattering angle and \( \lambda \) is the wavelength. Individual data sets were merged and radius of gyration (\( R_g \)) and maximum particle dimension (\( D_{\text{max}} \)) were calculated using the GNOM program.\(^{73}\) Ab initio modelling were carried out using the experimentally determined \( R_g \) and \( D_{\text{max}} \) parameters. Multiple low resolution solution conformations were generated by DAMMIN, a program that relies on temperature-dependent simulated annealing.\(^{74}\) Goodness of fit parameter (\( \chi \)) was calculated for each model to assess its quality and passing models were averaged using the DAMAVER program.\(^{75}\) HYDROPRO was used to calculate the
hydrodynamic parameters for the models. Following SAXS data collection, DLS measurements were carried out for each sample to check whether degradation occurred due to radiation exposure.

2.10 CRYSTALLIZATION TRIALS

In an effort to identify suitable conditions for crystallization of hTR$_{1-20}$ DNA$_{c1}$, hTR$_{1-20}$ DNA$_{c2}$, hTR$_{1-20}$ DNA$_{c1}$·RHAU$_{53-105}$ and hTR$_{1-20}$ DNA$_{c2}$·RHAU$_{53-105}$, crystallization trials were conducted using commercial crystallization kits from Hampton Research (CA, USA). The quadruplexes and complexes were prepared as described in the preceding sections of this report. Crystallization conditions were tested using the vapour diffusion method. Sitting drops composed of 1 µL of the species to be crystallized and 1 µL of reservoir solution were set-up in 96-well plates. Each well contained 50 µL of reservoir solution. All species were filtered using 0.2 µm filters prior to crystallization. The plates were covered with clear protective film and incubated at 20 °C. Crystal Screen HT and Natrix HT screening kits (Hampton Research, CA, USA) were used to test crystallization conditions for of hTR$_{1-20}$ DNA$_{c1}$ at 7.26 mg/ml and hTR$_{1-20}$ DNA$_{c2}$ at 5.41 mg/ml. Crystal Screen HT, PEG/ION HT, Index HT and Natrix HT screening kits (Hampton Research, CA, USA) were utilized to test possible crystallization conditions for hTR$_{1-20}$ DNA$_{c1}$·RHAU$_{53-105}$ and hTR$_{1-20}$ DNA$_{c2}$·RHAU$_{53-105}$. The complex of RHAU$_{53-105}$ with hTR$_{1-20}$ DNA$_{c1}$ was tested at 5.09 mg/ml and 8.65 mg/ml and the complex with hTR$_{1-20}$ DNA$_{c2}$ was tested at 5.41 mg/ml. The concentrations of the complexes refer to the concentration of RHAU$_{53-105}$ in the complex only. Concentration of the oligonucleotide component was determined spectrophotometrically and then the concentration of protein was calculated on the basis of equimolar binding between hTR$_{1-20}$ DNA and RHAU$_{53-105}$. 
To optimize the crystallization conditions for RHAU$_{53-105}$·hTR$_{1-20}$ DNA$_{c2}$ identified using the initial PEG/ION HT screen, forty-eight conditions were tested using the vapour diffusion method (Table 1). Hanging drops composed of 2 µL of the complex at 5.5 mg/ml (protein component) and 2 µl of the crystallization solution were set up in 24-well plates. Each well contained 1.00 ml of reservoir solution. The plates were incubated at 20 °C.

In addition, hTR$_{1-20}$ DNA$_{c2}$·RHAU$_{53-105}$ was further screened using Crystal Screen 1 and Crystal Screen 2 (Hampton Research, CA, USA). Hanging drops composed of 2 µL of complex containing 4.00 mg/ml of RHAU$_{53-105}$ and 2 µL of reservoir solution were prepared on slides and suspended above wells containing 1.00 ml of reservoir solution. The wells were sealed with grease and the plates were incubated at 20 °C. The crystal plates were observed daily for the first week, every three days for the next week, weekly for the subsequent month and monthly for the remainder of the time.
Table 1. Crystallization solutions screened in an effort to optimize crystals of RHAU$_{53-105}$·hTR$_{1-20}$ DNA$_{c2}$ initially obtained in 0.2 M formate salt, 20 % w/v PEG 3500.

<table>
<thead>
<tr>
<th>Plate</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>PEG 3350 18% w/v</td>
<td>PEG 3350 19% w/v</td>
<td>PEG 3350 20% w/v</td>
<td>PEG 3350 21% w/v</td>
<td>PEG 3350 22% w/v</td>
<td>PEG 3350 23% w/v</td>
</tr>
<tr>
<td></td>
<td>CHKO$_2$ 0.2 M</td>
<td>CHKO$_2$ 0.2 M</td>
<td>CHKO$_2$ 0.2 M</td>
<td>CHKO$_2$ 0.2 M</td>
<td>CHKO$_2$ 0.2 M</td>
<td>CHKO$_2$ 0.2 M</td>
</tr>
<tr>
<td>B</td>
<td>PEG 3350 18% w/v</td>
<td>PEG 3350 19% w/v</td>
<td>PEG 3350 20% w/v</td>
<td>PEG 3350 21% w/v</td>
<td>PEG 3350 22% w/v</td>
<td>PEG 3350 23% w/v</td>
</tr>
<tr>
<td></td>
<td>CHKO$_2$ 0.2 M</td>
<td>CHKO$_2$ 0.2 M</td>
<td>CHKO$_2$ 0.2 M</td>
<td>CHKO$_2$ 0.2 M</td>
<td>CHKO$_2$ 0.2 M</td>
<td>CHKO$_2$ 0.2 M</td>
</tr>
<tr>
<td></td>
<td>KCl 0.1 M</td>
<td>KCl 0.1 M</td>
<td>KCl 0.1 M</td>
<td>KCl 0.1 M</td>
<td>KCl 0.1 M</td>
<td>KCl 0.1 M</td>
</tr>
<tr>
<td>C</td>
<td>PEG 3350 18% w/v</td>
<td>PEG 3350 19% w/v</td>
<td>PEG 3350 20% w/v</td>
<td>PEG 3350 21% w/v</td>
<td>PEG 3350 22% w/v</td>
<td>PEG 3350 23% w/v</td>
</tr>
<tr>
<td></td>
<td>HCOONa 0.2 M</td>
<td>HCOONa 0.2 M</td>
<td>HCOONa 0.2 M</td>
<td>HCOONa 0.2 M</td>
<td>HCOONa 0.2 M</td>
<td>HCOONa 0.2 M</td>
</tr>
<tr>
<td>D</td>
<td>PEG 3350 18% w/v</td>
<td>PEG 3350 19% w/v</td>
<td>PEG 3350 20% w/v</td>
<td>PEG 3350 21% w/v</td>
<td>PEG 3350 22% w/v</td>
<td>PEG 3350 23% w/v</td>
</tr>
<tr>
<td></td>
<td>HCOONa 0.2 M</td>
<td>HCOONa 0.2 M</td>
<td>HCOONa 0.2 M</td>
<td>HCOONa 0.2 M</td>
<td>HCOONa 0.2 M</td>
<td>HCOONa 0.2 M</td>
</tr>
<tr>
<td></td>
<td>KCl 0.1 M</td>
<td>KCl 0.1 M</td>
<td>KCl 0.1 M</td>
<td>KCl 0.1 M</td>
<td>KCl 0.1 M</td>
<td>KCl 0.1 M</td>
</tr>
</tbody>
</table>

Plate 2

<table>
<thead>
<tr>
<th>Plate</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>PEG 3350 18% w/v</td>
<td>PEG 3350 19% w/v</td>
<td>PEG 3350 20% w/v</td>
<td>PEG 3350 21% w/v</td>
<td>PEG 3350 22% w/v</td>
<td>PEG 3350 23% w/v</td>
</tr>
<tr>
<td></td>
<td>NH$_4$HCO$_2$ 0.2 M</td>
<td>NH$_4$HCO$_2$ 0.2 M</td>
<td>NH$_4$HCO$_2$ 0.2 M</td>
<td>NH$_4$HCO$_2$ 0.2 M</td>
<td>NH$_4$HCO$_2$ 0.2 M</td>
<td>NH$_4$HCO$_2$ 0.2 M</td>
</tr>
<tr>
<td>B</td>
<td>PEG 3350 18% w/v</td>
<td>PEG 3350 19% w/v</td>
<td>PEG 3350 20% w/v</td>
<td>PEG 3350 21% w/v</td>
<td>PEG 3350 22% w/v</td>
<td>PEG 3350 23% w/v</td>
</tr>
<tr>
<td></td>
<td>NH$_4$HCO$_2$ 0.2 M</td>
<td>NH$_4$HCO$_2$ 0.2 M</td>
<td>NH$_4$HCO$_2$ 0.2 M</td>
<td>NH$_4$HCO$_2$ 0.2 M</td>
<td>NH$_4$HCO$_2$ 0.2 M</td>
<td>NH$_4$HCO$_2$ 0.2 M</td>
</tr>
<tr>
<td></td>
<td>KCl 0.1 M</td>
<td>KCl 0.1 M</td>
<td>KCl 0.1 M</td>
<td>KCl 0.1 M</td>
<td>KCl 0.1 M</td>
<td>KCl 0.1 M</td>
</tr>
<tr>
<td>C</td>
<td>PEG 3350 18% w/v</td>
<td>PEG 3350 19% w/v</td>
<td>PEG 3350 20% w/v</td>
<td>PEG 3350 21% w/v</td>
<td>PEG 3350 22% w/v</td>
<td>PEG 3350 23% w/v</td>
</tr>
<tr>
<td></td>
<td>Mg(HCO$_2$)$_2$ 0.2 M</td>
<td>Mg(HCO$_2$)$_2$ 0.2 M</td>
<td>Mg(HCO$_2$)$_2$ 0.2 M</td>
<td>Mg(HCO$_2$)$_2$ 0.2 M</td>
<td>Mg(HCO$_2$)$_2$ 0.2 M</td>
<td>Mg(HCO$_2$)$_2$ 0.2 M</td>
</tr>
<tr>
<td></td>
<td>KCl 0.1 M</td>
<td>KCl 0.1 M</td>
<td>KCl 0.1 M</td>
<td>KCl 0.1 M</td>
<td>KCl 0.1 M</td>
<td>KCl 0.1 M</td>
</tr>
<tr>
<td>D</td>
<td>PEG 3350 18% w/v</td>
<td>PEG 3350 19% w/v</td>
<td>PEG 3350 20% w/v</td>
<td>PEG 3350 21% w/v</td>
<td>PEG 3350 22% w/v</td>
<td>PEG 3350 23% w/v</td>
</tr>
<tr>
<td></td>
<td>CHKO$_2$ 0.2 M</td>
<td>CHKO$_2$ 0.2 M</td>
<td>CHKO$_2$ 0.2 M</td>
<td>CHKO$_2$ 0.2 M</td>
<td>CHKO$_2$ 0.2 M</td>
<td>CHKO$_2$ 0.2 M</td>
</tr>
</tbody>
</table>
2.11 TEST FOR EXPRESSION OF RHAU\textsubscript{53-1008} IN \textit{ESCHERICHIA COLI} (\textit{E. coli})

Expression of His\textsubscript{6}-tagged RHAU\textsubscript{53-1008} in pET-15b plasmid was attempted in \textit{E. coli} strain BL21(DE3) cells. A pre-culture containing 50 ml of LB media and 100 µg/ml ampicillin was inoculated and grown overnight at 37 °C with constant shaking at 200 rpm. In the morning, 1.5 ml of the pre-culture was used to inoculate another 50 ml of LB with ampicillin. The newly inoculated culture was incubated at 37 °C with constant shaking at 200 rpm until OD\textsubscript{600} reached 0.6, at which time expression was induced with addition of IPTG at 5 mM final concentration. The culture was incubated for two additional hours under the same conditions and sampled every half hour. The collected samples were then denatured by addition of 10% (w/v) sodium dodecyl sulfate and Laemmli buffer in a 1:1 ratio and heating at 95 °C for 4 minutes prior to analysis by SDS-PAGE.

Since expression of His\textsubscript{6}-RHAU\textsubscript{53-1008} in \textit{E. coli} strain BL21(DE3) proved unsuccessful, expression was carried out in \textit{E. coli} strain BL21(DE3) supplemented with the RIPL plasmid. \textit{E.coli} strain BL21(DE3)-RIPL cells were transformed with His\textsubscript{6}-RHAU\textsubscript{53-1008} in pET-15b plasmid by electroporation. The protocol described for expression of RHAU\textsubscript{53-105} in \textit{E. coli} strain BL21(DE3) cells was followed to express His\textsubscript{6}-RHAU\textsubscript{53-1008} in \textit{E. coli} BL21(DE3) RIPL cells. Following expression, the cells were collected by centrifugation and frozen at -80 °C.

2.12 PURIFICATION OF His\textsubscript{6}-RHAU\textsubscript{53-1008}

Purification of His\textsubscript{6}-RHAU\textsubscript{53-1008} was attempted by affinity chromatography using Ni-NTA agarose beads (Qiagen, USA). The frozen pellet was lysed in 40 ml of 20 mM Tris/Tris-HCl, pH 8.5, I=150 mM (NaCl). Serine protease inhibitors were added to the pellet to prevent protein degradation. The cells were lysed by homogenization followed by sonication. The
supernatant was separated from the pellet by centrifugation at 22,000 g for 30 minutes at 4 °C. Prior to addition of the lysate supernatant, the Ni-NTA agarose beads were washed with three volumes of 20 mM Tris/Tris-HCl, pH 8.5, I=150 mM (NaCl). The beads were pelleted by centrifugation at low speed for 2 minutes, the buffer was discarded, the lysate supernatant was added and the beads were incubated with the lysate supernatant for 1.5 hours. Then, the beads were pelleted and washed with two volumes of 20 mM Tris/Tris-HCl, pH 8.5, I=150 mM (NaCl) followed by two volumes of 20 mM HEPES, pH = 7.5, I=150 mM (NaCl). Elution of the His\textsubscript{6}-tagged protein was attempted using an imidazole gradient. The beads were successively washed with two volumes of 20 mM imidazole/imidazole HCl, pH 6.0, I=150 mM (NaCl), one volume of 250 mM imidazole/imidazole HCl, pH 6.0, I=150 mM (NaCl) and three volumes of 500 mM imidazole/imidazole HCl, pH 6.0, I=150 mM (NaCl). The beads were further washed and stored in 20% ethanol. The entire purification procedure was performed at 4 °C.

2.13 WESTERN BLOT FOR DETECTION OF RHAU\textsubscript{53-1008}

To check the success of purification of His\textsubscript{6}-RHAU\textsubscript{53-1008} by affinity chromatography, a western blot was performed using a mouse anti-RHAU antibody. A sample of the cell lysate as well as the samples of washing buffers collected during attempted purification of His\textsubscript{6}-RHAU\textsubscript{53-1008} were denatured and resolved by SDS-PAGE. A pre-stained molecular size ladder was loaded on the gel. The proteins were then transferred from the polyacrylamide gel to a nitrocellulose membrane soaked in 10 mM CAPS, pH 11.0, 10 % v/v methanol using semi-dry electroblotting at 0.11 A for 45 minutes. Following the transfer of the proteins to the nitrocellulose membrane, the membrane was blocked using Phosphate Buffered Saline with Tween 20 (PBS-T) with 5 % (w/v) Coffee Whitener (No Name Brand). After one hour incubation at room temperature, the
membrane was washed with PBS-T buffer for three consecutive 10 minute intervals. The membrane was then incubated for one hour with mouse anti-RHAU antibody in blocking solution, washed with PBS-T buffer and incubated for addition 1 hour with horseradish peroxidase (HRP) conjugated anti-mouse antibody in blocking solution. The membrane was again washed with PBS-T, incubated in Luminata Forte Western HRP substrate (EMD Millipore, USA) for 4 minutes at room temperature and imaged using a fluorescent detector.

2.14 CLONING OF RHAU$_{53-1008}$$^\Delta$14.pCEP-Pu.BM40

To obtain RHAU$_{53-1008}$$^\Delta$14.pCEP-Pu.BM40, RHAU$_{1-52}$ was first deleted in RHAU$_{1-1008}$$^\Delta$14 in pcDNA3 by in-vitro site-directed mutagenesis. RHAU$_{53-1008}$$^\Delta$14.pcDNA3 was copied and amplified from RHAU$_{1-1008}$$^\Delta$14.pcDNA3 using the polymerase chain reaction (PCR). The PCR reaction was performed using the Phusion High-Fidelity DNA Polymerase enzyme (Thermo-Scientific), which produces products with blunt ends. The following primers were used: 5’- CATCCCGGGCACCTGAAAGGCCGAAATCGGCATGTGG-3’ (forward) and 5’-AAGCTTGGGTCTCCCTATAGTGAGTCGTATTAATTTC-3’. The final reaction mixture contained the following components: 1X Phusion HF buffer (supplied by Thermo-Scientific), 0.1 mM each dNTP, 0.5 µM forward primer, 0.5 µM reverse primer, 100 ng template, 3% (v/v) dimethyl sulfoxide (DMSO) and 1 U of Phusion High-Fidelity DNA Polymerase. The components were mixed on ice and the PCR reaction was carried out in the S1000™ Thermal Cycler (BIO-RAD) using a three-step protocol. Initial denaturation lasted 3 minutes at 98 °C. Then, 32 cycles were completed under the following conditions: 15 sec denaturation at 98 °C, 30 sec annealing at 65 °C, 4 minutes and 15 seconds extension at 72 °C. The final extension step took place at 72 °C for 5 minutes. After completion of the PCR reaction, the PCR mixture was
loaded into a 0.8% agarose gel and resolved by gel electrophoresis. The plasmid corresponding to RHAU\textsubscript{53-1008}\textsuperscript{Δ14}.pcDNA3 was extracted from the gel using the GeneJET Gel Extraction Kit (Thermo-Scientific). The purified linear plasmid was circularized and ligated using T4 DNA Ligase (Thermo-Scientific). The ligation mixture was purified using the GeneJET PCR Purification Kit (Thermo-Scientific).

Then, RHAU\textsubscript{53-1008}\textsuperscript{Δ14} was copied and amplified from RHAU\textsubscript{53-1008}\textsuperscript{Δ14}.pcDNA3 by PCR. The forward primer 5’-
AGCCCGCGCTAGCTCATCATCACCACCACATGACGATGACGATAAGCATCCCGGGC
ACCTGAAAGGCCGCGAA -3’ contains a NheI restriction site followed by a His\textsubscript{6} tag and an enterokinase cleavage site. The reverse primer 5’-
GCCGCCCTCGAGTCAGCTGTAATATCCATCCTGGAATCGTGGCGG
ACCTGAAAGGCCGCGAA-3’ contains an XhoI restriction site. The PCR reaction contained the following components: 1X Phusion HF buffer (supplied by Thermo-Scientific), 0.1 mM each dNTP, 0.5 µM forward primer, 0.5 µM reverse primer, 100 ng template and 1 U of Phusion High-Fidelity DNA Polymerase. The reaction was performed the S1000\textsuperscript{TM} Thermal Cycler (BIO-RAD) under the following conditions: initial denaturation for 3 minutes at 98 °C, 32 cycles of 15 second denaturation at 98 °C, 30 second annealing at 65 °C and 2 minute extension at 72 °C, followed by a final 5 minute extension at 72 °C. After completion of the PCR reaction, the PCR mixture resolved by gel electrophoresis on a 0.8% agarose gel. The band corresponding to RHAU\textsubscript{53-1008}\textsuperscript{Δ14} was extracted from the gel using the GeneJET Gel Extraction Kit (Thermo-Scientific) and digested with NheI and XhoI restriction enzymes, which produce sticky ends upon DNA cleavage. The pCEP-Pu.BM40 vector was also digested with NheI and XhoI. Both RHAU\textsubscript{53-1008}\textsuperscript{Δ14} and pCEP-Pu.BM40 vector were purified using the GeneJET PCR Purification Kit (Thermo-Scientific) after completion of the digestion.
reactions, concentration of each species was measured spectrophotometrically and the two oligonucleotide segments were ligated using T4 DNA Ligase (Thermo-Scientific). The ligation mixture was incubated at room temperature for 15 minutes and used to transform z-competent *E. coli* strain DH5α cells. The transformed cells were grown on LB and ampicillin agar at 37 °C for 24 hours, individual colonies were sampled and grown in LB and ampicillin cultures overnight at 37 ° and the cells were lysed and plasmid extracted using the GeneJET Plasmid Extraction Kit (Thermo-Scientific). The size and sequence of the obtained plasmid was confirmed by gel electrophoresis and sequencing. All sequencing reactions were carried out by a commercial facility.
3. Results

3.1 DNA ANALOG OF THE hTR\textsubscript{1-20} SEQUENCE ASSUMES G-QUADRUPLEX SECONDARY STRUCTURE IN SOLUTION

In an effort to analyze the binding interaction between the RSM of RHAU and G-quadruplex oligonucleotides, synthetic DNA of the hTR\textsubscript{1-20} sequence was used. The DNA quadruplex was prepared in the presence of 100 mM KCl and purified on a preparative size exclusion column. hTR\textsubscript{1-20} DNA eluted as two peaks, suggesting that the two species differ in their conformations since they have identical molecular weights. Conformation 1, which eluted prior to conformation 2, comprised about two-thirds of the sample. To determine whether conformations 1 and 2 of the hTR\textsubscript{1-20} DNA quadruplex could interconvert, conformation 1 fractions collected during the first purification run were heated at 95 °C for ten minutes, allowed to passively cool to room temperature and purified on the preparative HiLoad Superdex 75 26/60 column (GE Healthcare). This procedure was repeated four times. The elution profiles for each of the four runs (Figure 3A) indicate that upon heating, a portion of hTR\textsubscript{1-20} DNA\textsubscript{c1} interconverts to hTR\textsubscript{1-20} DNA\textsubscript{c2}, although the proportion of hTR\textsubscript{1-20} DNA\textsubscript{c2} decreases with each subsequent heating-cooling cycle (Figure 3B).

The interconversion between hTR\textsubscript{1-20} DNA\textsubscript{c1} and hTR\textsubscript{1-20} DNA\textsubscript{c2} was further assessed using UV-Vis spectrophotometry. Figure 4A shows UV-Vis spectra collected for hTR\textsubscript{1-20} DNA at 20 °C, at 80 °C and at 20 °C following cooling to room temperature after heating at 80 °C. An increase in absorbance at 260 nm at 80 °C relative to 20 °C is indicative of G-quadruplex unfolding. Following the heating to 80 °C and cooling back to 20 °C, the sample produced a spectrum that is intermediate between the spectra obtained at 20 °C and at 80 °C.
The interconversion between hTR$_{1-20}$ DNA$_{c2}$ and hTR$_{1-20}$ DNA$_{c1}$ was also investigated by UV-Vis spectrophotometry. In contrast to hTR$_{1-20}$ DNA$_{c1}$, hTR$_{1-20}$ DNA$_{c2}$ produces identical UV-Vis spectra at 20 °C before and after heating at 80 °C, suggesting that the heating-cooling cycle did not induce a conformational change in the sample. Moreover, circular dichroism spectra for hTR$_{1-20}$ DNA$_{c2}$ were not affected by heating of the sample to 80 °C and cooling back to room temperature (Figure 4B). Neither UV-Vis spectrophotometry nor spectropolarimetry show evidence of conversion of hTR$_{1-20}$ DNA$_{c2}$ to hTR$_{1-20}$ DNA$_{c1}$.

The purified hTR$_{1-20}$ DNA was further analyzed on a Superdex 200 SEC column and conformation 1 eluted ahead of conformation 2 (Figure 5). Both conformations stained with SYBER Green II as well as with the quadruplex specific dye N-methylmesoporphyrin IX (Figure 6). In contrast, the hTR$_{1-43}$ RNA sequence that was mutated to prevent quadruplex formation did not stain with N-methylmesoporphyrin IX. UV-Vis measurements for hTR$_{1-20}$ DNA$_{c1}$ and hTR$_{1-20}$ DNA$_{c2}$ shown in Figure 7 indicate hyperchromicity at 260 nm and hypochromicity at 297 nm at 80°C in comparison to 20°C. The extinction coefficients at 297 nm for hTR$_{1-20}$ DNA$_{c1}$ and hTR$_{1-20}$ DNA$_{c2}$ decreased by 30.9% and 37.5%, respectively, at 80°C relative to 20°C. UV CD spectra for both conformations of hTR$_{1-20}$ DNA confirm the presence of quadruplex secondary structure (Figure 8). Both hTR$_{1-20}$ DNA$_{c1}$ and hTR$_{1-20}$ DNA$_{c2}$ were analyzed by spectropolarimetry in 20 mM HEPES, pH 7.5, 100 mM KCl and 20 mM sodium phosphate, pH 7.5, 100 mM KF. Due to substitution of HEPES and KCl with sodium phosphate and KF, it was possible to extend collection of circular dichroism data to lower wavelengths. The spectra in both buffers for each conformation of the quadruplex are comparable, except the lower wavelength maximum is shifted left (Figure 8). Both spectra show a maximum at 263 nm and a minimum at 242 nm. In addition, a second maximum is observed at 207 nm for hTR$_{1-20}$ DNA$_{c1}$ and at 209 nm
for hTR$_{1-20}$ DNA$_{c2}$. The two spectra also differ in the region between 280 and 315 nm, where conformation 1 produced a sloping shoulder that is absent for conformation 2. Peak dampening, indicative of the loss of the quadruplex structure was observed at 80°C for both conformations (Figure 9). Using the ellipticity measured at 263 nm, the melting temperature was estimated to be 70°C for hTR$_{1-20}$ DNA$_{c1}$ and 64°C for hTR$_{1-20}$ DNA$_{c2}$. The recorded melting curves are displayed in Figure 10.

SAXS experiments were performed to obtain low resolution structural information about hTR$_{1-20}$ DNA quadruplexes in solution. DLS measurements that were conducted prior to SAXS data collection reveal the presence of monodisperse species in solution (Figure 11). Each quadruplex conformation was analyzed by SAXS at multiple concentrations. The resulting data sets were merged to produce a scattering profile for each quadruplex species (Figure 12A). Pair distribution functions ($P(r)$), which describe the distances between all observed electron pairs in a molecule, were constructed using the scattering information. Bell-shaped $P(r)$ functions were obtained for each quadruplex and were used to extract the radii of hydration ($r_G$) and maximum distances of the molecules ($D_{max}$) (Figure 12B). The hydrodynamic parameters for the two conformations are similar. All hydrodynamic parameters are summarized in Table 1. Numerous low resolution models were calculated for each quadruplex and were later merged to produce an average surface envelope for each conformation. Each conformation of the quadruplex appears as a disk-like structure with two concave faces (Figure 12C). The normalized spatial discrepancy (NSD) values, ≤ 0.51 for both conformations of the quadruplex, indicate similarity between individually calculated models. In addition, $\chi$ values (0.8 for hTR$_{1-20}$ DNA$_{c1}$ and 0.9 for hTR$_{1-20}$ DNA$_{c2}$) suggest excellent agreement between experimentally obtained and model-based
hydrodynamic parameters. DLS measurements obtained after SAXS data collections showed absence of detectable degradation following exposure to radiation.
Figure 3. Purification of hTR$_{1-20}$ DNA by size exclusion chromatography. A. Elution profiles for hTR$_{1-20}$ DNA purified on a HiLoad Superdex 75 26/60 column (GE Healthcare). Following each run, the fractions corresponding to conformation 1 (peak on the left) were collected, heated at 95 °C for 10 minutes, cooled to room temperature and purified on the same column. This procedure was repeated four times. B. A bar graph showing the reduction in the amount of interconversion of hTR$_{1-20}$ DNA$_{c1}$ to hTR$_{1-20}$ DNA$_{c2}$ with each heating-cooling cycle.
Figure 4. The analysis of interconversion between the two conformations of hTR1-20 DNA by UV-Vis spectrophotometry and spectropolarimetry. UV-Vis spectra for hTR1-20 DNA\textsubscript{c1} (A) and hTR1-20 DNA\textsubscript{c2} (B) measured at 20 °C (solid line), 80 °C (dashed line) and 20 °C after sample was heated at 80 °C and passively cooled (dotted line). C. CD spectrum measured for hTR1-20 DNA\textsubscript{c2} at 20°C before (solid line) and after (dotted line) heating at 80 °C.
Figure 5. Elution profiles of N-terminal RHAU, hTR\textsubscript{1-20} G-quadruplexes and their complexes acquired from the Superdex 200 10/300 GL size exclusion chromatography column. The following colour scheme was used: RHAU\textsubscript{53-105}, red, solid; hTR\textsubscript{1-20} DNA\textsubscript{c1}, orange, solid; hTR\textsubscript{1-20} DNA\textsubscript{c2}, goldenrod, solid; RHAU\textsubscript{53-105}	extperiodcentered hTR\textsubscript{1-20} DNA\textsubscript{c1}, orange, dotted; hTR\textsubscript{1-20} DNA\textsubscript{c2}, goldenrod, dotted. All species except RHAU\textsubscript{53-105} were eluted in 20 mM HEPES, pH 7.5, \(I=100\) mM KCl. RHAU\textsubscript{53-105} was eluted in 10 mM HEPES, pH 7.5, \(I=154\) mM (NaCl).
Figure 6. Electrophoretic Mobility Shift Assay with RHAU$_{53-105}$ and hTR$_{1-20}$ DNA quadruplex.  

A. RHAU$_{53-105}$ shifts hTR$_{1-20}$ DNA$_{c1}$ under native conditions. The hTR$_{1-20}$ DNA$_{c1}$ stains with both SYBR Green II (left) and N-methylmesoporphyrin IX (right). The hTR$_{1-43}$ RNA mutated to prevent quadruplex formation only stains with SYBR Green II. B. RHAU$_{53-105}$ shifts hTR$_{1-20}$ DNA$_{c2}$ under native conditions. The hTR$_{1-20}$ DNA$_{c1}$ stains with both SYBR Green II (left) and N-methylmesoporphyrin IX (right). The hTR$_{1-43}$ RNA mutated to prevent quadruplex formation only stains with SYBR Green II.
Figure 7. Indication of G-quadruplex structure in UV-visible spectra obtained for hTR<sub>1-20</sub> DNA<sub>c1</sub> (orange) and hTR<sub>1-20</sub> DNA<sub>c2</sub> (goldenrod). The spectra were recorded at 20 °C (solid) and 80 °C in 20 mM HEPES, pH 7.5, I=100 mM KCl. Hypochromicity of 30.9% for hTR<sub>1-20</sub> DNA<sub>c1</sub> and 37.5% for hTR<sub>1-20</sub> DNA<sub>c2</sub> is evident at 297 nm (Inset). Slight hyperchromicity is seen at 260 nm for both quadruplexes.
Figure 8. Effect of buffer composition on the conformations of hTR$_{1-20}$ DNA$_{c1}$ (orange) and hTR$_{1-20}$ DNA$_{c2}$ (goldenrod). Circular dichroism spectra for each conformation of hTR$_{1-20}$ DNA in 20 mM HEPES, pH 7.5, 100 mM KCl (solid line) and 20 mM sodium phosphate, pH 7.5, 100 mM KF (dashed line) are compared.
Figure 9. Analysis of hTR_{1-20} DNA G-quadruplexes by spectropolarimetry. Far-UV CD spectra of hTR_{1-20} DNA_{c1} (orange) and hTR_{1-20} DNA_{c2} (goldenrod) and 20 °C (solid) and 80 °C (only standard deviations of three measurements are shown). Vertical error bars show standard deviations. All spectra were recorded in 20 mM HEPES, pH 7.5, I=100 mM KCl. Maxima and minima of the spectra are indicated by vertical lines.
Figure 10. Melting curves of the G-quadruplexes obtained by spectropolarimetry at 263 nm. The midpoints of transition indicated by vertical lines were estimated at ~70 °C for hTR$_{1-20}$ DNA$_{c1}$ (orange) and at ~64 °C for hTR$_{1-20}$ DNA$_{c2}$ (goldenrod).
Figure 11. Dynamic light scattering profiles of N-terminal RHAU constructs, quadruplexes and their complexes. The concentration (Conc.) dependence of the $r_H$ of RHAU$_{53-105}$ is shown in the inset. The following colour scheme was used: RHAU$_{53-105}$, red, solid; RHAU$_{53-150}$, black, solid; hT$_{1-20}$ DNA$_{c1}$, orange, solid; hTR$_{1-20}$ DNA$_{c2}$, goldenrod, solid; RHAU$_{53-105}$·hTR$_{1-20}$ DNA$_{c1}$, orange, dotted; hTR$_{1-20}$ DNA$_{c2}$, goldenrod, dotted. All species except RHAU$_{53-105}$ and RHAU$_{53-150}$ were analyzed in 20 mM HEPES, pH 7.5, I=100 mM KCl. RHAU$_{53-105}$ and RHAU$_{53-150}$ were analyzed in 10 mM HEPES, pH 7.5, I=154 mM (NaCl).
Figure 12. Solution conformation of hTR$_{1-20}$ DNA G-quadruplexes determine by SAXS. A. Scattering profiles of hTR$_{1-20}$ DNA$_{c1}$ (*orange*) and hTR$_{1-20}$ DNA$_{c2}$ (*goldenrod*). The merged measured data for each quadruplex is superimposed with the scattering profiles calculated based on the low resolution models in C. The data was collected in 20 mM HEPES, pH 7.5, I=100 mM (KCl). B. Pair distribution functions for the two G-quadruplexes. C. Average low resolution representations of the G-quadruplexes with $D_{max}$ indicated below the images.
3.2 RHAU\textsubscript{53-105} ADOPTS EXTENDED CONFORMATION WITH ORDERED SECONDARY STRUCTURE IN SOLUTION

RHAU\textsubscript{53-105} labelled with an N-terminal His\textsubscript{6} tag expressed in LB media, purified on the Talon\textsuperscript{®} cobalt column and ran as a single band at 8.3 kDa on a denaturing polyacrylamide gel (Figure 13A). Following His\textsubscript{6}-tag cleavage with thrombin, RHAU\textsubscript{53-105}, which is 6.4 kDa in size, ran with the 6.5 kDa marker (Figure 13B) and eluted as a single peak in 10 mM HEPES, pH 7.5, I=154 mM (NaCl) from a size exclusion column (Figure 5). The circular dichroism spectrum for RHAU\textsubscript{53-105} reveals a secondary structure with features of ordered structure and random coil (Figure 14). The CD signal was most prominent in the far UV spectrum and was absent beyond 250 nm. Circular dichroism spectra for RHAU\textsubscript{53-105} were measured in 10 mM HEPES, pH 7.0, I=154 mM (NaCl), 10 mM TRIS, pH 7.0, I=154 mM (NaCl) and 10 mM sodium phosphate, pH 7.0, I=154 mM (NaF). As shown in Figure 15, the spectra were identical in both HEPES and Tris buffers, but not in sodium phosphate. The proteins aggregated at high concentration in the sodium phosphate buffer, so it was not possible to obtain a meaningful circular dichroism spectrum for RHAU\textsubscript{53-105} in sodium phosphate below 200 nm. RHAU\textsubscript{53-105} was analyzed by DLS and SAXS at different concentrations. The protein was found to be monodisperse in solution (Figure 11). The $P(r)$ function obtained from the merged scattering profile was used to calculate $r_G$ and $D_{\text{max}}$ values of 3.25 (± 0.05) nm and 12.0 nm, respectively (Figure 16A-B, Table 2). The low resolution structure envelope reveals an elongated molecule with two bulges thought to correspond to two subdomains (Figure 16C). The NSD value of 0.51 (±0.02) and $\chi$ value of 1.2 presented in Table 1 point to agreement between individually calculated models as well as between experimentally measured and model based hydrodynamic parameters.
Figure 13. Expression and purification of RHAU$_{53-105}$. A. Image of 16% SDS-PAGE gel showing samples collected during expression (lanes 2-4 from left) and purification (lanes 5-11 from left) of RHAU$_{53-105}$. A band corresponding to RHAU$_{53-105}$ running at 8.3 kDa is present in the samples collected at 2 and 3 hours after induction of expression with IPTG and is absent at 0 hr. The cells were lysed in 20 mM sodium phosphate, 2 mM imidazole, 6 M guanidine hydrochloride, pH 7.0 and purified from the cell lysate by affinity chromatography. RHAU$_{53-105}$ was eluted in 150 mM imidazole, pH 7.0, I=300 mM (NaCl) (lanes 9-11 from left). B. Image of 16% SDS-PAGE gel showing RHAU$_{53-105}$ before (lane 2 from left) and after (lane 3 from left) the cleavage of the N-terminal His$_6$-tag with thrombin. Following His$_6$-tag cleavage, the protein was purified on a HiTrap benzamidine FF column (GE Healthcare) to remove thrombin.
Figure 14. Far UV-visible CD spectrum of RHAU_{53-105}. The spectrum was recorded at 20 °C in 10 mM HEPES, pH 7.5, I=154 mM (NaCl). Comparison of the recorded spectrum of RHAU_{53-105} (red) with reference spectra of secondary structure elements random coil (rc), α-helix (α) and β-strand (β) reveals evidence of ordered secondary structure in RHAU_{53-105}.\textsuperscript{77}
Figure 15. Effect of buffer composition on the conformation of RHAU$_{53-105}$. Circular dichroism spectra for RHAU$_{53-105}$ in 10 mM HEPES, pH 7.0, I=154 mM (NaCl) (solid line), 10 mM Tris, I=154 mM (NaCl) (dashed line) and 10 mM sodium phosphate, pH 7.0, I=154 mM (NaF) (dotted line).
Figure 16. N-terminal domain of RHAU adopts an extended ordered conformation in solution. A. Scattering profiles of RHAU$_{53-105}$ (red) and RHAU$_{53-150}$ (black). The merged measured data for each RHAU construct is superimposed with the scattering profiles calculated based on the low resolution models in C. The data was collected in 10 mM HEPES, pH 7.5, I=154 mM (NaCl). B. Pair distribution functions for N-terminal RHAU constructs RHAU$_{53-105}$ and RHAU$_{53-150}$ are indicative of extended conformations. C. Average low resolution representations of the N-terminal RHAU constructs with D$_{max}$ indicated beside the images. Superposition of the low resolution models of the two constructs allowed assignment of the N and C termini.
3.3 RHAU<sub>53-105</sub> BINDS hTR<sub>1-20</sub> DNA G-QUADRUPLE IN SOLUTION

The N-terminal construct of RHAU in complex with either hTR<sub>1-20</sub> DNA<sub>c1</sub> or hTR<sub>1-20</sub> DNA<sub>c2</sub> was analyzed in solution to elucidate the interaction between the RSM of RHAU and G-quadruplex. RHAU<sub>53-105</sub> was found to shift both hTR<sub>1-20</sub> DNA<sub>c1</sub> and hTR<sub>1-20</sub> DNA<sub>c2</sub> on a native polyacrylamide gel, suggesting that it binds both DNA quadruplex conformations (Figure 4). In addition, both RHAU<sub>53-105</sub>·hTR<sub>1-20</sub> DNA<sub>c1</sub> and RHAU<sub>53-105</sub>·hTR<sub>1-20</sub> DNA<sub>c2</sub>, each mixed in 1:1 molar ratios, eluted as single peaks from a size exclusion chromatography column (Figure 5). CD spectra obtained for both complexes are shown in Figure 17 along with the spectra for G-quadruplexes alone. The CD signal was influenced by both the protein component and the quadruplex below 250 nm and only the quadruplex component dominated the spectrum above 250 nm. The wavelengths of the maxima and minima remained unchanged from those observed for quadruplexes alone, but the horizontal shoulder to the right of the 263 nm maximum in hTR<sub>1-20</sub> DNA<sub>c2</sub> was transformed to a sloped shoulder in the complex of hTR<sub>1-20</sub> DNA<sub>c2</sub> and RHAU<sub>53-105</sub>.

SAXS data was collected for each complex at a range of concentrations and the merged scattering profiles (Figure 18A) were used to obtain two bell-shaped \( P(r) \) functions (Figure 18B). The low resolution structures appear as globular shapes without defined subdomains (Figure 14C). \( D_{max} \) values of 5.0 nm for RHAU<sub>53-105</sub>·hTR<sub>1-20</sub> DNA<sub>c1</sub> and 5.2 nm for RHAU<sub>53-105</sub>·hTR<sub>1-20</sub> DNA<sub>c2</sub> show that both complexes assume a compact conformation relative to the extended structure of RHAU<sub>53-105</sub> alone. The hydrodynamic parameters summarized in Table 1 are similar for both complexes. The NSD (≤0.56 for both complexes) and \( \chi \) (1.1 and 1.2 for complex with hTR<sub>1-20</sub> DNA<sub>c1</sub> and hTR<sub>1-20</sub> DNA<sub>c2</sub>, respectively) values reveal agreement between individual models used to calculate the merged average model for each complex as well as between experimentally derived and model-based parameters. DLS measurements (Figure 11) obtained
before and after SAXS data collection show the presence of monodisperse molecule in solution and absence of observable radiation damage after exposure to X-rays.
Figure 17. Conformational change associated with protein binding to DNA G-quadruplex. Far UV-visible CD spectra of hTR$_{1-20}$ DNA$_{c1}$ (orange, solid) and hTR$_{1-20}$ DNA$_{c2}$ (goldenrod, solid) to RHAU$_{53-105}$·hTR$_{1-20}$ DNA$_{c1}$ (orange, dotted) and RHAU$_{53-105}$·hTYR$_{1-20}$ DNA$_{c2}$ (goldenrod, dotted) were obtained at 20 °C in 20 mM HEPES, pH 7.5 and I=100 mM (KCl).
Figure 18. Solution conformations of G-quadruplex-protein complexes determined by SAXS. A. Scattering profiles of RHAU$_{53-105}$·hTR$_{1-20}$ DNA$_{c1}$ (orange) and RHAU$_{53-105}$·hTR$_{1-20}$ DNA$_{c2}$ (goldenrod). The merged measured data for each complex is superimposed with the scattering profiles calculated based on the low resolution models in C. The data for each complex was collected in 20 mM HEPES, pH 7.5, I=100 mM (KCl). B. Pair distribution functions for the two complexes. C. Average low resolution representations of the protein-G-quadruplex complexes with D$_{max}$ indicated beside the images.

A.  

B.  

C.  

RHA$_{53-105}$·hTR$_{1-20}$ DNA$_{c1}$

5.0 nm

RHA$_{53-105}$·hTR$_{1-20}$ DNA$_{c1}$

5.2 nm
Table 2. Summary of hydrodynamic data for the protein, quadruplexes and complexes. Errors are indicated in the parenthesis and dashes indicate values that could not be determined.

<table>
<thead>
<tr>
<th>Hydrodynamic parameter</th>
<th>RHAU&lt;sub&gt;53-105&lt;/sub&gt;</th>
<th>hTR&lt;sub&gt;1-20&lt;/sub&gt; DNA&lt;sub&gt;c1&lt;/sub&gt;</th>
<th>hTR&lt;sub&gt;1-20&lt;/sub&gt; DNA&lt;sub&gt;c2&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>Model-based&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Experimental</td>
<td>Model-based&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>&lt;sup&gt;a&lt;/sup&gt;r&lt;sub&gt;H&lt;/sub&gt;(nm)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.90 (±0.03)</td>
<td>1.90 (±0.02)</td>
<td>1.83 (±0.04)</td>
</tr>
<tr>
<td>&lt;sup&gt;b&lt;/sup&gt;r&lt;sub&gt;G&lt;/sub&gt;(nm)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.1 (±0.2)</td>
<td>2.2 (±0.2)</td>
<td>1.34 (±0.06)</td>
</tr>
<tr>
<td>&lt;sup&gt;c&lt;/sup&gt;D&lt;sub&gt;max&lt;/sub&gt;(nm)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>7.8</td>
<td>8.1 (±0.1)</td>
<td>3.5</td>
</tr>
<tr>
<td>&lt;sup&gt;d&lt;/sup&gt;r&lt;sub&gt;G&lt;/sub&gt;/r&lt;sub&gt;H&lt;/sub&gt;</td>
<td>1.1 (±0.2)</td>
<td>1.1 (±0.2)</td>
<td>0.71 (±0.08)</td>
</tr>
<tr>
<td>χ</td>
<td>-</td>
<td>1.1</td>
<td>-</td>
</tr>
<tr>
<td>NSD</td>
<td>-</td>
<td>0.58 (±0.06)</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hydrodynamic parameter</th>
<th>RHAU&lt;sub&gt;53-105·hTR&lt;sub&gt;1-20&lt;/sub&gt; DNA&lt;sub&gt;c1&lt;/sub&gt;</th>
<th>RHAU&lt;sub&gt;53-105·hTR&lt;sub&gt;1-20&lt;/sub&gt; DNA&lt;sub&gt;c2&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>Model-based&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Experimental</td>
</tr>
<tr>
<td>&lt;sup&gt;a&lt;/sup&gt;r&lt;sub&gt;H&lt;/sub&gt;(nm)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.7 (±0.1)</td>
<td>2.4 (±0.1)</td>
</tr>
<tr>
<td>&lt;sup&gt;b&lt;/sup&gt;r&lt;sub&gt;G&lt;/sub&gt;(nm)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.78 (±0.02)</td>
<td>1.9 (±0.1)</td>
</tr>
<tr>
<td>&lt;sup&gt;c&lt;/sup&gt;D&lt;sub&gt;max&lt;/sub&gt;(nm)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>5.0</td>
<td>5.5 (±0.05)</td>
</tr>
<tr>
<td>&lt;sup&gt;d&lt;/sup&gt;r&lt;sub&gt;G&lt;/sub&gt;/r&lt;sub&gt;H&lt;/sub&gt;</td>
<td>0.7 (±0.1)</td>
<td>0.8 (±0.2)</td>
</tr>
<tr>
<td>χ</td>
<td>-</td>
<td>1.2</td>
</tr>
<tr>
<td>NSD</td>
<td>-</td>
<td>0.56 (±0.02)</td>
</tr>
</tbody>
</table>

<sup>a</sup> The values were experimentally determined from DLS data with error obtained from linear regression analysis to infinite dilution for multiple concentrations.

<sup>b</sup> The values were experimentally determined from the SAXS data with error obtained from P(r) analysis by GNOM.

<sup>c</sup> The values were experimentally determined from SAXS data obtained from P(r) analysis by GNOM.

<sup>d</sup> Model-based parameters calculated from HEDROPRO with errors obtained as the S.D. from multiple models.
3.4 CRYSTALLIZATION TRIALS

Commercial crystallization screens Crystal Screen HT and Natrix HT (Hampton Research, CA, USA) were used to test one hundred ninety-two crystallization solutions for a condition suitable for crystallization of hTR$_{1-20}$ DNA quadruplexes. Conformation 1 of hTR$_{1-20}$ DNA did not yield any crystals. In contrast, fine needle-like crystals appeared in the drop of hTR$_{1-20}$ DNA$_{c2}$ at 5.41 mg/ml with 0.4 M potassium sodium tartrate tetrahydrate. The crystals, pictured in Figure 19, were first observed two months after the crystal screens were set-up. Because the crystals were too small to mount, they were not tested for X-ray diffraction.

Commercial crystallization screens Crystal Screen HT, Natrix HT, PEG/ION HT and Index HT (Hampton Research, CA, USA) were used to check three hundred eighty-four solutions for conditions appropriate for crystallization of the RHAU$_{53-105}$·hTR$_{1-20}$ DNA quadruplex complexes. The complex of RHAU$_{53-105}$ and hTR$_{1-20}$ DNA$_{c1}$ was tested at two concentrations (5.09 mg/ml and 8.65 mg/ml protein), but no crystals were obtained. In contrast, due to limited availability of hTR$_{1-20}$ DNA$_{c2}$, hTR$_{1-20}$ DNA$_{c2}$·RHAU$_{53-105}$ was only tested at 5.41 mg/ml (protein concentration). The following four conditions in the PEG/ION HT screen produced crystals: 0.2 M sodium formate, 20 % w/v PEG 3350; 0.2 M potassium formate, 20 % w/v PEG 3350; 0.2 M ammonium formate, 20 % w/v PEG 3350 and 0.2 M sodium formate pH 7.0, 20 % w/v PEG 3350. All crystals had the shape of rectangular prisms with smooth faces, but they dissociated into clusters of fine needle-like crystals upon removal of the protective film from the top of the well to check the crystals for X-ray diffraction. Figure 20A shows the image of the crystals obtained in 0.2 M sodium formate, 20 % w/v polyethylene glycol (PEG) 3350 after the well had been opened. To verify whether the crystals were composed of the protein-DNA complex or salt, an X-ray diffraction pattern was obtained (Figure 20B). The crystal was
cryoprotected in 24% PEG 3350 and 10% ethylene glycol and its ability to diffract X-rays was tested at three angles, for ten minutes at each angle. The diffraction pattern presented in Figure 20B shows that although the crystal diffracted to as far as 4.1 Å, the spots are smeared and weak, likely due to the dispersion of the original crystal into parallel fine needle-like crystals. The attempt to optimize the conditions with PEG 3350 in order to obtain crystals less prone to dissociation failed.

A subsequent screen of RHAU53-105·hTR1-20 DNA2 at 4.0 mg/ml (protein component) using commercial kits Crystal Screen 1 and Crystal Screen 2 (Hampton Research, CA, USA) identified 1.5 M sodium chloride, 10% v/v ethanol as a promising condition for crystallization of this complex. The crystal appeared two weeks after the screens were set-up. The crystal, shown in Figure 21A, appears to be composed of numerous microcrystals, so further optimization is necessary. Nevertheless, it stains with methyl blue (Figure 22B), suggesting it contains protein.
Figure 19. Image of crystals of hTR$_{1-20}$ DNA$_{c2}$ grown by vapour diffusion in 0.4 M potassium sodium tartrate tetrahydrate at 20 °C. Sitting drops were prepared by adding 1 µL of the reservoir solution to 1 µL of hTR$_{1-20}$ DNA$_{c2}$ at 7.26 mg/ml. The well contained 50 µl of the reservoir solution. The observed crystals appeared after two months.
Figure 20. A. Image of crystals of hTR₁-20 DNAₐ₂·RHAU₅₃-₁₀₅ grown by vapour diffusion in 0.2 M sodium formate, 20 % w/v polyethylene glycol 3350. Sitting drops were prepared by adding 1 µL of the reservoir solution to 1 µL of the complex (5.41 mg/ml protein concentration in complex). The well contained 50 µl of the reservoir solution. The observed crystals appeared after two months. B. X-ray diffraction pattern obtained for crystal shown in A after exposure to X-rays for 10 minutes. The crystal was cryoprotected in 24% polyethylene glycol 3350 and 10 % ethylene glycol. The resolution is 4.1 Å.
Figure 21. A. Image of crystals of hTR$_{1-20}$ DNA$_{c2}$-RHAU$_{53-105}$ grown by vapour diffusion in 1.5 M sodium chloride, 10 % v/v ethanol. Hanging drops were prepared by adding 2 µL of the reservoir solution to 2 µL of the complex (4.00 mg/ml protein concentration in complex). The well contained 1.00 mL of the reservoir solution. The observed crystals appeared after two weeks. B. The crystals stain with methyl blue.
3.5 EXPRESSION AND PURIFICATION OF THE FULL-LENGTH RHAU PROTEIN

In order to extend the biophysical analysis to the full length RHAU protein, expression of His<sub>6</sub>-RHAU<sub>53-1008</sub> was attempted in *E. coli*. RHAU<sub>53-1008</sub> is 110 kDa in size and, except for the first N-terminal fifty-two amino acids, represents the full length RHAU protein. The glycine-rich N-terminus is predicted to be disordered and was omitted due to concern over its ability to crystallize. The initial effort to express His<sub>6</sub>-tagged construct of the RHAU protein composed of amino acids 53 to 1008 in *E. coli* strain BL21(DE3) was unsuccessful (Figure 22A). To aid in expression, the His<sub>6</sub>-RHAU<sub>53-1008</sub>.pET-15b plasmid was transformed into *E. coli* strain BL21(DE3)-RIPL, which corrects for the codon bias in *E. coli* with extra copies of *argU*, *ileY*, *leuW* and *proL* tRNA genes. A band running immediately above the 100 kDa marker, comparable to the expected size of His<sub>6</sub>-RHAU<sub>53-1008</sub>, is evident in samples collected from the BL21(DE3)-RIPL culture, suggesting that RAHU<sub>53-1008</sub> was produced (Figure 22B). Then, purification was attempted by affinity chromatography using Ni-NTA agarose beads. Although RHAU protein was detected in the cell lysate (Figure 23), it was not present in any of the imidazole buffers used to wash the beads following incubation with the lysate in an attempt to elute any bound protein. Some RHAU protein was also detected in washing buffer fractions at the very top of the gel, suggesting that RHAU might have aggregated.

In an effort to produce the RHAU protein in a eukaryotic system, such as Human Embryonic Kidney 293T (HEK293T) cells, RHAU<sub>53-1008</sub>Δ<sub>14</sub> was cloned in the pCEP-Pu.BM40 vector optimized for overexpression of proteins in HEK293T cells. The RHAU<sub>53-1008</sub>Δ<sub>14</sub> construct is missing the nuclear localization signal located at amino acids 516 to 529. The nuclear localization signal was removed because it would have competed with the extracellular secretion signal contained in the pCEP-Pu.BM40 vector.
Figure 22. Expression of RAHU_{53-1008} in *Escherichia coli* cells. A. Samples collected during expression of His_{6}-tagged RAHU_{53-1008} in *E. coli* strain BL21(DE3) cells. The samples were mixed with reducing Laemmli buffer and heated at 95 °C for 4 minutes before loading on the gel. B. After induction of expression of His_{6}-tagged RAHU_{53-1008} in *E. coli* strain BL21(DE3)-RIPL cells with IPTG, samples were collected every thirty minutes for two hours. The samples were mixed with 10% sodium dodecyl sulfate and reducing Laemmli buffer in a 1:1 ratio and heated at 95 °C for 4 minutes before loading on the gel.
Figure 23. A. Purification of His$_6$-RHAU$_{53-1008}$ expressed in *Escherichia coli* BL21(DE3)-RIPL cells. A. An image of a 10% polyacrylamide gel resolved by electrophoresis showing the cell lysate and the fractions collected during purification of His$_6$-tagged RHAU$_{53-1008}$ protein by batch purification with nickel beads. The purification procedure was carried out under native conditions and protein elution was attempted with imidazole gradient. B. Western Blot for RHAU$_{53-1008}$ superimposed over the gel shown in A. RAHU$_{53-1008}$, indicated by the black arrow, was detected using a mouse anti-RAHU primary antibody and horse radish peroxidase conjugated anti-mouse secondary antibody.
4. Discussion

G-quadruplexes have been identified in key areas of the human genome, including telomeres, promoter regions of genes and 5’ and 3’-UTRs of mRNAs.\textsuperscript{6,18,22,24–26,78} Non-random distribution of G-rich sequences potentially capable of forming G-quadruplexes suggests that these structures are of physiological importance to gene metabolism.\textsuperscript{22} In fact, G-quadruplexes have been implicated in transcriptional and translational regulation, gene splicing, and disease.\textsuperscript{20} These highly stable structures must be unwound to allow for correct functioning of the genome. Although unwinding of double-stranded oligonucleotides has been studied in detail, little information is available about quadruplex recognition and unraveling. To date, RHAU has been identified as a major ATP-dependent G-quadruplex resolvase in human cells.\textsuperscript{52} The conserved N-terminal RSM motif was shown to be required for quadruplex recognition and binding, but not for quadruplex unwinding.\textsuperscript{56,79} Although RHAU has been shown to bind both DNA and RNA G-quadruplexes \textit{in vitro}, its main target in cell lysate was found to be a G-quadruplex in the 5’ terminus of hTR.\textsuperscript{52,53,55,56} The sequence upstream of the P1 helix in hTR contains several guanine runs and can form G-quadruplexes thought to serve a protective function during TERT maturation.\textsuperscript{56} The P1 helix, conserved among all known vertebrates except rodents, is a structural element required for boundary definition of the telomere RNA template sequence and its formation may be impeded by incorporation of nucleotides required for its assembly into upstream G-quadruplex structures.\textsuperscript{58,59}

The binding between an N-terminal construct of RHAU and the DNA quadruplexes formed by the hTR\textsubscript{1-20} sequence was investigated using an integrated biophysical approach. RHAU\textsubscript{53-105} and two conformations of the quadruplex were purified and were used to prepare RHAU\textsubscript{53-105}·hTR\textsubscript{1-20} DNA complexes in a 1:1 molar ratio. A DNA analog of the hTR\textsubscript{1-20}
sequence was used in this analysis to compare the structural features of the hTR\textsubscript{1-20} quadruplex to its DNA analogs and to determine whether the DNA analogs would be suitable for subsequent use in X-ray crystallography. The hTR\textsubscript{1-20} DNA was purified in two conformations in the presence of potassium cations. Since both conformations are of the same molecular weight, they likely fold differently from each other, resulting in a larger solvated volume for conformation 1. In fact, conformation 1 elutes prior to conformation 2 from a size exclusion column. The presence of quadruplex secondary structure was confirmed by CD, hypochromicity measurements, SAXS and staining with quadruplex specific dye N-methylmesoporphyrin IX. Hypochromicity of 30.9\% and 37.5\% was observed for hTR\textsubscript{1-20} DNA\textsubscript{c1} and hTR\textsubscript{1-20} DNA\textsubscript{c2}, respectively, at 297 nm at 80°C relative to 20°C. Such significant decrease in absorbance at 297 nm is indicative of G-quadruplex unfolding and is associated with the helix-coil transition.\textsuperscript{80} High melting temperatures (~70 °C for hTR\textsubscript{1-20} DNA\textsubscript{c1} and ~64 °C for hTR\textsubscript{1-20} DNA\textsubscript{c2}) further confirm that hTR\textsubscript{1-20} DNA assumes a G-quadruplex secondary structure in solution. These temperatures are higher than the melting temperature of double-stranded DNA helix, although they are lower than the 80 °C obtained for the melting temperature of the hTR\textsubscript{1-20} RNA G-quadruplex.\textsuperscript{81}

The possibility of interconversion between the two conformations of the hTR\textsubscript{1-20} DNA G-quadruplex was investigated by SEC, UV-Vis spectrophotometry and spectropolarimetry. After a heating-cooling cycle, a portion of conformation 1 was shown to refold into conformation 2, although the portion of conformation 2 formed decreased with each successive heating-cooling cycle. This finding suggests that hTR\textsubscript{1-20} DNA\textsubscript{c1} might be a mixture of more than one conformation that due to their similarities cannot be easily resolved on a HiLoad Superdex 75 26/60 column (GE Healthcare). In contrast, hTR\textsubscript{1-20} DNA\textsubscript{c2}, which makes up about one-third of
the hTR\textsubscript{1-20} DNA after the first heating-cooling cycle, does not change its conformation following additional heating and cooling. No interconversion of hTR\textsubscript{1-20} DNA\textsubscript{c2} to hTR\textsubscript{1-20} DNA\textsubscript{c1} was detected by either UV-Vis spectrophotometry or spectropolarimetry.

The two conformations of the quadruplex were differentiated by spectropolarimetry. The CD spectra for both hTR\textsubscript{1-20} DNA conformations are characteristic of parallel quadruplexes with characteristic maximum and minimum at \(\sim 260\) nm and \(\sim 240\) nm, respectively. In contrast, CD spectra for anti-parallel quadruplexes typically display a maximum at \(\sim 290\) nm and a minimum at \(\sim 260\) nm.\textsuperscript{82} In addition, the N-methylmesoporphyrin IX compound used to stain the hTR\textsubscript{1-20} DNA has previously been shown to be specific to parallel quadruplexes.\textsuperscript{11}

Recent analysis of a shorter 5’ hTR sequence, hTR\textsubscript{1-18}, in 100 mM K\textsuperscript{+} reveals similar findings regarding its conformation in solution.\textsuperscript{7} Martadinata \textit{et al.} studied hTR\textsubscript{1-18} by gel electrophoresis, UV, CD and NMR, although they did not present a high resolution structure of the quadruplex.\textsuperscript{7} In addition, they show that the quadruplexes formed by hTR\textsubscript{1-18} likely dimerize in solution by stacking at their 5’ termini.\textsuperscript{7} Since hTERT has been shown to function as a dimer \textit{in vivo}, the authors suggest that the observed quadruplex stacking might be involved in dimerization of TERT.\textsuperscript{7,83}

The hTR\textsubscript{1-20} alone and in complex with RHAU\textsubscript{53-105} was separately analysed in a comprehensive biophysical approach.\textsuperscript{81} The RNA quadruplex formed by hTR\textsubscript{1-20} assumed a single conformation in the presence of potassium cations. Like the two conformations of the hTR\textsubscript{1-20} DNA quadruplex, the hTR\textsubscript{1-20} quadruplex exhibits a minimum at 242 nm, although the maximum for the RNA occurs at 264 nm instead of 263 nm.\textsuperscript{81} The hydrodynamic parameters obtained for the DNA quadruplexes are similar to the parameters of the hTR\textsubscript{1-20} quadruplex for which \(r_H\), \(r_G\) and \(D_{\text{max}}\) were found to be 1.91 (\(\pm 0.02\)) nm, 1.31 (\(\pm 0.08\)) nm and 3.4 nm,
respectively. As for the DNA quadruplexes, the SAXS envelope for hTR\textsubscript{1-20} has a disk-like shape with two concave faces. However, the hTR\textsubscript{1-20}-RHAU\textsubscript{53-105} complex appears to be more extended ($r_H = 3.10 \pm 0.05$ nm, $r_G = 2.60 \pm 0.04$ nm and $D_{max} = 8.5$ nm) than either hTR\textsubscript{1-20} DNA\textsubscript{c1}·RHAU\textsubscript{53-105} or hTR\textsubscript{1-20} DNA\textsubscript{c2}·RHAU\textsubscript{53-105} complex. The SAXS envelope for the hTR\textsubscript{1-20}·RHAU\textsubscript{53-105} complex appears similar to the low resolution model for RHAU\textsubscript{53-105} alone, except the N-terminal region of the RHAU construct is enlarged. NMR analysis reveals the perturbation of same backbone residues in RHAU\textsubscript{53-105} upon binding to either hTR\textsubscript{1-20} or hTR\textsubscript{1-20} DNA analogs, suggesting that the 2’-OH group of RNA is not required for binding to RSM of RHAU.\textsuperscript{81} However, the absence of the 2’-OH group in hTR\textsubscript{1-20} DNA quadruplex likely permits an interaction between the quadruplex and a region of RHAU\textsubscript{53-105} downstream of the RSM, accounting for the compact shape of the low resolution models calculated for the hTR\textsubscript{1-20} DNA-RHAU\textsubscript{53-105} complexes based on SAXS data.\textsuperscript{81} This structure reveals that the N-terminus of RHAU\textsubscript{53-105} is involved in binding to the G-quadruplex. The N-terminus of the RHAU construct was identified by superposition of the low resolution models calculated from the SAXS data for RHAU\textsubscript{53-105} and RHAU\textsubscript{53-150}. Moreover, the low resolution model of the complex of hTR\textsubscript{1-20} and RHAU\textsubscript{53-105} suggests that G-tetrad and not the sugar phosphate backbone of the quadruplex likely participates in the binding to RHAU\textsubscript{53-105}.\textsuperscript{81}

The N-terminal construct of the RHAU protein was used in this analysis in order to study the binding between RHAU and the DNA analog of its quadruplex target formed by the hTR\textsubscript{1-20} sequence. RHAU\textsubscript{53-105} was chosen because it contains the RSM motif (amino acids 54-66) previously shown to be required for quadruplex binding.\textsuperscript{79} Moreover, CD spectra show that the construct possesses defined secondary structure. Although the full length RHAU has appreciably higher affinity for hTR\textsubscript{1-20} than RHAU\textsubscript{53-105}, the N-terminal construct was
nevertheless shown to bind hTR\textsubscript{1-20} and hTR\textsubscript{1-20} DNA by EMSA.\textsuperscript{57,81} In addition, His\textsubscript{6}-tagged-RHAU\textsubscript{53-105} was able to bind hTR in cell lysate depleted of endogenous full length RHAU, suggesting that the chosen construct is physiologically relevant.\textsuperscript{81}

Additionally, expression of RHAU\textsubscript{53-1008} was attempted in \textit{E.coli} strain BL21(DE3) supplemented with rare \textit{E.coli} codons, but purification of the protein by affinity chromatography was unsuccessful and evidence of aggregation was detected. Since RHAU is a fairly large human protein, it might be better suited for expression in eukaryotic rather than in bacterial cells. For this purpose, His\textsubscript{6}-RHAU\textsubscript{53-1008}\textsubscript{Δ14} was cloned into the pCEP-Pu.BM40 vector optimized for overexpression of proteins in HEK293T cells. Moreover, HEK293T cells produce post-translational modifications not possible in \textit{E. coli}, but which might nevertheless be physiologically relevant. Further research should in part focus on production and crystallization of full length RHAU protein.

A comprehensive biophysical approach was used to characterize the complex formed between DNA analogs of hTR\textsubscript{1-20} and an N-terminal construct of RHAU. The findings presented here confirm and further describe the previously established interaction between RHAU and hTR\textsubscript{1-20} DNA and compare the hydrodynamic parameters of the hTR\textsubscript{1-20}-RHAU\textsubscript{53-105} complex to the complex of RHAU\textsubscript{53-105} with the hTR\textsubscript{1-20} quadruplex DNA analog. Since the RSM of RHAU appears to interact with hTR\textsubscript{1-20} and hTR\textsubscript{1-20} DNA analogs in identical manner, it is worthwhile to pursue the attainment of the high resolution structure of the DNA complex. A high resolution structure depicting the RHAU\textsubscript{53-105}-hTR\textsubscript{1-20} DNA interaction is required to provide further insight into RHAU-quadruplex binding. Equally important, a high resolution structure of the full-length RHAU protein would yield new understanding of the mechanism of quadruplex unwinding by DEA-H-box helicases.
In order to determine suitable crystallization conditions for hTR1-20 DNA, G-quadruplexes and their complexes with RHAU53-105, several commercial crystallization screening kits from Hampton Research (CA, USA) were tested. Although neither hTR1-20 DNAc1 nor hTR1-20 DNAc2·RHAU53-105 crystallized in any of the screened solutions, crystals of hTR1-20 DNAc2·RHAU53-105 were obtained in 0.2 M sodium formate, 20 % w/v PEG 3350. These crystals diffracted X-rays weakly, but they also had the tendency to dissociate into parallel fine needle-like crystals unsuitable for good quality X-ray diffraction collection upon removal from the crystallization drop. Substitution of sodium with either potassium or ammonium also yielded hTR1-20 DNAc2·RHAU53-105 crystals that behaved similarly. An attempt to optimize this condition by varying the concentration of PEG 3350, pH and nature of the cation in the salt component of the crystallization solution failed to produce any usable crystals. However, in a subsequent screen for crystallization of hTR1-20 DNAc2·RHAU53-105, 1.5 M sodium chloride with 10 % v/v ethanol was revealed as a promising condition. Crystals that appeared two weeks after the screens were set-up absorbed methyl blue, suggesting that protein was present in them. This finding confirms that the observed crystals were not composed purely of salt. One crystal was tested for X-ray diffraction and was found to diffract X-rays weakly. Close examination of the crystal and its X-ray diffraction pattern suggests that it is composed of multiple microcrystals. Optimization of the 1.5 M sodium chloride with 10 % v/v ethanol condition is necessary. Furthermore, fine needle-like crystals were obtained for hTR1-20 DNAc2 in 0.4 M potassium sodium tartrate tetrahydrate, but they were too fine to measure for X-ray diffraction. Due to limited availability of hTR1-20 DNA, this condition was not optimized. Several promising condition for crystallization of hTR1-20 DNAc2 in complex with the N-terminal domain of the RHAU protein have been identified and are in need of further optimization. The N-terminal
domain investigated here is unique to RHAU and a high resolution structure of hTR$_{1-20}$ DNA$_{c2}$:RHAU$_{53-105}$ is anticipated for the input it might provide into the understanding of the interaction between the RSM motif of RHAU and G-quadruplex oligonucleotides.
5. Summary and Future Directions

G-quadruplexes are thermodynamically stable secondary structures that have the propensity to form within guanine-rich single stranded polynucleotides. G-quadruplexes consist of consecutive guanine tetrads associated into ring-like structures through Hoogsteen hydrogen bonds that stack on top of each other. G-quadruplex structures have been identified within telomeres, promoter regions of genes, 5’ and 3’ UTRs of mRNAs and implicated in transcriptional and translational control. Since G-quadruplexes have been shown to act as cis-regulatory element within the genome and formation of aberrant G-quadruplexes within single-stranded regions of DNA has been associated with disease, it is wise to investigate the mechanisms of G-quadruplex recognition and unwinding by specialized helicases. RHAU, which has been identified as a protein that binds and unwinds a G-quadruplex in the 5’ terminus of hTR, is a major source of G-quadruplex resolving activity in human cell lysates. The RSM region within the N-terminal domain of RHAU is responsible for recognition of G-quadruplexes by RHAU.

The binding between the DNA analog of hTR$_{1-20}$ and the N-terminus of RHAU, RHAU$_{53-105}$, was studied using biophysical methods. Gel electrophoresis, UV-visible spectroscopy, spectropolarimetry, dynamic light scattering and small angle X-ray scattering were used to conduct a comprehensive analysis of hTR$_{1-20}$ DNA, RHAU$_{53-105}$ and their complexes. The data show that hTR$_{1-20}$ DNA assumes two conformations in the presence of potassium ions in solution. Moreover, both conformations of hTR$_{1-20}$ assume parallel G-quadruplex secondary structure characterized by a disk-like shape with two concave faces. Formation of G-quadruplex is confirmed by staining with G-quadruplex specific dye N-Methylmesoporphyrin IX, significant hypochromicity at 297 nm at 80 °C relative to 20 °C, circular dichroism spectra with
characteristic minima and maxima at 242 and 263 nm, respectively, as well as estimated melting temperatures of ~70 °C for hTR\textsubscript{1-20} DNA\textsubscript{c1} and ~64 °C for hTR\textsubscript{1-20} DNA\textsubscript{c2}. Further, CD polarimetry and SAXS demonstrate that RHAU\textsubscript{53-105} assumes an extended ($D_{\text{max}} = 7.8$ nm, $r_G = 2.1$ (±0.2) nm) and ordered conformation in solution. Electrophoretic mobility shift assays show that RHAU\textsubscript{53-105} shifts both hTR\textsubscript{1-20} DNA\textsubscript{c1} and hTR\textsubscript{1-20} DNA\textsubscript{c2}. Complex formation between RHAU\textsubscript{53-105} and hTR\textsubscript{1-20} DNA was further confirmed by spectropolarimetry, dynamic light scattering and SAXS. The low resolution SAXS models for the complexes appear more compact than the extended model for RHAU\textsubscript{53-105} construct ($D_{\text{max}} = 5.0$ nm for complex with hTR\textsubscript{1-20} DNA\textsubscript{c1}; $D_{\text{max}} = 5.2$ nm for complex with hTR\textsubscript{1-20} DNA\textsubscript{c2}; $D_{\text{max}} = 7.8$ nm RHAU\textsubscript{53-105} alone). CD spectra of the complexes are dominated by the G-quadruplex component beyond 250 nm, suggesting that hTR\textsubscript{1-20} retains its G-quadruplex structure upon binding with RHAU\textsubscript{53-105}. These findings characterize hTR\textsubscript{1-20} DNA, RHAU\textsubscript{53-105} and their complexes, although a high resolution structure of the complex is required for a more specific elucidation of the this interaction. Currently, there is an ongoing attempt to crystallize the complex of RHAU\textsubscript{53-105} and hTR\textsubscript{1-20} DNA\textsubscript{c2} in order to obtain a high resolution structure of this complex by X-ray diffraction. Two conditions, 0.2 M sodium formate, 20 % (w/v) PEG 3350 and 1.5 M sodium chloride, 10 % (v/v) ethanol have been identified as suitable solutions for crystallization of hTR\textsubscript{1-20} DNA\textsubscript{c2}·RHAU\textsubscript{53-105}, but further optimization of these conditions is necessary. In addition, the gene for the full-length RHAU protein was cloned into the pCEP-Pu.BM40 vector to be used in expression of RHAU in HEK293T cells. Purification and analysis of the full length RHAU protein alone and in complex with hTR would contribute toward the understanding of the mechanisms involved in G-quadruplex binding and unwinding.
6. References


7. Appendix

7.1 LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALS</td>
<td>Amyotrophic lateral sclerosis</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>ATR</td>
<td>Alternatively translated region</td>
</tr>
<tr>
<td>BLM</td>
<td>Bloom syndrome protein</td>
</tr>
<tr>
<td>CD</td>
<td>Circular dichroism</td>
</tr>
<tr>
<td>c-KYT</td>
<td>Gene that encodes tyrosine-protein kinase Kit protein</td>
</tr>
<tr>
<td>c-MYC</td>
<td>Gene that encodes transcription factor Myc</td>
</tr>
<tr>
<td>CR</td>
<td>Conserved region</td>
</tr>
<tr>
<td>DHX36</td>
<td>Alternative name for RHAU protein</td>
</tr>
<tr>
<td>DLS</td>
<td>Dynamic light scattering</td>
</tr>
<tr>
<td>$D_{\text{max}}$</td>
<td>Maximum particle dimension</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>FANCJ</td>
<td>Fanconi anemia group J protein</td>
</tr>
<tr>
<td>FGF-2</td>
<td>Fibroblast growth factor 2</td>
</tr>
<tr>
<td>FTD</td>
<td>Frontotemporal dementia</td>
</tr>
<tr>
<td>His$_6$</td>
<td>Hexahistidine</td>
</tr>
<tr>
<td>HRP</td>
<td>Horseradish peroxidase</td>
</tr>
<tr>
<td>hTERT</td>
<td>Human telomerase reverse transcriptase</td>
</tr>
<tr>
<td>hTR</td>
<td>Human telomerase reverse transcriptase RNA</td>
</tr>
<tr>
<td>HuR</td>
<td>Human antigen R</td>
</tr>
<tr>
<td>IPTG</td>
<td>β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>IRES</td>
<td>Internal ribosome entry site</td>
</tr>
<tr>
<td>LB</td>
<td>Luria broth</td>
</tr>
<tr>
<td>miRNA134</td>
<td>Micro-RNA-134</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>NFAR1</td>
<td>Interleukin enhancer binding factor 3</td>
</tr>
<tr>
<td>NMM</td>
<td>N-methyl mesoporphyrin IX</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>NRAS</td>
<td>Neuroblastoma RAS viral (v-ras) oncogene homolog</td>
</tr>
<tr>
<td>NSD</td>
<td>Normalized spatial discrepancy</td>
</tr>
<tr>
<td>$P(r)$</td>
<td>Electron pair distribution function</td>
</tr>
<tr>
<td>p53</td>
<td>Tumor protein 53</td>
</tr>
<tr>
<td>P68</td>
<td>P68 RNA Helicase, also known as DDX5</td>
</tr>
<tr>
<td>P72</td>
<td>P72 RNA Helicase, also known as DDX17</td>
</tr>
<tr>
<td>PARN</td>
<td>Polyadenylate-specific ribonucleases</td>
</tr>
<tr>
<td>PEG</td>
<td>Polyethylene glycol</td>
</tr>
<tr>
<td>PITX1</td>
<td>Paired-like homeodomain 1 protein</td>
</tr>
<tr>
<td>$R_g$</td>
<td>Radius of gyration</td>
</tr>
<tr>
<td>$R_H$</td>
<td>Radius of hydration</td>
</tr>
<tr>
<td>RHAU</td>
<td>RNA associated with AU-rich element</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RSM</td>
<td>RHAU-specific motif</td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
</tr>
<tr>
<td>SAXS</td>
<td>Small angle X-ray scattering</td>
</tr>
<tr>
<td>SEC</td>
<td>Size exclusion chromatography</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SF2</td>
<td>Superfamily2</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris/Borate/EDTA (Ethylenediaminetetraacetic acid)</td>
</tr>
<tr>
<td>TERRA RNA</td>
<td>Telomeric non-coding ribonucleic acid</td>
</tr>
<tr>
<td>TERT</td>
<td>Catalytic component of telomerase reverse transcriptase holoenzyme</td>
</tr>
<tr>
<td>TNAP</td>
<td>Tissue-nonspecific alkaline phosphatase</td>
</tr>
<tr>
<td>TP53</td>
<td>Gene that encodes p53 protein</td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated region</td>
</tr>
<tr>
<td>UV-Vis</td>
<td>Ultraviolet-visible</td>
</tr>
<tr>
<td>WRN</td>
<td>Werner protein</td>
</tr>
<tr>
<td>YY1</td>
<td>Yin Yang 1 protein</td>
</tr>
<tr>
<td>( \chi )</td>
<td>Chi value</td>
</tr>
</tbody>
</table>