



# **INSTITUTO IMDEA NANOCIENCIA**

# UNIVERSIDAD AUTÓNOMA DE MADRID

## FACULTAD DE CIENCIAS

**TESIS DOCTORAL – PhD THESIS:** 

# The folding complexity of TERRA G-quadruplex unveiled at the single-molecule level

Irene Gutiérrez Pérez

Memoria presentada para optar al título de Doctor en Física de la Luz y de la Materia

Septiembre 2017

**Director: Dr. José Ricardo Arias González** Optical Nanomanipulation Lab Instituto IMDEA Nanociencia



A mi hermana

# Agradecimientos

Llegó el final de esta aventura que comenzó en septiembre de 2013. Durante estos años he tenido la suerte de conocer y contar con gente maravillosa que me ha ayudado no sólo a crecer profesionalmente sino también a nivel personal. Espero saber expresar mi gratitud a todos ellos en estas pocas líneas.

En primer lugar, quiero agradecer al Instituto Madrileño de Estudios Avanzados en Nanociencia (IMDEA Nanociencia) la financiación otorgada para realizar esta tesis doctoral. A mi director, Dr. J. Ricardo Arias González, por darme la oportunidad de trabajar en un campo totalmente distinto a mi formación previa y confiar en mi para desarrollar este proyecto. Gracias Ricardo por haber hecho todo lo posible porque llegase este día pese a las dificultades que encontramos en el camino, desgraciadamente la mayoría de ellas no científicas. También, no puedo dejar de dar las gracias al Dr. Borja Ibarra por haber hecho posible mi continuidad en IMDEA en un momento crítico de mi doctorado y por sus buenos consejos.

Querría resaltar también la inestimable ayuda del Prof. Alfredo Villasante en la solicitud de becas. Gracias Alfredo por transmitirme tu amor por la ciencia. Fue un placer conocerte y una pena que la vida no nos haya dejado seguir aprendiendo más de ti.

Mi más sincera gratitud al Prof. Carlos González por permitirme formarme en su laboratorio en el Instituto de Química-Física "Rocasolano" (IQFR-CSIC) junto a su maravilloso equipo, Dra. Irene Gómez Pinto y Dr. Miguel Garavís. Gracias Irene por tu ayuda en la búsqueda de los "plásmidos perdidos", ¡al final los encontramos! Mil y una veces gracias Miguel por enseñarme a sintetizar los G-quadruplexes, por tus consejos y por siempre estar dispuesto a ayudarme incluso cuando ya estabas realizando tu *postdoc* fuera del grupo. No sabes cuánto he valorado tus *e-mails* explicándome todas las dudas que me iban surgiendo con los experimentos. Ha sido un placer trabajar contigo.

También me gustaría agradecer su contribución a esta tesis al grupo de Dr. Fernando Falo y Dr. Alessandro Fiasconaro, de la Universidad de Zaragoza, por desarrollar un modelo teórico a partir de nuestros datos experimentales, y al de la Dra. Ana M. Pizarro, por la síntesis del compuesto de rutenio con el que se hicieron los experimentos preliminares de estabilización de G-quadruplexes de DNA.

Al Dr. Santiago Casado, por enseñarme a manejar el AFM y a la ya doctora María Eugenia Fuentes, antigua estudiante del grupo del Dr. Fernando Moreno, por su amabilidad y buenos consejos para llevar a cabo las medidas con dicho instrumento.

A todo el personal de IMDEA que ha hecho y/o sigue haciendo nuestro día a día más fácil, especialmente a Juan Carlos y a Antonio. También querría agradecer a Adriana Arnaiz su buena gestión de los equipos comunes para que todos podamos trabajar seguros y de manera organizada y a Warren Smith, thank you so much for your technical support and for being the best handyman!

A mi grupo en IMDEA, Dra. Sara de Lorenzo y Héctor Rodríguez. Sara, gracias no sólo por construir las "minis", pues sin ellas no hubiese podido terminar el proyecto, sino también por estar siempre ahí para ayudarme y apoyarme. Gracias por tu contribución en las medidas de más cadena aleatoria para que pudiese llegar a tiempo y por tus consejos para la escritura de la tesis. A Héctor, mi compi de doctorado, gracias por esas risas en el laboratorio, por tu amabilidad y por siempre estar dispuesto a compartir tus conocimientos con los demás. También me gustaría recordar a Elisa Poyatos, que estuvo realizando su trabajo fin de máster en el grupo en mi primer año de doctorado. Gracias Eli por tu bondad y honestidad. De verdad deseo que tengas suerte. Y, por supuesto, al antiguo *postdoc* de mi laboratorio, Dr. Elías Herrero Galán. Gracias Elías por enseñarme a trabajar con rigor y precisión en un laboratorio de bioquímica. Fue un verdadero placer hacer mi trabajo fin de máster bajo tu supervisión y he tenido muy presente tus consejos durante estos cuatro años. Ojalá encuentre a más gente como tú en mi futuro profesional.

A mis compañeros del laboratorio 161: Rebe, Fer, Kateryna y Carlos. Rebe, gracias por siempre sacar un hueco para resolver mis dudas de bioquímica y por tus consejos cuando me atascaba con los experimentos. A mi compi de mesa, Fer, con quien empecé casi a la vez en este mundo "pincero", gracias por todos los buenos ratos que hemos compartido, las canciones para los días tristes y tu ayuda con los programas de Matlab. De verdad, espero y deseo que te vaya genial en el doctorado. Aunque estuviésemos en grupos diferentes, no imagino un mejor compañero de aventura. A Kateryna, gracias por tu amabilidad y por la tranquilidad que transmites y a Carlos, nuestra última incorporación, gracias por tu alegría y las risas en esta última etapa de tantos nervios. ¡Mucha suerte con vuestras tesis!

También quiero mostrar mi agradecimiento a otros compañeros de IMDEA con los que, aun no trabajando de forma directa, han sido un gran apoyo en estos años. A Aitor, Sara y Patri P por compartir nuestros comienzos en IMDEA. Gracias Aitor por tu simpatía y tus consejos con el AFM. Sara, gracias por ayudarme a adaptarme a IMDEA, explicándome dónde estaba todo y cómo iban los equipos. Gracias Patri P por tus consejos y esos momentos en el AFM, compartiendo días buenos y, no tan buenos, de medidas. David, Patri B, Dani, Elena, Tere, Sofi, Alejandra, Maite, Jenny y Adrián fueron completando la familia imdeana. Gracias a todos por esos buenos ratos a la hora de la comida. ¡Mucha suerte y ánimo con vuestras tesis!

A Fran, por su sonrisa, sus abrazos y esos viajes en tren de vuelta a casa que servían de "terapia". Gracias por tu cariño, tu bondad, tu honestidad y tus buenos consejos. Sé que contigo me llevo un amigo de verdad, de esos que duran para toda la vida. A mi familia de Zaragoza, Mari Carmen y Javier, gracias por hacerme sentir como en casa y enseñarme lo bonito de vuestra ciudad.

A Felipe. Sin duda lo más bonito que me llevo de esta etapa de mi vida. Gracias por tu amor, por tu confianza en mí, por tus mil y un consejos, por siempre estar ahí y darme fuerza para continuar. No podría haber imaginado nada mejor: empezar, compartir y terminar esta etapa juntos. Esto es solo el principio de todo.

Y por último, a mi apoyo constante, mi familia. Gracias a mis abuelos, Adolfina y Justino, por su cariño y por preocuparse siempre por mí. Gracias a mis padres, Mari Carmen y Rafael, por su amor y confianza en mí, por haber hecho posible con su esfuerzo que haya llegado hasta aquí y por recordarme qué es lo importante en esta vida. Gracias papá por tus consejos, por siempre pensar que soy capaz de conseguirlo y por esos besos en los días tristes. Gracias mamá por creer en mí y por tus mil y un achuchones cuando llego a casa. Y, finalmente, a mi todo, mi otra mitad, mi mejor amiga y la mejor hermana que me podía dar la vida, mi Paula. Gracias por estar ahí siempre, por resolver mis mil y una dudas de biología y, aun en la distancia, hacerme sentir que estás conmigo y transmitirme que puedo lograrlo. Ojalá la vida te dé todo lo bonito que mereces. Daros las gracias me parece poco para todo lo que habéis hecho y seguís haciendo por mí. Os quiero mucho.

*«The future belongs to those who believe in the beauty of their dreams»* Eleanor Roosevelt

# Contents

RESUMEN	13
ABSTRACT	15
LIST OF ABBREVIATIONS	17
1 INTRODUCTION	19
1.1. Fundamentals of nucleic acids	19
1.1.1. The building blocks of nucleic acids	20
1.1.2. The Watson-Crick model of DNA	24
1.1.3. Non-B DNA structures	25
1.1.4. Structural and functional diversity of RNA molecules	30
1.1.4.1. Three-dimensional RNA structures	30
1.1.4.2. Functions of RNA structures	31
1.2. G-quadruplexes	32
1.2.1. Structure and topology	32
1.2.2. Location of potential G-quadruplex-forming sequences (PG4s)	37
1.2.2.1. DNA and RNA G-quadruplexes in human telomeres	37
1.2.3. Biological relevance	39
1.2.4. Applications	40
1.3. Mechano-chemistry and single-molecule manipulation techniques	42
1.3.1. Optical tweezers	43
1.3.1.1. Optical trap	44
1.3.1.2. Basic elements and up-to-date configurations	45
1.3.1.3. Single-molecule manipulation of nucleic acids	48
1.3.1.4. Limitations and drawbacks	51
1.3.2. Atomic force microscopy	52
1.3.2.1. Operating principle	52
1.3.2.2. Limitations and drawbacks	55

2	TERRA G-QUADRUPLEXES IMAGED BY ATOMIC FORCE MICROSCOPY	57
	2.1. Introduction	57
	2.2. Materials and methods	58
	2.2.1. Synthesis of TERRA molecules	58
	2.2.1.1. TERRA-25 molecules	58
	2.2.1.2. TERRA-25 molecules with double-stranded RNA handles	60
	2.2.1.3. TERRA-16 molecules	61
	2.2.2. AFM imaging	63
	2.2.3. Sample preparation for AFM measurements	64
	2.3. Results	64
	2.3.1. AFM imaging of TERRA-25	64
	2.3.2. AFM imaging of TERRA-25 with double-stranded RNA handles	65
	2.3.3. AFM imaging of TERRA-16 molecules	67
	2.4. Discussion	69
3	THE UNFOLDING DYNAMICS OF SHORT TERRA MOLECULES	71
	3.1. Introduction	71
	3.2. Materials and Methods	72
	3.2.1. Synthesis of TERRA molecules	72
	3.2.1.1. TERRA reference construct and RNA control systems 1 and 2	2 74
	3.2.1.2. RNA control system 3	77
	3.2.2. Synthesis of DNA analogue molecules	78
	3.2.3. Single-molecule experiments	80
	3.2.3.1. Optical Tweezers setups	80
	3.2.3.2. Force-induced unfolding experiments	82
	3.3. Results	83
	3.3.1. Single rupture events	85
	3.3.1.1. Comparison with the DNA analogous system	91
	3.3.1.2. Comparison with the RNA control systems	93
	3.3.2. Consecutive rupture events	100
	3.4. Discussion	106
4	COLLABORATIONS	111
	4.1. Mesoscopic model for DNA G-quadruplex force-induced unfolding	111

4.2. Stabilization of DNA G-quadruplexes by ruthenium complexes		
4.2.1. Materials and methods	119	
4.2.1.1. Synthesis of the ruthenium complex	119	
4.2.1.2. Synthesis of human telomeric DNA G-quadruplexes	120	
4.2.1.3. Synthesis of dsDNA molecules for control experiments	120	
4.2.1.4. Optical-tweezers experiments	121	
4.2.2. Preliminary results and discussion	122	
CONCLUSIONS		
CONCLUSIONES		
APPENDIX A: SIMULATIONS OF TERRA SEQUENCES BY MFOLD		
APPENDIX B: DNA G-QUADRUPLEX MESOSCOPIC MODEL: THEORY		
APPENDIX C: LIST OF PUBLICATIONS		
BIBLIOGRAPHY		

# Resumen

Los telómeros son estructuras nucleoproteicas que se encuentran en los extremos de los cromosomas lineales para protegerlos frente a la degradación. Dentro del genoma humano, se les considera los poseedores de la mayor cantidad de secuencias nucleicas propensas a formar estructuras tipo *G-quadruplex*, debido a sus extremos 3' de cadena sencilla ricos en guaninas (Gs).

Las estructuras tipo *G-quadruplex* se forman por el apareamiento, mediante enlaces tipo *Hoogsteen*, de cuatro hebras de ADN o ARN con una alta concentración de Gs. Concretamente, surgen del apilamiento de dos o más tétradas de Gs estabilizadas mediante cationes monovalentes, típicamente sodio o potasio. La demostración de su existencia *in vivo* en los telómeros sugiere que dichas estructuras no canónicas podrían desempeñar una función protectora en los mismos y que, por consiguiente, estarían involucradas en la estabilidad cromosómica.

Recientemente, se ha demostrado que los telómeros pueden ser transcritos, desde las regiones subteloméricas hacia los extremos de los cromosomas, en moléculas *TERRA* (del inglés, *TElomeric Repeat-containing RNA*). Al igual que sus equivalentes en ADN, estos tránscritos no codificantes podrían plegarse formando tándems de *G-quadruplexes* y desempeñar funciones relevantes a nivel celular, tales como la regulación de la longitud de los telómeros, la inhibición de la actividad enzimática de la telomerasa, la formación de la heterocromatina telomérica y/o la protección de los telómeros.

En esta tesis se ha investigado la estabilidad mecano-química de moléculas *TERRA* mediante las técnicas de microscopía de fuerzas atómicas (AFM; del inglés,

*Atomic Force Microscopy*) —en modo imagen— y pinzas ópticas (OT; del inglés, *Optical Tweezers*) —como técnica de espectroscopía de fuerzas a nivel de molécula individual—.

Mediante AFM, se ha logrado, por primera vez, la visualización directa de moléculas largas de TERRA en aire (*Capítulo 2*). Para ello, se ha desarrollado un protocolo bioquímico, riguroso y eficaz, que permite la comparación *in situ* de estructuras de ácidos nucleicos tipo cuádruplex y dúplex por medio de este tipo de microscopía. Además, a través de las imágenes obtenidas, se ha conseguido dilucidar la configuración adoptada por dichas moléculas sobre una superficie, lo cual ha resultado consistente con lo observado previamente para *G-quadruplexes* de DNA.

Mediante el uso de pinzas ópticas, se ha estudiado su dinámica de plegamiento, revelándose una inesperada complejidad estructural inherente a las especies de ARN (*Capítulo 3*). Concretamente, se han analizado moléculas de ARN que contienen entre cinco y siete repeticiones de la secuencia telomérica GGGUUA; es decir, secuencias que sólo pueden dar lugar a un único *G-quadruplex*. De acuerdo a los resultados obtenidos, el plegamiento de las moléculas *TERRA* en forma de *G-quadruplexes* se encuentra limitado por la capacidad inherente al ARN de autoplegarse sobre sí mismo de manera aleatoria, formando así estructuras más favorables a nivel entrópico. Este bloqueo estocástico confiere a estas moléculas un patrón de plegamiento bimodal, el cual no ocurre en sus análogos en ADN.

El bloqueo estocástico que sufren los *G-quadruplexes* de TERRA debido a la condensación del ARN puede ser relevante tanto para el entendimiento de sus posibles funciones reguladoras dentro de la célula, como para el diseño de fármacos con unión selectiva a dichas estructuras cuaternarias para el tratamiento de diversas enfermedades, incluyéndose entre ellas el cáncer.

# Abstract

Telomeres are nucleoprotein structures located at the end of linear chromosomes that are essential to protect them from degradation. In the human genome, they concentrate the highest amount of potential G-quadruplex-forming sequences due to their 3'-end G-rich single-stranded DNA overhangs.

A G-quadruplex (G4) is a four-stranded structure that arises from the stacking of two or more planar arrangements of four guanines (*G-tetrads*) stabilized by Hoogsteen hydrogen bonds and monovalent cations, typically sodium or potassium. Telomeric G4 visualization *in vivo* suggests that these non-canonical structures would have a role in telomere protection and, therefore, in chromosome stability.

Recently, it has been found that telomeres can be transcribed into <u>*TElomeric*</u> <u>*Repeat-containing*</u> <u>*RNA*</u> (*TERRA*) molecules. As their DNA counterparts, these noncoding transcripts are susceptible to fold into multiple G4 blocks and could perform various cellular functions, such as regulation of telomere length, inhibition of telomerase, telomeric heterochromatin formation and telomere protection.

In this thesis, TERRA molecules have been investigated at the single-molecule level by atomic force microscopy (AFM), in the imaging mode, and optical tweezers (OT), as a force-spectroscopic tool, to understand their mechano-chemical stability.

By AFM, a direct visualization of long TERRA molecules in ambient air has been obtained for the first time (*Chapter 2*). In this regard, a rigorous and efficient biochemical protocol that allows an *in situ* comparison of quadruplex and duplex nucleic acids structures has been developed. The resulting images have afforded the elucidation of the assembly mode of TERRA G4 tandems over a surface, finding consistency with earlier images of DNA G4s.

By OT, the folding dynamics of TERRA molecules has been characterized revealing an unexpected structural complexity inherent to the RNA species (*Chapter 3*). More in depth, RNA molecules with five to seven repeats of the G-rich TERRA sequence GGGUUA, thus capable of forming only one G4, have been analysed finding that the potential of RNA telomeric sequences to fold into a G4 conformation is limited by the intrinsic capacity of RNA to self-associate randomly and further condense into entropically more favourable structures. This behaviour confers on these short molecules a unique bistable-folding capacity that is not found in their DNA analogues.

The stochastic blockage of TERRA G4 assembly generated by RNA condensation may show relevant to understand the potential regulatory functions of TERRA molecules inside the cell, and in the design of G4-binding drugs for the treatment of several human diseases, including cancer.

# **List of Abbreviations**

А	Adenine
AFM	Atomic Force Microscopy
bp	Base-pair
С	Cytosine
DNA	Deoxyribonucleic acid
ds	Double-stranded
dsDNA	Double-stranded DNA
dsRNA	Double-stranded RNA
G	Guanine
G3	G-triplex
G4	G-quadruplex
mT	Minitweezers
MFP	Molecular Force Probe
МТ	Magnetic Tweezers
n	Number of telomeric repeats
Ν	Number of realizations
nt	Nucleotide
ОТ	Optical Tweezers
PDB	Protein Data Bank
PG4s	Potential G-quadruplex-forming sequences
POT1	Protection Of Telomeres protein 1
PSD	Position-Sensitive Detector
Rap1	Rhoptry associated protein-1
RNA	Ribonucleic acid
RuD	Ruthenium Dimer
SMT	Single Molecule Technique
Sp6 promoter	5' - ATTTAGGTGACACTATAG - 3'
SS	Single-stranded
ssDNA	Single-stranded DNA

ssRNA	Single-stranded RNA
Т	Thymine
TERRA	<u>Te</u> lomeric <u>R</u> epeat-containing <u>R</u> N <u>A</u>
TERRA-16	TERRA molecules with 16 GGGUUA repeats
TERRA-25	TERRA molecules with 25 GGGUUA repeats
TIN2	TRF1 Interacting Nuclear Factor 2
TPP1	Tripeptidyl peptidase I
TRF1	Telomeric Repeat-binding Factor 1 protein
TRF2	Telomeric Repeat-binding Factor 2 protein
U	Uracil
UTR	Untranslated Region

# Chapter 1 Introduction

## 1.1. Fundamentals of nucleic acids

Nucleic acids are key biological macromolecules in the origin and maintenance of life. In particular, they are responsible for storing, transmitting and expressing genetic information from one generation to the next. The structure of every protein, and ultimately of every biomolecule and cellular component, is a product of the information programmed into a specific nucleic acid sequence.

Two types of this material basis of genetic inheritance are distinguished in nature: deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). Although both of them are capable of carrying and preserving information using a four-letter alphabet, DNA was the selected molecule by most organisms, leaving RNA confined to some viral genomes. Over the past few years, RNA has emerged as a far more relevant molecule than previously expected thanks to the discovery of RNA interference [1-3].

Beyond their biological relevance, DNA and RNA have also become powerful platforms for nanotechnological applications because of their exceptional structural and functional diversity: from compatible biomaterials or new treatments for several genetic diseases to potential alternatives to current data storage systems [4-13].

This section provides an overview of the chemical nature of nucleic acids as well as a detailed examination of their structures and functions.

## 1.1.1. The building blocks of nucleic acids

Nucleic acids, DNA and RNA, are formed by the polymeric arrangement of four different subunits, called nucleotides, covalently connected through phosphodiester bonds. These monomeric subunits could be further divided into phosphate groups and four distinct nucleosides. The latter were subsequently identified as consisting of a pentose sugar and one of four nitrogen-containing heterocyclic bases. Thus, each repeating unit in a nucleic acid polymer comprises these three characteristic components linked together: a phosphate group, a pentose and a nitrogenous base (*Fig.1.1*).



**FIGURE 1.1.** *Structure of nucleotides.* General structure of a **(A)** deoxyribonucleotide and **(B)** ribonucleotide, showing their three units: phosphate group, nitrogenous base and pentose ring with its numbering convention.



**FIGURE 1.2.** *Types of pentoses in nucleic acids.* **(A)** 2'-deoxy-D-ribose, in the case of DNA and **(B)** D-ribose, for ribonucleotide units. Numbering convention is also specified.

Pentoses are cyclic monosaccharides of five carbon atoms. Nucleic acids contain two kinds of them: 2'-deoxy-D-ribose in the case of DNA and D-ribose for ribonucleotide units of RNA. Both types of pentoses are in their  $\beta$ -furanose form

(closed five-membered ring) and their carbon atoms are numbered as 1', 2', 3'... to distinguish them from those of the nitrogenous bases (*Fig. 1.2*).

The nitrogenous bases are planar aromatic heterocyclic compounds derived from pyrimidine (based on a single aromatic ring) or purine (based on two connected aromatic rings). Both DNA and RNA contain two major purine bases, adenine and guanine, and two major pyrimidines. In both nucleic acids, one of the pyrimidines is cytosine, but the second one is not the same: it is thymine in DNA and uracil in RNA. Their structures are shown in *Figure 1.3*.



**FIGURE 1.3.** *Major forms of nitrogenous bases in nucleic acids.* **(A)** Purine and **(B)** pyrimidine- derived bases showing the numbering convention.

The base of each nucleotide is covalently linked to the sugar through a  $\beta$ -N-glycosidic bond.  $\beta$  means that the base is above the plane of the sugar when viewed onto the plane and, therefore, on the same face of the plane as the 5' hydroxyl substituent. In pyrimidine bases, this bond is created between *N*-1 and the sugar *C*-1' atom whereas in the purines, *N*-9 is the nitrogen atom involved. The formation of this linkage provokes the release of a water molecule, which is generated from a hydroxyl group of the pentose and a hydrogen atom of the base (*Fig. 1.4*).



**FIGURE 1.4.** *Formation of*  $\beta$ *-N- glycosidic linkage between a nitrogenous base and a pentose.* Here, it is shown how an adenine bonds to a deoxyribose by a  $\beta$ -N- glycosidic linkage, releasing a water molecule.

The successive nucleotides of DNA and RNA are covalently linked by phosphodiester bonds between the 5'-hydroxyl group of a pentose and the 3'-hydroxyl group of the following, conferring to the nucleic acid chains a specific polarity and differentiated 5' and 3' ends. By definition, the 5' end lacks a nucleotide at the 5' end position and the 3' end, at the 3' position. Other groups (most often one or more phosphates) may be present on one or both ends. Therefore, the covalent chains of nucleic acids could be described as a backbone of alternating phosphate and pentose residues with different nitrogenous bases attached at regular intervals, giving rise to a particular sequence in which the genetic information is encoded (*Fig. 1.5*).



**FIGURE 1.5.** *Phosphodiester linkages in the covalent backbone of DNA and RNA.* The phosphodiester bonds linking successive nucleotide units are highlighted with light blue, dashed lines.

### 1.1.2. The Watson-Crick model of DNA

Under the name of "*nuclein*", DNA was firstly isolated and characterized from leukocyte cell nuclei by Friedrich Miescher in 1868. In the 1930s, Albrecht Kossel and Phoebus Levene established that nuclein was the deoxyribonucleic acid, defined as a molecule with different repetitive subunits. However, its function as a genetic information carrier was not discovered until the 1940s: firstly, with the work of Oswald T. Avery, Colin MacLeod and Maclyn McCarty in 1944 and, then, in 1952, with the experiments by Alfred D. Hershey and Martha Chase [14, 15].

At the same time, hard efforts were being done to elucidate the molecular structure of DNA. Erwin Chargaff and his colleagues found that in DNA the molar ratios of adenine:thymine and cytosine:guanine were both unity [16]. Meanwhile, Rosalind Franklin and Maurice Wilkins unravelled the helical structure of DNA using X-ray diffraction [17]. In 1953, taking into account all of these findings, James Watson and Francis Crick postulated a three-dimensional model of the DNA structure [18].

The **Watson-Crick model of DNA**, also referred to as B-form DNA or B-DNA, consists of two antiparallel helical DNA chains wound around the same axis to form a right-handed double helix (*Fig.1.6*). The hydrophilic backbones of alternating deoxyribose and phosphate groups are on the outside of the double helix, facing the surrounding water, whereas the hydrophobic purine and pyrimidine bases of both strands remain stacked inside the double helix due to a combination of van der Waals and dipole-dipole interactions between them. In addition to the base-stacking interactions, the DNA double helix maintains a regular structure thanks to hydrogen bonding between complementary base pairs: A=T, connected through two hydrogen bonds, and G≡C, through three. Another characteristic features are the opposite and unequally sized grooves that run continuously along the entire length of the molecule chain. So-called *minor* and *major grooves*, they appear because of the offset pairing of the two DNA strands.

Regarding the periodicities observed in the X-ray diffraction patterns of DNA fibres by Franklin and Wilkins, Watson and Crick proposed a structure in which the vertically stacked bases inside the double helix would be 3.4 Å apart and justified the secondary repeat distance of about 34 Å in the presence of 10 base pairs in each complete turn of the double helix. However, subsequent measurements in aqueous solution, revealed 10.5 base pairs per helical turn [19].



**FIGURE 1.6.** *Watson-Crick model for the structure of DNA.* Schematic representation of a double DNA helix, showing its dimensions. The offset pairing of the two strands creates two types of grooves on the surface of the duplex, whose names reflect their differences in size. Right insets represent the hydrogen-bonding patterns (light blue dashed lines) between the complementary base pairs A=T and G=C.

## 1.1.3. Non-B DNA structures

B-DNA is the most stable structure for a random sequence DNA molecule under physiological conditions. However, structural transitions into other DNA forms can occur induced by changes in environmental conditions, protein binding or superhelical tension.

**A-form** appears when the relative humidity is reduced to less than about 75%. As B-DNA, it is a right-handed double helix but wider and shorter. It measures 2.3 nm wide and has 11 base pairs per turn. Moreover, its base pairs are tilted about 20° to the helix axis rather than being perpendicular as in B-form. These structural changes deepen the major groove while making the minor groove shallower [20-23]. Biochemical, crystallographic and computer simulation analyses of the A-DNA structure as well as protein-DNA complexes indicate that this conformation may either form upon binding of certain proteins to DNA, or be an important intermediate step in forming the strongly distorted DNA conformation observed within at least some complexes with proteins [24-26]. Recently, single-molecule experiments have proved that the A-form is not an intrinsic conformation of dsDNA, unlike in dsRNA (see next subsection). In particular, these experiments have found no cooperative B-A transition, which should be promoted by a bistability of the molecule itself [27-29]. A contour-length reduction compatible with the so-called A-DNA may be then stabilized during protein binding or by the DNA condensation involved at low humidity in bulk experiments [30].

**Z-form**, whose name is due to the zigzag appearance of the backbone, is adopted by DNA segments with alternating sequence of purine and pyrimidine bases, especially C and G or 5-methyl-C and G residues. This double helix is longer and narrower than B-form; it measures 3.8 nm wide and has 12 base pairs per turn and, unlike A and B-forms, is left-handed. High salt concentrations are required to minimize electrostatic repulsion between the backbone phosphates, which are closer to each other than in A- and B-DNA. These Z-DNA tracts may play a role in regulating the expression of some genes or in genetic recombination, but it is still under investigation [31].



**FIGURE 1.7.** *Comparison of A, B and Z-forms of DNA.* **(A)** Crystal structure of an A-DNA dodecamer featuring an alternating pyrimidine-purine sequence (PDB ID: 5MVT) [32]. **(B)** Structure of a B-DNA dodecamer, d(CpGpCpGpApApTpTpCpGpCpG) (PDB ID: 1BDNA) [33]. **(C)** Structure of a Z-DNA dodecamer, generated using a computer program from standard coordinates [34]. Nitrogenous bases are represented in light blue and the backbone of alternating phosphate and pentose residues in dark blue.

In addition to Z-form, another non-B DNA structures related to rare sequences have been identified. Hairpins, cruciforms, triplex or quadruplex structures are some examples of that.

**Hairpins and cruciforms**, also known as stem-loop structures, arise from palindromic DNA regions; that is, DNA fragments with complementary inverted repeats (*Fig. 1.8*). In the formation of a hairpin, only a single DNA strand is needed so it can be produced during replication, bacterial conjugation, natural transformation or viral infections. To create a cruciform, two palindromic strands are required and its stability is enhanced by DNA supercoiling. This latter stem-loop structure is fundamentally important for a wide range of gene biological processes, including replication, regulation of gene expression, nucleosome structure and recombination. Moreover, it is implicated in the evolution and development of some diseases, such as cancer and Werner's syndrome [35, 36].



**FIGURE 1.8.** *Schematic representation of hairpin and cruciform structures.* Palindromic DNA (or RNA) sequences can form unusual structures with intrastrand base pairing: **(A)** hairpins, when only a single DNA strand is involved or **(B)** cruciform, if both strands participate.

Other non-canonical DNA structures involve three or even four DNA strands. Nucleotides participating in a Watson-Crick base pair (A=T, G=C) can form additional hydrogen bonds, particularly with functional groups arrayed in the major groove, giving rise to *Hoogsteen-type base pairing*. This special type of

hydrogen bonding allows pairing between, for example, a thymine with an A=T pair, T=A\*T, or a cytidine residue (if protonated) with a G $\equiv$ C nucleotide pair, C $\equiv$ G\*C+ (*Fig. 1.9 (A)*). **Triple-stranded (***triplex***) DNA structures**; that is, B-form double helices with a third strand coupled in their major grooves, are generated using this non-Watson-Crick base-pairing [37-40]. Four DNA strands can also pair in that way forming a **quadruplex**, but this occurs readily only for sequences with a very high proportion of guanosine residues at physiological conditions (see *Section 2*). Cytosine-rich DNA sequences are able to self-fold into tetraplex structures, named *i-motif*, at slightly acidic conditions [41, 42].



**FIGURE 1.9.** *Hoogsteen-type base pairing.* Common base-pairing patterns in: **(A)** a DNA triplex,  $T=A^*T$  (left) and  $C\equiv G^*C^+$  (right), in which the Hoogsteen pair is highlighted in blue; **(B)** a guanosine tetraplex, and **(C)** in an i-motif arrangement structure,  $C^*C^+$ . Hydrogenbonds are represented in red dashed lines.

### 1.1.4. Structural and functional diversity of RNA molecules

The relationship between structure and function in RNA remains unclear. More and more scientists are aware of the importance of elucidating this issue, considering it as the next step to be solved on the road to understand the molecular basis of life. Here, a brief review about the structural complexity of RNA and their large number of functional roles is presented.

### 1.1.4.1. Three-dimensional RNA structures

All RNA molecules, except the RNA genomes of certain viruses, are produced during transcription from information permanently stored in DNA. Basically, an enzyme system converts the genetic information in a segment of double-stranded DNA into a single-stranded RNA molecule with a base sequence complementary to one of the DNA chains. Although only differing in that RNA has a hydroxyl group at the 2' position of the pentose, and uracil instead of thymine (see *subsection 1.1.1*), the structural diversity of RNA is much greater than in DNA. [43]

**Single-stranded RNA molecules** tend to assume a right-handed helical conformation dominated by base-stacking interactions, which are stronger between two purines than between a purine and pyrimidine or between two pyrimidines [43]. Nevertheless, more complex RNA structures are found in nature, from hairpin loops created by self-complementary regions of a single strand (see *Fig.1.8*) to structures that involve two, three or four strands [44-47].

**Double-stranded RNA structures**, either hybrids or not [48, 49], follow the same base pairing pattern as DNA (G $\equiv$ C and A=U, or with occasional T residues in some RNAs), but also including the interaction between G and U residues when complementary sequences in both strands pair with each other. If both strands pair perfectly, they adopt an A-form right handed conformation —B-form has not been observed and Z-form has only been obtained synthetically under very high salt or high temperature conditions—. Recent experiments by optical and magnetic tweezers, besides AFM, have confirmed, on a one-by-one basis, this type of

arrangement [50] in contrast to DNA, whose A-like conformation has only been observed in bulk experiments where aggregation cannot be prevented [27-30]. However, perfectly paired RNA strands are uncommon over long sequences. Breaks in a regular A-form helix, caused by mismatched or unmatched bases in one or both strands, are common and result in bulges or internal loops [45]. RNA duplexes can also arise from non-complementary antiparallel strands by ribose zipper motifs; that is, consecutive hydrogen-bonding interactions between ribose 2'-hydroxyls groups of each strand [45].

Regarding other RNA complex structures, **triplexes** are found in many pseudoknots and almost all large structured RNAs [51]. Mainly, they consist of RNA double helices with a third strand coupled in their minor or major grooves through Hoogsteen base-pairing, without disrupting the preformed duplex structure. Naturally occurring triplexes, such as those formed in telomerase RNAs [52-57], long non-coding RNAs [58-60] and metabolite-sensing riboswitches [61-63], seem to be important in the stabilization of three-dimensional RNA structures and for diverse biological functions [64-66].

Concerning **RNA quadruplexes**, they follow the same folding pattern as their DNA counterparts, being found in guanine rich regions of telomeres and mRNAs [67-73]. A deeper insight into this special structure can be found in *Section 1.2* of this chapter.

### 1.1.4.2. Functions of RNA structures

A variety of RNA forms have been distinguished based on its diverse and often complex biological functions, including *messenger RNA*, *transfer RNA*, and *ribosomal RNA*.

*Messenger RNA (mRNA)* is essentially a copy of a section of DNA, commonly called transcript, which carries to ribosomes the necessary genetic information to

synthesise the amino acid sequence of a functional protein. *Transfer RNA (tRNA)*, folded into a cloverleaf-shaped structure, serves as molecular adaptor in protein synthesis, carrying amino acids to the ribosomes and arranging them according to the nucleotide sequence of the mRNA. *Ribosomal RNA (rRNA)*, together with a set of basic proteins, forms the ribosomes, the cellular organelles where protein synthesis, i.e. translation, takes place. [43]

There are still more RNA forms. For instance, short RNAs are not only part of organelles like ribosomes and spliceosomes, but also of some enzymes like the RNA polymerases. Other classes of RNA species include *microRNAs (miRNAs), small interfering RNAs (siRNAs),* and *piwi-interacting RNA (piRNAs),* all of which perform important functions in the cell but are not translated into proteins. The discovery of these RNAs has been one of the most exciting advances in recent years, and there is currently a lot of interest in the use of these molecules for therapies [74-80].

## 1.2. G-quadruplexes

In spite of the early breakthrough of G-quadruplexes in 1962 [81], it has not been until the last two decades when the interest in this four-stranded structures has increased. They had been considered as structural curiosities observed *in vitro* [82, 83] until quite recently, when researchers recognized them as potential mechanisms for regulating multiple biological processes *in vivo*, such as transcription, replication, translation and telomere maintenance [84-87].

In this section, the structure, topology and biological roles of G-quadruplexes will be reviewed as well as their budding nanotechnological applications.

### 1.2.1. Structure and topology

Guanine-rich regions of nucleic acids (*G-tracts*) have a high propensity to selfassociate into four-stranded structures known as *G-quadruplexes (G4s)*. Their building blocks are planar aromatic arrays of four Hoogsteen hydrogen-bonded guanines, called *G*-quartets or *G*-tetrads, in which each guanine base forms two hydrogen bonds with its neighbours (*Fig.1.10 (D*)). Then, a stable G4 arises from the stacking of at least two of these G-quartets, which are held together by loops created by the nucleotides not involved in the tetrads (*Fig.1.10 (C*)).



**FIGURE 1.10.** *G*-quadruplex structure. Front **(A)** and top **(B)** views of the crystal structure of the intramolecular G-quadruplex formed in the human telomeric DNA sequence AGGG(TTAGGG)<sub>3</sub> obtained from potassium-containing solution (PDB ID: 1KF1) [88]. It is 41 Å wide and 6.3 Å tall. The local G-quartet rise is 3.13 Å, with an average 30° twist between successive planes. Potassium ions, here in red, are positioned between the stacked G-quartets, 2.7 Å from each of the eight O6 carbonyl groups, in a bipyramidal antiprismatic arrangement. Nitrogenous bases are represented in light blue and the backbone of alternating phosphate and pentose residues in dark blue. **(C)** Schematic plot of the same structure shown in (A), where blue rectangles represent guanine bases and red spheres, potassium ions. **(D)** Chemical structure of the basic structural motif of a G-quadruplex conformation, i.e. the G-quartet. Hydrogen bonds and the negatively charged channel created by O6 atoms (in green) are shown as blue and green dashed lines, respectively. M<sup>+</sup> stands for metal ion.

The stability of this stacking is strongly dependent on monovalent cations (M<sup>+</sup>). It has been ascribed to the strong negative electrostatic potential created by the guanine O6 oxygen atoms, which results in a central negatively charged channel inside the G-tetrad stack where cations are located. The precise position of these counter-ions depends on their nature. For instance, potassium and ammonium ions tend to be situated equidistant from two consecutive tetrads so that they are eight coordinated to O6 atoms in a symmetric tetragonal bipyramidal configuration. The smaller sodium ion is able to being either in the plane of a G-quartet or accommodated midway between two successive ones [89-95].

In principle, G4s can be folded from a single G-rich strand intramolecularly, being required four or more G-tracts in the strand, or by the intermolecular association of two (bimolecular) or four (tetramolecular) separate strands (*Fig.1.11*). More in detail, intramolecular G4-forming sequences can be described as  $G_mX_nG_mX_oG_mX_pG_m$ , where *m* is the number of G residues in each short G-tract and  $X_m$ ,  $X_o$  and  $X_p$  can be any combination of residues, including G, of equal —as in vertebrate telomeric sequences— or unequal length, that form the loops. Meanwhile, bimolecular G4s reported to date respond to the association of two identical sequences  $X_nG_mX_oG_mX_p$ —where n and p may be zero— whereas tetramolecular configurations would be formed by four  $X_nG_mX_o$  or  $G_mX_nG_m$  strands associating together [94, 95].



**FIGURE 1.11.** *Types of G-quadruplexes according to the number of strands involved.* Gquadruplexes can be folded from **(A)** a single G-rich sequence intramolecularly or by the intermolecular association of **(B)** two or **(C)** four separate strands (highlighted in different colours). For simplicity, the twist between successive G-tetrad planes is not represented.

These non-canonical conformations exhibit extensive structural polymorphism relative to duplex DNA. The relative arrangement of strand polarity together with variations in loop size and sequence and/or in the number of stacked G-tetrads lead to a multitude of G4 structures, as indeed is found experimentally [88, 96-108].

Regarding the polarities of the four strands in a G4, they can be parallel, three parallel and one anti-parallel, or alternating parallel, resulting in different conformations named as *parallel*, *3+1* or *hybrid* and *antiparallel*, respectively (see *Fig.1.12*). These variations in strand polarity affect the location of the loops between G-rich segments in intramolecular and bimolecular G4s.



**FIGURE 1.12**. *G*-quadruplex folds depending on strand polarity: (A) Parallel; (B) 3+1 or hybrid and (C) antiparallel. Blue rectangles and red spheres represent guanine bases and metal ions, respectively. Strands highlighted in orange represent a change in polarity. For simplicity, the twist between successive G-tetrad planes as well as the anti- and synconformations of guanosine nucleosides are not taken into account in these schematic plots.

As shown in *Figure 1.13*, adjacent parallel strands require a connecting loop to link the bottom G-quartet with the top one, leading to *propeller* type loops. Anti-parallel G-strands can be linked either by *diagonal* or *lateral* (sometimes termed *edge-wise*) loops, depending on whether the strands are adjacent or diagonally opposed. In addition to its function as a linker, loops can themselves form stacking and hydrogen bonding interactions through their residues, further stabilizing (or not) G-quadruplex folds [90, 106].



**FIGURE 1.13.** *Types of loops in G-quadruplex structures.* **(A)** Propeller-type; **(B)** diagonal-type and **(C)** lateral-type. For simplicity, the twist between successive G-tetrad planes as well as the anti- and syn- conformations of guanosine nucleosides are not taken into account in these schematic plots.

In addition to the location of the loops, the orientation of individual strands in a G4 has also consequences for the glycosidic angle conformations *—anti* or *syn* (see *Fig.1.14*) *—*, of the guanosines involved in a G-tetrad. Thus, an all parallel orientation for all four strands, as in the tetramolecular G4 shown schematically in *Figure 1.12 (A)*, requires that all glycosidic angles adopt an *anti* conformation. Antiparallel G4s have both *syn* and *anti* conformations, arranged in a way that is characteristic of a given topology and a set of strand orientations [94, 109].



**FIGURE 1.14.** *Glycosidic angle conformations in a G-tetrad.* **(A)** Anti- and (B) synconformations of the guanosines involved in a G-tetrad.
### 1.2.2. Location of potential G-quadruplex-forming sequences (PG4s)

Both prokaryotic and eukaryotic genomes are rich in potential G4-forming sequences (PG4s), even some human DNA and RNA viruses [110-118]. Thanks to genome-wide analyses, the location of these sequences has been revealed recently. Specifically, it was found that, in addition to being non-randomly distributed, pG4s are evolutionary conserved, indicating a selection pressure to retain these sequences at specific functional genomic domains [119].

Focusing on human genome, it contains almost four hundred thousand of PG4s, being found the highest concentration at telomeres. Inside the cell nucleus, other opportunities to form G4s arise during DNA replication, transcription and repair, when DNA becomes transiently single stranded due to the breaking of Watson-Crick base pairings [84, 86]. RNA is also known to form these secondary structures, having been reported as many as 3000 mRNAs with PG4s located in their 5' untranslated regions (UTRs) [67-69, 71, 72]. Long G-rich RNA transcripts of telomeric DNA (TERRA) are also identified as genomic regions capable of folding into G4 structures [70, 120, 121].

#### 1.2.2.1. DNA and RNA G-quadruplexes in human telomeres

Telomeres are nucleoprotein structures located at the end of linear chromosomes that are essential to protect these regions from recombination and degradation activities [122]. In vertebrates, they consist of long tandem arrays of a short oligonucleotide sequence (~10-15 kb long) —TTAGGG in one strand (so-called *TG strand*) and AATCCC in the complementary [123] —, and an associated six-protein complex (TRF1, TRF2, Rap1, TIN2, TPP1 and POT1), termed *shelterin* [124, 125]. As shown in *Figure 1.15*, the TG strand of each telomere is longer than its complement, leaving a region of G-rich single stranded DNA (*G-overhang*) of up to a few hundred nucleotides (150-200 nt) at the 3' end. The 3'- OH group at the termini of this G-overhang is specifically recognised by telomerase, a reverse transcriptase in charge of telomere elongation in those cell types in which it is expressed [126, 127].

Telomere shortening and telomerase activity are directly related to cellular aging and cancer [128-133]. After each cell division, telomeres shorten owing to the so-called "*end-replication problem*"; that is, the incomplete replication of linear chromosomes by conventional DNA polymerases. Most somatic cells do not express telomerase activity so when a critical telomere length is reached, they undergo a growth arrest called *replicative senescence*, leading to chromosomal instability and subsequent loss of cell viability. In contrast, high levels of telomerase have been detected in approximately 90% of human malignant tumour cells, enabling their division indefinitely [128]. Consequently, telomerase has become both a highly attractive biomarker and a target for the development of new strategies for cancer detection, diagnosis and treatment [129, 133, 134].



**FIGURE 1.15.** *Telomeres at the end of linear chromosomes.* Here, it is shown their structure in vertebrates: long tandem arrays of double stranded TTAGGG repeats (~10-15 kb long) in which the TG strand is longer than its complement, leaving a region of G-rich single stranded DNA overhang of up to a few hundred nucleotides (150-200 nt) at the 3' end. Reproduced with permission from [135].

Telomere synthesis by telomerase requires that G-overhangs remain singlestranded. However, to protect themselves from nucleases and the enzymes that repair double-stranded breaks, these single-stranded regions can either fold back and pair with its complement in the duplex region of the telomere, giving rise to what is known as a *t- loop* [122, 136, 137], or form consecutive intramolecular G4s spontaneously [138-141]. These two capping structures need not be mutually exclusive and could represent different functional states [84].

In addition to DNA G4s, RNA analogous structures are also possible at the end of telomeres. Although telomeres had always been considered transcriptionally silent because of their compact heterochromatin structure, recent findings have demonstrated the opposite. In particular, the telomeric C-rich strand is transcribed from subtelomeric regions towards chromosome ends into **TElomeric Repeatcontaining RNA (TERRA)** largely by the DNA-dependent RNA polymerase II (RNAPII) [142-145]. Ranging in size from 100 to 9000 nucleotides, TERRA molecules contain subtelomere-derived sequences and several UUAGGG repeats at their 3'end, which are susceptible to fold into multiple G4 blocks. The existence of these large, heterogeneous and non-coding transcripts may reveal a new level of regulation and protection of chromosome ends [146-153] and seems to correlate with their potential to adopt a G4 conformation [154].

Unlike the wide structural polymorphism exhibited by their DNA counterparts —highly dependent on the involved sequences and existing environments [94, 155]—, TERRA of various lengths solely adopt a **parallel G4 structure** in dilute solutions [120, 121, 156-161] and even in living cells [70, 162]. TERRA G4 parallel topology is less sensitive to the surrounding environment, being found a similar structure both in Na<sup>+</sup> [156] and K<sup>+</sup> solution [161, 163]. Additionally, the TERRA G4 is more thermo-stable than the DNA counterpart [164, 165].

#### 1.2.3. Biological relevance

The recent evidence that G4 structures occurs *in vivo* suggests that these DNA and RNA higher order structures may play important cellular regulatory roles,

such as protecting telomere ends against nucleases or specifying the initiation site of DNA replication. However, G-rich sequences also pose particular problems for cells, being necessary spatio-temporal regulation mechanisms for the formation and resolution of G4s.

In telomeres, G4 structures provoke the inhibition of telomerase activity, leading to telomere shortening [166, 167]. Then, a G4 folding and unfolding regulation mechanism is needed to allow telomere synthesis by telomerase when required [168]. The phosphorylation of *telomere end-binding protein*  $\beta$  (TEBP $\beta$ ) in ciliates or the preferential function of several RecQ helicases at vertebrate telomeres, such as *Werner syndrome* (WRN) and *Bloom syndrome* (BLM) proteins, would be some examples of that [138, 169-173].

Similarly, during replication, G4s formed when DNA is transiently singlestranded need to be resolved by specialized DNA polymerases or unwinding helicases, like *Fanconi anemia complementation group J* (FANCJ) helicase, to allow replication [174, 175].

In transcription, the presence of a G4 could either impair its initiation by the RNA polymerase or inhibit the process if it is present in the antisense strand.

Finally, G4s formed in the 5'-UTR of mRNAs can regulate translation and lead to aborted RNA transcripts in hexanucleotide repeat expansion diseases, such as amyotrophic lateral sclerosis (ALS) or frontotemporal dementia (FTD) [71, 176].

#### 1.2.4. Applications

Since the demonstration of their presence in living cells [138, 140], G4 structures have raised a lot of interest in the field of nanomedicine. Thanks to their unique geometry, these four-stranded structures have become important drug-design targets for the treatment of various human disorders, ranging from viral infections to cancer [177-179].

The specific recognition of G4s by small ligands can occur through various binding modes in a manner analogous to that of double-helical DNA intercalators: *external stacking mode*, just on the surface of the terminal G-quartet (*Fig. 1.16 (A*)); *intercalating mode*, between the stacks of G-tetrads (*Fig. 1.16 (B*)) or *groove binding mode* (*Fig. 1.16 (C*)). Nonetheless, not all of them are considered to be equally probable. Intercalating binding of the ligand could be considered as the most improbable scenario due not only to the high stability and rigidity of the G4 structure, but also to the large amount of energy needed to distort their structural integrity. Thus, stacking of the drug on the outer planes of G-tetrads or groove binding seem to be more energetically favourable and probable [179]. Cationic porphyrin (TMPyP4) [180], acridine (BSU6039) [181], polycyclic acridine (RHPS4) [182] and several peptide-hemicyanine conjugates [183] are some examples of the compounds that have been identified as interacting specifically with DNA G4s following the latter patterns.



**FIGURE 1.16.** *Interaction modes between ligands and G-quadruplexes:* (A) external stacking mode, on the surface of the terminal G-quartet; (B) intercalating mode, between the stacks of G-tetrads, and (C) groove binding mode. Red spheres represent metal ions and dark blue disks, G4-binding drugs.

Over the last decade, studies about stabilization of G4 structures by small ligands have been mainly focused in the G-rich single-stranded regions of human telomeres. As described previously, each 3'-telomeric overhang is capable of adopting a G4 conformation, blocking the elongation step catalysed by telomerase and, consequently, leading to cellular senescence. Then, achieving stable G4 configurations in these regions by using small ligands has become a compelling antitumoral strategy [184-192]. But, without any doubt, the major challenge will be identifying non-toxic drugs that are truly selective for G4 assemblies [184].

Besides their potential application as antitumoral agents, G4s are also considered promising candidates in the field of nanotechnology. In particular, G-quartets have been proposed as building blocks of new electronic nanodevices due to their capability to form higher order structures, termed G-wires, as well as their rather high conductance [193-200]. Furthermore, some G-quartet-forming DNA aptamers have been reported as attractive stationary phase materials in bioanalytical chemistry [201-203], effective catalysts [204, 205] and, labelled with fluorescent dyes, as new types of biosensors [206-213].

# **1.3. Mechano-chemistry and single-molecule manipulation techniques**

The essential role that forces play in biochemical processes soon attracted the interest of many scientists, giving rise to a new, dynamic, exciting and interdisciplinary research field called *mechano-chemistry* [214]. It includes the study of biochemical processes as diverse as protein folding [215-222], the protein-induced bending of DNA [223, 224], the stress-induced enzymatic catalysis [225, 226], the behaviour of molecular motors [227-233] and the elasticity of nucleic acid structures [50, 234-238].

The cell environment where all these fundamental processes take place is stochastic. It means that proteins and enzymes exist in a background of random thermal fluctuations that affect their dynamics. The study of these fluctuations in real time and at a single-molecule level was not possible until quite recently. Traditional techniques, such as circular dichroism (CD), nuclear magnetic resonance (NMR) and fluoresce microscopy, can only investigate chemical processes on a bulk level, which only gives access to the average dynamics. New experimental tools should be developed and, over the last two decades, outstanding progresses have been made. In fact, *single-molecule manipulation techniques* (SMTs) have emerged as powerful tools to investigate the movements, forces and strains generated by single biological molecules in real time [214, 239-243]. Their most important feature is the absence of problems associated with populating averaging. Whereas bulk experiments hinders the access to fluctuations and the identification of multispecies distributions, single-molecule approaches allow the access to this information as there are not constrictions associated to the multiple interactions that take place in an ensemble of equivalent molecules. Nowadays, a wide variety of SMTs can be distinguished depending mainly on force and displacement ranges and spatial and temporal resolution. The most commonly used are optical tweezers (OT), magnetic tweezers (MT) and atomic force microscopy (AFM). A comparison of them can be found on *Table 1.1*.

	Optical tweezers	Magnetic tweezers	AFM
Force range (pN)	0.1-100	0.01-10	10-10 <sup>5</sup>
Displacement range (nm)	0.1-10 <sup>5</sup>	5-104	0.5-104
Spatial resolution (nm)	0.1-2	5-10	0.5-1
Temporal resolution (s)	10-4	10-1-10-2	10-3
Practical advantages	Specific manipulation High force resolution	Specificity to magnets Ability to induce torque	High spatial resolution

**TABLE 1.1.** Comparison of the most commonly used single-molecule manipulation *techniques.* Based on Ref. [214] and [243].

In this section the operating principle of the SMTs used in this thesis, optical tweezers and AFM, will be described in detail, being also highlighted their practical limitations and drawbacks.

## 1.3.1. Optical tweezers

Optical tweezers (OT) are extremely well suited for the measurement of force and motion in a large variety of biomolecular and cellular processes. It can be used to exert small forces on particles ranging in size from  $\sim 20$  nm to several micrometres and, simultaneously, measure the three-dimensional displacement of the trapped particle with sub-nanometre accuracy and sub-millisecond time resolution. Single cells, organelles within cells, lipid vesicles and polystyrene or silica microspheres —used alone or as probes linked to a biomolecule of interest—are examples of trapped objects by an optical field [243-247].

#### 1.3.1.1. Optical trap

The force sensor of this single-molecule technique is an optical trap. It is created by focusing a laser beam to a diffraction-limited spot with a high numerical aperture (NA) microscope objective [248]. Not all types of particles can be optically trapped; it is dependent on their refractive index. Whereas a dielectric particle is attracted to regions of higher field intensity, a metallic one may have either an attractive or a repulsive nature towards them [249]. Cells, organelles and biomolecules have been shown to behave as dielectric [250-253].



**FIGURE 1.17.** *Optical forces acting on a dielectric particle.* **(A)** Schematic representation of the scattering ( $F_{scat}$ ) and gradient ( $F_{grad}$ ) components of the optical force on a dielectric sphere originated by a Gaussian laser beam. **(B)** Axial stability of a dielectric sphere on a focused laser beam due to the gradient force along the centre line of the beam, towards the focal region. Reproduced with permission from [254].

Two kind of forces act on an optically trapped dielectric particle: the *scattering force*, produced by the photons striking the particle along their propagation direction, and the *gradient force*, produced by a gradient of field intensity. To form a stable optical trap, the gradient force along the optical axis must overcome the

scattering force, which is achieved if the focusing of the laser beam is strong enough (see *Fig.1.17*).

For small displacements (~ 150 nm) of the trapped object from its equilibrium position, the force is linearly proportional to the displacement, so the optical trap can be interpreted as a Hookean spring. The spring constant, or stiffness, depends on how tightly the laser is focused, the laser power and the polarizability of the trapped object and it needs to be well characterized to attain accurate force data [243, 255].

#### 1.3.1.2. Basic elements and up-to-date configurations

Although the range of OT experimental configurations is diverse and keeps growing, there are some basic elements that are present in all of them:

- Laser illumination. For high-precision optical trapping measurements, a laser beam with a Gaussian output intensity profile and superior power stability is required. Thus, the achievement of the smallest focal spot is possible, producing the largest and most stable optical gradient and reducing drastically measurement noise. Moreover, to minimize optically induced damage in biological samples, a trapping laser operating in the nearly-infrared region from 800 to 1100 nm is typically selected. In that optical window, water and biological specimens exhibit an adequate transparency. Dual-beam systems, which basically consist of two equal counter-propagating laser beams brought to the same focus, are also proposed as an alternative to reduce photodamage [256]. Regarding the output power, a couple of hundred of mW is generally appropriate to generate forces up to hundreds of pN. [243, 255]
- Strongly-focusing objective lens. To achieve the steep focus needed to create a stable optical trap, the NA of the trapping microscope objective should be near or above 1.0. This requirement is met by both water and oil

immersion objectives. However, and although the latter offers the highest NA, water immersion objectives are the most commonly used for optical trapping since they do not introduce spherical aberrations that degrade optical trap performance deep in solution. [243, 257, 258]

Microfluidics chamber. It is the place where the single-molecule experiments are carried out in buffer solution at a controlled rate. It is usually built by using two glass coverslips sandwiching two parafilm layers with an imprinted system of channels (see *Fig.1.18*). Single cells or coated beads with chemically attached biomolecules are introduced inside the fluidic chamber by the upper channel and flowed to the focal region, located in the central one, through a glass capillary. In some OT configurations, a micropipette is placed inside the chamber not only to act as the frame of reference of the optical trap, but also to hold a second coated bead by suction. This second microsphere, to which the biological specimen will also bind chemically, is flowed through the lower channel, also connected with the central one by a glass capillary.



**FIGURE 1.18.** *Schematic representation of a microfluidics chamber.* Microfluidics chambers are usually built by using two microscope coverglasses sandwiching parafilm layers with imprinted channels.

- Condenser lens. It is used to collect the light scattered by the optically trapped biological specimen and convert it into a parallel beam.
- > **Light-detection system**. The interference between the light scattered by

the trapped bead and unscattered light is measured with a quadrant photodiode or position-sensitive photodetector (PSD) to further obtain the force and displacement measurements that experience the sample of interest. When a trapped bead is at its equilibrium position, a symmetric interference pattern and a null detector signal is produced. Any displacement of the bead results, then, in an asymmetric interference profile, which generates a detector signal proportional to the bead displacement. Therefore, registering the light deflected by the bead and characterizing the optical trap as a spring constant, both bead movement and force can be determined. [243]

A schematic representation of the OT set-up previously described and used in this thesis (see *Chapter 3*) is shown in *Figure 1.19*.



**FIGURE 1.19.** *Basic elements in a OT instrument.* Laser light (typically infrared) is highly focused by a microscope objective creating an optical trap inside a microfluidic chamber, where the single-molecule experiments are carried out. A bead or a single cell can be trapped in the focal region whereas a second bead is held by suction on top of a micropipette. The light scattered by the sample in the trap is firstly collected by a condenser lens and, then, redirected to a position-sensitive photodetector.

Nowadays, more sophisticated OT instruments are pursued by several laboratories around the world, looking for better and high-resolution optical trapping measurements. The so-called miniTweezers set-up [259], also used in this thesis (*see Chapter 3*), is an example of that, which has been subsequently improved by adding a temperature controller [260]. Within this new generation of OT systems are the hybrid configurations; i.e., the combination of OT with other SMTs to take advantage of the best features of each technique. Magneto-optical tweezers [261] —which combines the nanometer resolution of OT and the easy manipulation of MT— and OT-FIONA (*Fluorescence Imaging with One Nanometer Accuracy*) instrument [262] —which allows the detection of biomolecule position, conformation or biochemical state simultaneously with its mechanical manipulation— are particularly fruitful examples of these hybrid tools.

## 1.3.1.3. Single-molecule manipulation of nucleic acids

Thanks to the advent of the OT technique, elastic properties and folding dynamics of diverse nucleic acid structures have been investigated from their characteristic force-extension curves (*F-x curves*): from double-stranded DNA and RNA polymers at different environmental conditions [28, 29, 50, 263, 264] to non-canonical conformations, like G4s [265-269].



**FIGURE 1.20.** Scheme of the experimental configuration in OT measurements. A single dsDNA molecule was tethered by opposite DNA duplex ends between two polystyrene microspheres: an anti-digoxigenin ( $\alpha$ -DIG)-coated bead, optically trapped, and a streptavidin-coated bead, held by suction on top of a micropipette. "B" stands for biotin and "D" for digoxigenin.

These single-molecule experiments are usually carried out in an experimental OT set-up like that shown in *Figure 1.19*. A single molecule is grabbed by two different, specific sites by means of two short nucleic acid stretches, usually called *handles*, to properly functionalized dielectric microspheres. Typically, nucleic acid fibres are prepared containing modified nucleotides with biotin groups at one end and digoxigenin at the opposite, so that they can bind to stretptavidin- and anti-digoxigenin-coated beads, respectively (see *Fig.1.20*). The purpose of the handles is to avoid direct contact between the molecule of interest and the surfaces of the beads, which may produce artefacts owing to surface effects [270]. F-x curves are obtained by the stretching, and subsequent relaxation, of a single biomolecule by moving the micropipette relative to the laser trap at a constant rate. The extension of the nucleic acid strain is determined from the distance between the beads with a video camera, whereas the force acting on it is calculated from the displacement of the laser beams on the PSDs [264].



**FIGURE 1.21.** *Characteristic force-extension curve of a single dsDNA molecule.* The graph shows an individual dsDNA molecule from the [30286, 34286] fragment of  $\lambda$  DNA (47.3% GC, 4.0 kbp) in 150 mM NaCl, 150 mM NaCl, 10 mM Tris-Cl, 1 mM EDTA, pH 7.5. Insets illustrate the strand arrangement in each elasticity regime according to the literature (see text for details). Extension axis has been normalized to the contour length of the molecule (L<sub>0</sub>).

In *Figure 1.21* a typical F-x curve of a single double-stranded DNA molecule is shown. Three different elasticity regimes can be distinguished during its stretching pathway [30, 264]:

- Entropic elasticity regime. Below some 5 pN, the molecule, initially bent over itself in a random way as a result of thermal fluctuations, is stretched like a Hookean spring until its characteristic contour-length; i.e. its length at maximum physically possible extension.
- Enthalpic or intrinsic elasticity regime. From 5 to 60 pN, approximately, the molecule is stretched beyond its contour length (L<sub>0</sub>).
- Overstretching transition. At about 65 pN —in non-torsionally constraint molecules and depending on ionic strength, pH and sequence—, the molecule undergoes a structural transition to an almost totally unwound state with an elongation of 1.7 times its contour length. This sudden change happens over a very narrow force range (~2 pN) and involves an increase in the rise per phosphate from 3.4 to 5.8 Å. The bases unstack and, depending on environmental conditions, the strands may either remain linked by partial base pairing —like a parallel ladder (*S-form*)—, or separate into a totally molten state (*force-induced melting*) or in a heterogeneous combination of both [271-281]. The stretching cycle is always reversed before 100 pN because the biotin-streptavidin bond often breaks after a few seconds at this force.

When the stress on the molecule decreases, it begins to return to its natural state but following a slightly different pathway, provoking a certain hysteresis between the stretching and the relaxation traces. This absence of overlapping in the F-x curve is a hallmark of the reannealing of the strands out of equilibrium. The DNA regions which have remained base-paired rapidly contract to the B-form but the frayed tracts take more time to reanneal, perhaps due to the secondary structure they could adopt in the overstretched molecule [264].

Regarding non-canonical nucleic acid structures, as G4s, F-x curves provide a great deal of information, including their mechanical stability ( $F_{rupture}$ ) and unfolding extension ( $\Delta x$ ). Both parameters are extracted from the inherent rupture events, which become observable as jumps over the characteristic F-x curves of the handles –dsDNA, dsRNA or hybrids– (see *Chapter 3*).

#### 1.3.1.4. Limitations and drawbacks

The awesome versatility and outstanding precision of this single-molecule technique are accompanied unfortunately by important limitations and drawbacks.

Some of these important difficulties are associated with the fact of using light to generate force. As trap stiffness depends on the gradient of the optical field, any optical perturbation in the intensity distribution affects the performance of the OT. As a result, high-resolution measurements are only achieved if trapped samples are optically homogenous.

OT also lack selectivity and exclusivity. Any dielectric particle near the focus region can be optically captured, being possible the simultaneous trapping of a quite large number of objects. For this reason, freely diffusing samples must be kept at extremely low concentrations to prevent this undesirable effect [243].

Local heating in the vicinity of the optical trap due to the high light intensity at the focus of the trapping laser  $(10^9-10^{12} \text{ W}\cdot\text{cm}^{-1})$  is another drawback but it can be quantified, both theoretically and experimentally [282-284], and even kept limited [260].

Optical damage induced in trapped biological specimens is less well understood. Although the use of trapping lasers operating in the near-infrared spectrum minimizes photodamage, singlet oxygen or other reactive oxygen species could persist throughout this wavelength range [244].

Last but not least, the artificial nature of the single-molecule linkage strategy has to be considered. Whereas individual cells can be held directly in an optical trap, handling single molecules requires the use of molecular handles as explained in the previous section. The absence of this type of configuration in living conditions is considered as a handicap for *in vivo* applications of this technique. [243, 285]

#### 1.3.2. Atomic force microscopy

Atomic force microscopy (AFM) is a powerful and widely used imaging technique for visualizing the surface of samples with sub-nanometre resolution. Initially developed to overcome limitations of scanning tunnelling microscopy (STM) in imaging non-conductive samples [286, 287], AFM soon became an ideal tool for biological applications [288]. The possibility of operating under near-physiological conditions and manipulating individual molecules represent the major steps forward [289, 290].

#### 1.3.2.1. Operating principle

Basically, an atomic force microscope consists of a soft cantilever of 100 to 200  $\mu$ m in length with a sharp tip at its end —a pair of microns long and less than 100 Å in diameter—, which is used either to scan topography or to extend molecules attached to a surface [291].

When AFM is used in an **imaging mode** (*Fig. 1.22*), the cantilever scans the surface of the sample line after line. When the tip is brought into proximity of the surface, forces between the tip and the sample lead to a deflection of the cantilever according to Hooke's Law:

$$F = -kz \tag{[1.1]}$$

where F is the force exerted on the cantilever, k the spring constant determined by the type of cantilever, and z the displacement experienced by the cantilever along a scanning direction perpendicular to the surface of the sample. Typically, this deflection is measured using a laser spot reflected from the top surface of the cantilever into an array of photodiodes. This position-sensitive photodetector (PSD) allows the separation of the signals originated by the vertical and torsional movements of the cantilever during the scanning of a surface. The information of the topography of the sample resides in this vertical displacement whereas the frictional forces between the tip and the sample are provided by torsional data [292].



**FIGURE 1.22.** *Basic AFM configuration operating in imaging mode.* The force originated by the interaction between the atoms of the tip and those of the surface is a signal strongly dependent on the distance, which is recorded by a photodetector through the deflection of the cantilever during scanning.

When AFM is used for manipulating single molecules (*Fig.1.23*), this specialized version is called a *molecular force probe* (MFP) and its use is fundamentally distinct from imaging AFM. Specifically, the cantilever tip is

moved only vertically with respect to the surface by piezoelectric actuators. After the initial contact with the surface, the cantilever is pressed into the surface on which the sample is deposited with a predetermined constant force. The tip is then retracted and, because of the attachment of the sample (DNA, RNA or proteins) to it, the cantilever bends toward the surface. The displacement of the cantilever is monitored directly with a PSD. It is important to note that both the cantilever and the surface have to be chemically modified to achieving these specific attachments to the ends of the biomolecule of interest.



**FIGURE 1.23.** *Basic AFM configuration operating as a molecular force probe.* A typical AFM pulling experiment is represented in which a polyprotein molecule (displayed in yellow) is attached to a sample surface (displayed in orange) and an AFM cantilever tip. The piezoelectric scanner is retracted along the axial direction, increasing the separation between the cantilever and the sample surface. The force on the molecule is provided by the cantilever deflection, and the extension of the molecule is equal to the separation between the AFM tip and the sample surface.

As a result of MFP measurements, two types of characteristic data can be obtained: force and extension. The force exerted on the molecule during its pulling can be calculated by Hooke's law, using as *z* the cantilever deflection, and its extension is equal to the separation between the MFP tip and the sample surface. The precision of force and extension data depends on the quality of the calibration of the cantilever and the piezoelectric actuator, respectively. Piezoelectric scanners used in MFPs provide angstrom-level resolution and, together with a closed-loop position feedback control, are accurate and not susceptible to drift and hysteresis.

## 1.3.2.2. Limitations and drawbacks

The main drawbacks and limitations of AFM stem from the large size (100-200  $\mu$ m in length) and relatively high stiffness (10- 10<sup>5</sup> pN/nm) of the cantilevers. Consequently, a large minimal force is imposed, hindering AFM applications in the study of many biological processes and structures in which the forces involved are so minute.

In addition, unlike OT [253, 285], AFM is inappropriate for experiments *in vivo* because the AFM tip cannot penetrate the cell membrane without breaking it.

## **Chapter 2**

# **TERRA G-quadruplexes imaged by atomic force microscopy**

## 2.1. Introduction

Numerous studies have proven the utility of atomic force microscopy (AFM) for imaging and analysing nucleic acid structures at the single-molecule level, even under near-physiological conditions [50, 194, 293-307]. The data collected can be displayed as a 3-D topography of the sample, allowing sub-Angstrom accuracy in the determination of the z dimension of the sample of interest, i.e. height. Here, AFM has been used, for the first time, to make direct measurements of human telomeric repeat-containing RNA (TERRA) molecules in ambient air.

TERRA is a large non-coding RNA in animals and fungi, which forms an integral component of telomeric heterocromatin [142, 143, 148]. It could be involved in various cellular regulatory functions, such as regulation of telomere length, inhibition of telomerase and telomeric heterochromatin formation [149]. As its DNA counterpart, it is able to fold into multiple G-quadruplex (G4) blocks, each formed by a four repeat segment [121, 308-310]. Until now, several AFM studies have been done on the possible arrangements of G4s along human telomeric DNA [302, 304]. However, direct visualization of TERRA G4s formed in long RNA molecules by AFM had not been reported yet. We have designed different RNA substrates with telomeric G4-forming sequences, proposing an efficient method to check the formation of these special structures before further analysis.

## 2.2. Materials and methods

## 2.2.1. Synthesis of TERRA molecules

Three different RNA substrates with telomeric G4-forming sequences were visualized by AFM. In particular, two of them *—TERRA-25 and TERRA-25 with double stranded RNA handles—* contain twenty five GGGUUA repeats, giving rise to a maximum of six G4s, whereas the other one *—TERRA-16—* only has sixteen, being able to form four G4s at the most. The synthesis protocol of each of them is detailed below.

## 2.2.1.1. TERRA-25 molecules

The DNA template for this TERRA substrate was obtained by PCR (*Polymerase Chain Reaction*) amplification of a conveniently modified pUC18 plasmid (with an insert of twenty five GGGTTA repetitions) using the following pair of primers:

**F1**: 5'-Sp6 promoter-GAATAAGGGCGACACGGAAATGTTG-3' **R1**: 5'-TTTACGGTTCCTGGCCTTTTGCTG-3'

where F stands for forward primer and R for the reverse one. This technique, developed by Kary Mullis [311], is widely used in molecular biology to amplify a specific DNA segment in an easy, cheap and reliable way, generating thousands to millions of copies in a few hours and without using living cells. The reaction mixture (50  $\mu$ l) and PCR cycling conditions used here are detailed in *Table 2.1* and *Table 2.2*, respectively.

Reagent	V(µl)
Ultrapure Milli-Q-filtered water	22.0
Modified pUC18 plasmid (10 ng/µl)	1.0
Primer F1 (Integrated DNA Technologies, 10 μM)	1.0
Primer R1 (Integrated DNA Technologies, 10 $\mu$ M)	1.0
SapphireAmp® Fast PCR Master Mix (Takara ClonTech, Ref. RR350A)	25.0

TABLE 2.1. PCR mixture for DNA telomeric G4-forming sequen	ce amplification.
--	-------------------

Cycle		Temperature (°C)	Time (min)
Hot Start Automatic		94.0	3
Start cycle (x 40)	Denaturation	95.0	0.5
	Annealing	55.0	0.5
	Elongation	72.0	1.5
Temperature		72.0	5
Store		4.0	8

 TABLE 2.2. PCR cycles for DNA telomeric G4-forming sequence amplification.

The obtained reaction product was purified using *Wizard*® *SV Gel and PCR Clean-Up System* (Promega, Ref. A9281), eluting the sample in 50 µl of nuclease-free water. Its size was checked by 1% agarose gel electrophoresis. Subsequently, it was in vitro transcribed by Sp6 RNA polymerase (New England Biolabs, NEB, Ref. M0207S; 3h, 40°C), whose promoter sequence was incorporated at the 5' end of the forward PCR primer (F1). The generated transcript was purified using *RNeasy MinElute Cleanup Kit* (Qiagen, Ref. 74204), eluting the sample in 20 µl of RNase-free water. Its size was checked by 6% denaturing polyacrylamide/urea gel electrophoresis. Then, the sample was equilibrated in 15 mM KPi pH=7 buffer by diafiltration using *Amicon*® *Ultra centrifugal filter units* (Merck Millipore, Ref. UFC901024). Finally, to favour the formation of the G-quadruplexes, the sample is subjected to heating at 90°C for 5 min, after which it is gradually cooled at room temperature. A schematic representation of the whole process is shown in *Figure 2.1*.

Reagent		V(µl)
RNase-free water		14.80
Sp6 polymerase buffer 10X (NEB, Ref. M0207S)		2.50
rNTPs dil 1:10	А	
	С	1.25
	G	
	U	
DNA template (~50-60 ng/µl)		1.20
RNase Inhibitor (NEB, Ref. M0314S)		0.50
Sp6 polymerase (NEB, Ref. M0207S)		1.0

TABLE 2.3. Mixture reaction for in vitro transcription of DNA telomeric G4-forming sequences.



**FIGURE 2.1.** *Schematic representation of the steps involved in the synthesis of TERRA-25 molecules.* DNA template is generated by PCR amplification of a conveniently modified pUC18 plasmid and subsequently transcribed by Sp6 RNA polymerase, generating singlestranded (ss) RNA molecules with the desired number of telomeric repetitions.

## 2.2.1.2. TERRA-25 molecules with double-stranded RNA handles

To control that the structures visualized by AFM actually corresponded to G4s, TERRA-25 molecules were modified so that the telomeric sequence was flanked by asymmetric double-stranded RNA (dsRNA) at both sides. To synthesise the DNA templates for these handles, two more PCR amplifications were needed using the following pairs of primers:

F2: 5'-GAATAAGGGCGACACGGAAATGTTG-3'
R2: 5'-Sp6 promoter-GAGTCGTATTAAGCTTGGCACTGG-3'
F3: 5'- CTTCGAATTCGTAATCATGGTCATAGCTGT-3'
R3: 5'- Sp6 promoter-GGGTTCCTGGCCTTTTGCTG-3'

where F stands for forward primer and R for the reverse one, and the numbers 2 and 3 refer to handle A and handle B, respectively. Reaction mixtures and PCR cycling conditions are the same as previously described in *Table 2.1, 2.2* and *2.3*.

The final assembly is achieved by mixing equimolar amounts of the three transcripts (main molecule, handle A and handle B) and annealing them in the presence of 100 mM potassium chloride, which favours the G4 stabilization, in a PCR apparatus following the procedure specified in *Table 2.4*. A schematic representation of the whole process is shown in *Figure 2.2*.

Cycle	Temperature (°C)	Time (min)
Denaturation	90.0	5
Annealing	85.0	10
	62.0	90
	52.0	90
	20.0	8

 TABLE 2.4. Annealing procedure for TERRA assemblies.

#### 2.2.1.3. TERRA-16 molecules

To study the deposition geometry of the G-tandems with respect to the mica surface, TERRA molecules with fewer telomeric repeats were synthesised. Specifically, they contain sixteen GGGUUA repetitions; that is, molecules that can give rise to a maximum of four G4s. In addition, to favour their identification during AFM imaging, the 5' end of these single-stranded molecules was hybridized with a complementary RNA fibre to form a double helix following the same protocol described in the previous section. Then, two DNA templates were obtained by PCR amplification of a conveniently modified pUC18 plasmid (with an insert of sixteen GGGTTA repetitions) using the former pairs of primers F1-R1, for the main single-stranded molecule, and F2-R2, for its 5'-end-complementary. Reaction mixtures and conditions for PCR, in vitro transcription and annealing procedures are the same as previously detailed in *subsection 2.2.1.2*. A schematic representation of the whole process is shown in *Figure 2.3*.



**FIGURE 2.2.** Schematic representation of the steps involved in the synthesis of TERRA-25 molecules with dsRNA handles. (A) Synthesis of TERRA sequence; (B) synthesis of ssRNA to construct handle A; (C) synthesis of ssRNA to create handle B and (D) annealing of the three transcripts.



**FIGURE 2.3.** Schematic representation of the steps involved in the synthesis of TERRA-16 molecules with a single dsRNA handle. (A) Synthesis of TERRA sequence; (B) synthesis of ssRNA to construct the handle and (C) annealing of the two transcripts.

## 2.2.2. AFM imaging

TERRA imaging was performed using an atomic force microscopy from Nanotec Electronica S.L., operating in tapping mode in air, and using NSG-30 cantilevers (NT-MDT; nominal spring constant, k = 40 N/m; resonance frequency,  $v_R$  = 320 kHz). Under this working mode, the tip is oscillated near its resonance frequency and either the reduction of the oscillation amplitude or the shift in the resonance frequency is kept constant as an image is acquired. This allows imaging molecules weakly attached to a surface in ambient air [301]. Raw topographical images were processed using WSxM freeware [312]. Height measurements were done with respect to the basal salt layer.

#### 2.2.3. Sample preparation for AFM measurements

Before TERRA deposition, the top layer of a mica substrate was cleaved using Scotch tape to reveal an atomically flat and negatively charged surface. Immediately after, it was functionalized with 10 mM NiCl<sub>2</sub> during 1 min to allow the subsequent adsorption of the RNA molecules of interest. Excess of NiCl<sub>2</sub> was rinsed out with 3 ml of ultrapure Milli-Q-filtered water. Functionalized mica substrate was then dried under a gentle stream of nitrogen gas and checked by AFM to ensure a uniform surface. Finally, TERRA molecules were diluted to a concentration of ~ 0.5 pmol/µl in 10 mM Tris-Cl (pH=7.8) supplemented by 100 mM KCl, deposited onto functionalized mica substrate and incubated at room temperature for 3 min. RNA in excess was removed from the surface using 3 ml of ultrapure water. The surface was then blown dry in a gentle stream of nitrogen gas, being ready for AFM imaging.

## 2.3. Results

### 2.3.1. AFM imaging of TERRA-25

First experiments were carried out to observe specific G4 structures in TERRA molecules with twenty-five GGGUUA repeats; i.e. molecules that can give rise to a maximum of six G4s. As shown in *Fig.2.4B*, G4s appear as mountain-shaped protrusions over the mica, with a mean height ( $\pm$  SE) of 1.15  $\pm$  0.03 nm (*N*=76, *Fig.2.4D*). Single-stranded RNA overhangs were not visualized, probably because of either RNA degradation or a process of folding around the corresponding G4 structure, forming something like a ball of yarn.



**FIGURE 2.4.** *AFM imaging of TERRA-25 molecules.* **(A)** Dynamic-mode topographical AFM image of TERRA-25 molecules on Ni<sup>2+</sup> pre-treated mica taken in ambient air. **(B)** Several 3-D topographical AFM images where G-quadruplexes appear as mountain-shaped protrusions over the mica. **(C)** Height profile of a single TERRA molecule. **(D)** Height histogram of the analysed structure (*N*=76), fitting to a Gaussian distribution (mean ± s.d.,  $1.1 \pm 0.2$  nm).

#### 2.3.2. AFM imaging of TERRA-25 with double-stranded RNA handles

To ensure that the previously visualized structures actually correspond to G4s and not to salt agglomerations, molecules were modified in order to have asymmetric dsRNA handles at both ends of the G4-forming sequence (see *section 2.2.1.2*). As shown in *Figure 2.5 (A)* and (*B)*, topographical AFM images confirm this assembly; G4s appear again as mountain-shaped protrusions over the mica, but this time flanked by fibres of different lengths. Means in height (± SE) of G4-type structures and dsRNA handles are 0.91 ± 0.03 nm (*N*=23, *Fig.2.5 (C*)) and 0.50 ± 0.02 nm (*N*=23, *Fig.2.5 (D*)), respectively.



**FIGURE 2.5.** *AFM imaging of TERRA-25 molecules with dsRNA handles.* **(A, B)** Examples of dynamic-mode topographical AFM images taken in air, where height profiles of G-quadruplex structures (light blue) and RNA handles, A (648 bp; green) and B (397 bp; orange), are also shown. **(C)** Height histogram of G-quadruplex structures, fitting to a Gaussian distribution (mean  $\pm$  s.d., 0.9  $\pm$  0.1 nm). **(D)** Height histograms of dsRNA handles. In green, data corresponding to handle A with Gaussian fitting (mean  $\pm$  s.d., 0.5  $\pm$  0.1 nm) and, in orange, to handle B with log-normal fitting (median  $\pm$  s.d., 0.5  $\pm$  0.1 nm).

With regards to the lengths of the handles, they were measured using WSxM processor, which allows calculating the total length of a fibre by drawing a series of very short lines along its contour. Then, the mean lengths (± SE) obtained for

handle A (648 bp) and handle B (397 bp) were 191.9  $\pm$  1.6 nm (*N*=23) and 111.4  $\pm$  1.5 nm (*N*=23), respectively. This results in a nanometre-to-base pair conversion factor of 0.296  $\pm$  0.002 nm/bp for handle A, and of 0.281  $\pm$  0.004 nm/bp for handle B, values consistent with A-family conformations adopted by dsRNA molecules [30, 313].

#### 2.3.3. AFM imaging of TERRA-16 molecules

For the purpose of unraveling the geometry that the G-tandems adopt when they are deposited over a surface, TERRA molecules with sixteen telomeric repeats were synthesised; i.e. molecules that can give rise to a maximum of four G4s. In addition, to favour their identification during AFM imaging, the 5' end of these single-stranded molecules was converted into a double helix (see *Fig.2.6 (A)* and *(B)*). Means in height ( $\pm$  SE) of the G4-type structures and the dsRNA handles are 1.20  $\pm$  0.04 nm (*N*=28, *Fig.2.6 (C)*) and 0.75  $\pm$  0.03 nm (*N*=28, *Fig.2.6 (D)*), respectively. The mean length of the handles ( $\pm$  SE) is 188.2  $\pm$  0.9 nm, given a nanometre-to-base pair conversion factor of 0.290  $\pm$  0.001 nm/bp, again consistent with A-family conformations adopted by dsRNA molecules [30, 313].



**FIGURE 2.6.** *AFM imaging of TERRA-16 molecules.* **(A, B)** Examples of dynamic-mode topographical AFM images taken in air, where height profiles of G-quadruplex structures (light blue) and dsRNA handles (green) are also shown. **(C)** Height histogram of G-quadruplex structures, fitting to a Gaussian distribution (mean  $\pm$  s.d., 1.2  $\pm$  0.2 nm). **(D)** Height histogram of dsRNA handles, fitting to a Gaussian distribution (mean  $\pm$  s.d., 0.8  $\pm$  0.2 nm).

## 2.4. Discussion

A clear feature in all the obtained AFM images is the appearance of mountainshaped protrusions. Their identification with tandems of G4s arises from the analysis of the same molecules modified with handles (see *sections 2.3.1* and *2.3.2*), obtaining similar heights on average ( $1.15 \pm 0.03$  nm vs.  $0.91 \pm 0.03$  nm). In contrast, by reducing the number of telomeric repeats, no appreciable differences in height were found ( $1.20 \pm 0.04$  nm; see *section 2.3.3*), suggesting that consecutive G4s blocks (usually called *G-wires*) could be adsorbed with G-tetrad planes orthogonal to the surface (see *Fig. 2.7*).



**FIGURE 2.7.** *Elastic deformation of biomolecules during AFM imaging induced by tipsample interaction.* The height of G-wires is not as significantly diminished as in duplex structures, due to the stabilizing capacity of the coordinated cations.

The reduced heights of both G4s structures and dsRNA handles compared with their theoretical dimensions [313, 314] are probably caused by tip-sample interaction during scanning, provoking a strain that compresses the biomolecules [315]. However, the height of handles is significantly diminished compared to the height of the G-wires. A possible explanation could be in the coordinated cations of the quadruple helical structures. They provide an additional stability and, consequently, greater resistance to conformational changes upon surface binding [194].

## **Chapter 3**

# The unfolding dynamics of short TERRA molecules

### **3.1. Introduction**

The unfolding of a biomolecule has been majorly studied by chemical denaturation or temperature ramps. However, these methods, which are applied to an ensemble of similar molecules, do not represent the dynamical scenarios in which localized forces are applied by specialised proteins, like telomerase, to a single molecular structure in the cell. In addition, bulk denaturation methods usually mask the heterogeneity of statistically less populated intermediates by the ensemble averaging process, progressing through pathways with different thermally induced rates relative to those of the mechanical unfolding [316, 317]. Then, single-molecule manipulation techniques have emerged as powerful tools to investigate the movements, forces and strains generated by individual biological species in real time as well as to measure directly multistate or multispecies distributions [214, 240, 242].

G-quadruplex (G4) unfolding is a halfway problem in complexity between protein and duplex nucleic acid unfolding. It has been proposed that the configurational pathways between the folded and unfolded state comprise a few major intermediate conformations, such as G-hairpins and G-triplexes [318-320]. This is in contrast with the inherent complexity of the protein folding intermediates and the simplicity of a hairpin unzipping. Therefore, G4 unfolding also represents an intermediate molecular model system from the point of view of non-equilibrium thermodynamics [321].

Single-molecule mechanical folding and unfolding processes of long TElomeric Repeat-containing RNA (TERRA) molecules —susceptible to fold into multiple G4 blocks; see *section 1.2* for a deeper insight— have been studied previously, finding weak cooperativity in the tandem G4 unfolding [267]. In subsequent studies on the human telomeric DNA overhang, a similar behaviour was found [269]. Yangyuoru et al. analysed, by using a similar single-molecule procedure, a TERRA sequence with only four GGGUUA repeats, which can only fold into a single G4 [268]. These and other authors found that RNA G4s are more stable than previously analysed DNA G4s [270, 322-325].

The next step towards the global understanding of G4 stability, its dynamic processing by motor proteins and its interaction with drugs is unveiling the influence of neighbouring tracts in the folding pathway of a single TERRA G4. This chapter will be focused on the study of the unfolding dynamics of short TERRA molecules by using force spectroscopic optical tweezers (OT). In contrast to earlier bulk studies that showed dimorphic competition between a hairpin and a G4 in either a purposely-prepared RNA sequence [326] or a DNA tract [327, 328], here a single-strand wide conformational competition within a TERRA tract is demonstrated at the single-molecule level.

## 3.2. Materials and Methods

### 3.2.1. Synthesis of TERRA molecules

Four different RNA substrates with human telomeric G4-forming sequences were synthesised for their single-molecule manipulation by OT. In particular, three of them consist of a central ssRNA fragment of *n* repeats (n= 5, 6 or 7) of the telomeric sequence GGGUUA flanked by non-G-rich, random sequences at both sides, UAUA at the 5' end and UAGU at the 3' end (see *Fig.3.1 (A)*). The last one is made up of a central fragment of five GGGUUA repeats flanked by four extra
ribonucleotides (UAUA) at the 5' end and by ten extra nucleotides (UAGUCUUCGA) at the 3' end (see *Fig. 3.1 (B)*). In all cases, the telomeric RNA sequence together with the flanking extra ribonucleotides are located between two hybrid DNA:RNA handles, which are conveniently labelled for their later attachment to beads through antigen-antibody linkages in the OT experiments.



**FIGURE 3.1.** *Scheme of RNA molecular constructions.* **(A)** RNA substrates with a central single-stranded fragment of *n* repeats (n = 5, 6 or 7) of the human telomeric G4-forming sequence GGGUUA flanked by non-G-rich, random sequences at both sides, UAUA at the 5' end and UAGU at the 3' end. In all cases, the complete single-stranded sequence is linked to two hybrid DNA:RNA handles, one of 650 bp at the 5' end (handle A) and the other of 401 bp at the 3' end (handle B), conveniently labelled with digoxigenin (represented as D) and biotin (represented as B), respectively, for their later attachment to beads in the OT experiments. **(B)** RNA substrate with a central single-stranded fragment of five GGGUUA repeats flanked by four extra ribonucleotides (UAUA) at the 5' end and by ten extra nucleotides (UAGUCUUCGA) at the 3' end. The complete single-stranded sequence is also linked to two hybrid DNA:RNA handles, but this time handle B is shorter (395 bp).

By choosing the number of telomeric repeats between five and seven, the formation of TERRA G4 tandems is avoided, since four GGGUUA repeats are needed to give rise to a single conformation of this type. Thereby, the study of the possible influence of neighbouring tracts in the folding pathway of a single TERRA G4 is allowed. The five-repeat RNA molecule with four ribonucleotides at both sides will be considered as the reference system *—TERRA reference construct.* The six and seven-repeat RNA species *—RNA control systems 1* and *2*, respectively— have been designed to analyse the effect of additional telomeric repetitions in the formation of this secondary structure, whereas the five-repeat RNA molecules with a longer random sequence overhang *—RNA control system 3*— will be used to study the influence of adjacent heterogenous ssRNA over a telomeric G4-forming sequence. The synthesis protocol of each of them is detailed below.

# 3.2.1.1. TERRA reference construct and RNA control systems 1 and 2

The synthesis protocol of these three RNA constructions involves the following steps:

## 1. Production of modified pUC18 plasmids.

pUC18 plasmids with inserts of five, six and seven GGGTTA repeats were obtained by extraction from E.coli DH5 $\alpha$  transformant cells [267] using the *Wizard*® *Plus SV Minipreps DNA Purification System* (Promega, Ref. A1330), being subsequently sequenced by STAB VIDA to crosscheck the presence of these telomeric repetitions.

### 2. PCR amplifications of telomeric G4-forming sequences.

The DNA templates for these TERRA substrates (n=5, 1112 bp; n=6, 1118 bp; n=7, 1124 bp) were obtained by PCR amplification of conveniently modified pUC18 plasmids —with an insert of five, six or seven GGGTTA repetitions as required— using the following pair of primers:

F1: 5'-Sp6 promoter-GAATAAGGGCGACACGGAAATGTTG-3' R1: 5'-TTTACGGTTCCTGGCCTTTTGCTG-3'

where F stands for forward primer and R for the reverse one, and the number 1 refers to TERRA. The reaction mixture (50  $\mu$ l) and PCR cycling conditions used here are detailed in *Table 3.1* and *Table 3.2*, respectively. The obtained reaction products were purified using the *Wizard*® *SV Gel and PCR Clean-Up System* (Promega, Ref. A9281), eluting each sample in 50  $\mu$ l of nuclease-free water. Their sizes were checked by 1% agarose gel electrophoresis.

## 3. <u>PCR amplifications of the DNA templates for the hybrid handles</u>.

The DNA template for the hybrid handles, handle A (650 bp) and handle B (401 bp), were also obtained by PCR amplifications of the conveniently modified pUC18 plasmids using the following two pairs of primers:

F2: 5'-GAATAAGGGCGACACGGAAATGTTG-3'
R2: 5'-GTGAGTCGTATTAAGCTTGGCACTGG-3'
F3: 5'- CTTCGAATTCGTAATCATGGTCATAGCTGT-3'
R3: 5'- Biotine-TTTACGGTTCCTGGCCTTTTGCTG-3'

where F stands for forward primer and R for the reverse one, and the numbers 2 and 3 refers to handle A and handle B, respectively. The reaction mixture (50  $\mu$ l) and PCR cycling conditions are the same as those used for the synthesis of the templates of TERRA sequences (see *Table 3.1* and *Table 3.2*). The obtained reaction products were purified using the *Wizard*® *SV Gel and PCR Clean-Up System* (Promega, Ref. A9281), eluting each sample in 50  $\mu$ l of nuclease-free water. Their sizes were checked by 1% agarose gel electrophoresis.

Reagent	V(µl)
Ultrapure Milli-Q-filtered water	22.0
Modified pUC18 plasmid (10 ng/µl)	1.0
Forward Primer (Integrated DNA Technologies, 10 µM)	1.0
Reverse Primer (Integrated DNA Technologies, 10 µM)	1.0
SapphireAmp® Fast PCR Master Mix (Takara ClonTech, Ref. RR350A)	25.0

 TABLE 3.1. PCR mixture for DNA telomeric G4-forming sequence amplification.

Cycle		Temperature (°C)	Time (min)
Hot Start Automatic		94.0	3
Start cycle (x 40)	Denaturation	95.0	0.5
	Annealing	55.0	0.5
	Elongation	72.0	1.5
Temperature		72.0	5
Store		4.0	8

<b>TABLE 3.2.</b>	PCR cycles	for DNA t	telomeric	G4-forming	sequence	amplification.
	, ,					

## **4.** <u>Labelling of the handles</u>.

Handle A was labelled at its 3' end with a very short tail of digoxigenin-dUTPs using the *DIG Oligonucleotide Tailing Kit, 2nd Generation* (Roche, Ref. 3353583910) at 37°C during 15 min. Then, the sample was purified using the *Wizard SV Gel and PCR Clean-Up System* (Promega, Ref. A9281). Handle B was

labelled with a biotin using a 5'-biotinylated primer (primer R3) during its PCR amplification.

5. <u>Synthesis of ssRNA sequences</u>.

The ssRNA sequences were synthesised by in vitro transcription of the doublestranded (ds) DNA obtained previously using Sp6 RNA polymerase (New England Biolabs; 3h, 40°C), whose promoter sequence was incorporated at the 5' end of the forward PCR primer (primer F1). The obtained transcripts were purified using RNeasy MinElute Cleanup Kit (Qiagen, Ref. 74204), eluting the samples in 20  $\mu$ l of RNase-free water. Its size was checked by 6% denaturing polyacrylamide/urea gel electrophoresis.

Rea	V(µl)		
RNase-free water		14.80	
Sp6 polymerase buffer	2.50		
rNTPs dil 1:10	А		
	С	1.25	
	G		
	U		
DNA template	1.20		
RNase Inhibitor (I	0.50		
Sp6 polymerase (	1.0		

 TABLE 3.3. Mixture reaction for in vitro transcription of DNA G4-forming sequences.

6. <u>Annealing</u>.

The final assembly is achieved by mixing equimolar amounts of the corresponding ssRNA and the two DNA labelled handles and annealing them in the presence of 100 mM potassium chloride, which favours the G4 stabilization, in a PCR apparatus following the procedure specified in *Table 3.4*. A schematic representation of the whole process is shown in *Figure 3.2*.

Cycle	Temperature (°C)	Time (min)	
Denaturation	90.0	5	
Annealing	85.0	10	
	62.0	90	
	52.0	90	
	20.0	$\infty$	

 TABLE 3.4. Annealing procedure of TERRA assemblies for OT experiments.



FIGURE 3.2. Schematic representation of the steps involved in the synthesis of the *TERRA reference contruct (n=5) and the RNA control systems 1 (n=6) and 2 (n=7) for OT experiments.* The single-stranded (ss) DNA of the 5' hybrid handle was labelled at its 3'end with a very short tail of digoxigenin-dUTP (represented as D) whereas the ssDNA of the 3' hybrid handle was labelled during the PCR amplification using a 5'-biotinylated (represented as B) primer (primer R3). The primers flanking the telomeric repeats, R2 and F3, start at -4 and +4 nucleotides away from the telomeric repeats, respectively. The final molecular construct has as a result extra ssRNA on both sides of the sequence (GGGUUA)<sub>n</sub>, being n= 5, 6 or 7 (represented in pink in (D)).

### 3.2.1.2. RNA control system 3

The synthesis protocol of the *RNA control system* 3 – i.e. molecules with a central ssRNA fragment of five GGGUUA repeats flanked by four extra nucleotides (UAUA) at the 5' end and by ten extra nucleotides (UAGUCUUCGA) at the 3'end – follows the same steps described in the previous subsection except for the hybrid

handle B (see *Fig. 3.3*), which is now shorter (395 bp) to release more random ssRNA at the 3' end of the central fragment (ten ribonucleotides instead of four). Then, the forward primer F3 is substituted by F3', whose sequence is: F3': 5'-ATTCGTAATCATGGTCATAGCTGTTTCCTG-3'.



**FIGURE 3.3.** Schematic representation of the steps involved in the synthesis of the RNA control systems 3 for OT experiments. The same protocol shown in *Figure 3.2* was followed but substituting forward primer F3 by F3', which releases 10 nt of ssRNA instead of 4 nt.

### 3.2.2. Synthesis of DNA analogue molecules

DNA analogue molecules to the TERRA reference construct were prepared for comparison purposes. The central ssDNA fragment consists of five repeats of the telomeric sequence GGGTTA flanked by four extra nucleotides at both sides: TATA at the 5' end and TAGT at the 3' end. The handles are made up of two complementary DNA strands with the same sequences as those used in the TERRA hybrid handles. As the transcription step is not needed now, the forward primer F1 does not include the Sp6 promoter at its 5' end. The rest of the protocol follows the same steps as previously described in *subsection 3.2.1.1*. A schematic representation of the whole process is shown in *Figure 3.4*.



**FIGURE 3.4.** *Schematic representation of the steps involved in the synthesis of the DNA analogous molecules to TERRA reference construct for OT experiments.* The same protocol shown in Figure 3.2 was followed but without the Sp6 promoter at the 5' end of the forward primer F1 (now renamed as F1' in (A)).

### 3.2.3. Single-molecule experiments

### 3.2.3.1. Optical Tweezers setups

Measurements have been performed using two optical tweezers (OT) setups with different measurement stability and subsequent resolution, which further served to crosscheck results. The first one is a table-based instrument (*Fig.3.5*), whereas the second one, so-called *miniTweezers* (mT, *Fig.3.6*), is a compact instrument suspended from the ceiling with reduced mechanical drift and larger measurement stability over time due to the reduction of optical path lengths [256, 260].



**FIGURE 3.5.** *Table-based OT instrument used in this thesis to perform force-induced unfolding experiments.* (A) Frontal view of the whole set-up. (B) Detail of the central part of the instrument in which optical trapping is performed, being highlighted one of the water-immersion objectives (1), the microfluidics chamber (2) and the piezo-electric actuators that allow the displacement of the chamber (3). (C) Detail of the right side where one of the PSDs (4) and the CCD camera (5) is located.

Beyond these differences, both instruments are based on the same optical design: two diode lasers (the former, 200 mW at maximum power, 835 nm

wavelength from *JDS Uniphase* and the latter, 250 mW and 808 nm from *Lumics*) brought to the same focus with orthogonal polarizations, which allow their optical paths to be separated using polarizing beam splitters. Beams are focused according to a counter-propagating configuration through opposing water-immersion objectives, 60X, NA=1.20 (*Nikon CFI Plan-Apochromat* for the table-based setup and *Olympus, UPLSAPO 60XW* for the mT) to form the optical trap in a microfluidics chamber, which also contains a micropipette. The light exiting the trap from each beam is collected by the opposite objective lens and redirected to position-sensitive detectors, which monitor the three force components acting on the trapped bead. For the calibration based on the conservation of light momentum, it is necessary that all the light exiting the trap is collected; the numerical aperture of the beams is then reduced to approximately 0.5, which allows the objective lenses to act as both light focusing elements and collectors.



**FIGURE 3.6.** *Minitweezers instrument used in this thesis to perform force-induced unfolding experiments.* (A) Frontal view of the whole set-up, where the main parts have been highlighted: (1) head of the mT, which can be suspended from the ceiling by means of a system of pulleys, (2) electronic tower and (3) laser driver box. (B) Detail of the interior of the head of the instrument where all the components placed in the optical table in the previous set-up are compacted.

Sample force-extension curves recorded with each instrument are shown along section 3.3 (compare, for example, replicate curves of *Fig. 3.9* with those of *Figs. 3.14-3.17*). Although both instruments gave rise to similar results

(approximate force resolution of 0.1 pN and a distance resolution near 0.5 nm), the mT allowed faster valid data collection due to its improved stability. Specifically, force-extension curves, enveloped by thermal noise, were less influenced by artifactual noise enhancing the discernibility of the rupture events associated with the unfolding of the RNA structures.

#### 3.2.3.2. Force-induced unfolding experiments

To perform force-induced unfolding experiments, a single nucleic acid construct made up of a central single-stranded human telomeric G4-forming sequence flanked by duplexes – hybrid in the case of RNA systems – was tethered by opposite ends between two dielectric polystyrene microspheres: an antidigoxigenin ( $\alpha$ -Dig) -coated bead, optically trapped in the laser tweezers, and a streptavidin-coated bead, held by suction on top of a micropipette (see *Fig.3.7*) – located inside a microfluidics chamber—. Streptavidin-coated, 2.10 µm-diameter beads and G-protein-coated, 2.88 µm-diameter polystyrene beads, purchased from Spherotech (Libertyville, IL), were used.  $\alpha$ -Dig was linked to the G-coated beads using dimethyl pimelimidate (DMP) dihydrochloride.

In the table-based instrument, force on the studied molecule was generated by moving the micropipette —fixed to the microfluidics chamber— relative to the stationary optical trap through a piezo-controlled stage. In the miniTweezers, it is generated by moving the trap relative to the stationary micropette through bending the lasers' optical fibres using piezoelectric crystals. The extension of the molecule was determined, in both systems, from the distance between the centres of the beads. Stretch-relaxation cycles were performed at  $10.2 \pm 2.7$  nm/s, which corresponds to a pulling rate of  $5 \pm 1$  pN/s near the rupture events. All tweezers experiments were carried out at room temperature ( $23 \pm 2$  °C) in 10 mM Tris-HCl (pH 7.8) supplemented by 100 mM KCl and 1 mM EDTA. To avoid RNA degradation, diethylpyrocarbonate (DEPC)-treated water was used (DEPC to 0.1% v/v; incubation at 37°C overnight followed by autoclaving).



**FIGURE 3.7.** *Experimental set-up for OT measurements.* **(A)** Experimental configuration for TERRA reference construct and RNA control systems. A single RNA molecular construction was tethered by opposite hybrid duplex ends between two polystyrene microspheres: and  $\alpha$ -digoxigenin-coated bead, optically trapped, and a streptavidin-coated bead, held by suction on top of a micropipette. **(B)** Experimental configuration for DNA analogue molecules to TERRA reference construct. In this case, a single DNA molecular construction was attached to polystyrene beads by DNA duplex ends.

### 3.3. Results

We singled out several repetitions, n, of the human telomeric sequence GGGUUA and added extra nucleotides (nts) of non-G-rich, random sequences on both the 3' and 5' ends of ssRNA (see *Materials and Methods* section for further details on the synthesis). We chose 4 < n < 8, first, to make sure that the sequence can only fold into a single G4 conformation and, second, to study the effect of additional repetitions in the formation of this secondary structure.

For reference purposes, we will focus on the five-repeat molecule (n = 5) with four extra nucleotides on both the 3' and 5' ends. Earlier experiments by circular dichroism and NMR with this RNA fragment demonstrated its potential to form an intramolecular parallel G4 [267]. We will next show that the 8-nt ssRNA excess together with an extra GGGUUA repetition provides ample configurational flexibility to this telomeric sequence to fold into several conformations. In fact, this 38-nt sequence can give rise to: (i) a four-stranded, **G4 structure** interacting with different lengths of excess ssRNA on both ends (see *Fig.3.8*), (ii) a three-stranded, **G-triplex** (G3) similarly coexisting with two ssRNA lateral fragments of variable lengths and (iii) either a **hairpin-like structure or more general condensed conformations of ssRNA**. These structural and subsequent kinetic scenarios as well as the probabilities for the conformations discussed above can only be analysed by real-time, single-molecule experiments because bulk approaches in which an ensemble of equal preparations take place can only give rise to average signals thus hindering the access to the less populated conformations [285, 329].



**FIGURE 3.8.** *Possible G-quadruplex assembly configurations for the TERRA reference construct.* It is made up of five GGGUUA repeats with eight random nucleotides spread at the 3' and 5' ends to a total of 38 nucleotides. In green, random ribonucleotides are indicated; in blue, telomeric repetitions that forms a TERRA G4 and, in orange, the remnant GGGUUA.

We have used OT to study single-molecule mechanical folding and unfolding processes under controlled *in vitro* conditions. Rupture events become observable as jumps over the characteristic force-extension curves of the double-stranded hybrid handles, whose elastic response is similar to those of dsDNA and dsRNA [30, 50]. Although not deterministically, conformations can be tracked by force and

extension signals and the interference or compatibility among the different conformations can be distinguished by their simultaneous occurrence in single unfolding/refolding traces, as reflected in the force-extension curves and in the further statistical analysis. In this regard, we registered **two types of behaviours** in the force-extension experiments: **curves with single or two consecutive rupture events**, which will be analysed next.

#### 3.3.1. Single rupture events

Force-extension curves with only one rupture event represented the typical behaviour. *Figure 3.9* shows the results of several single molecules with the reference sequence. As mentioned above, the traces mostly reveal the characteristic elasticity of the double-stranded handles interrupted by rupture events at intermediate forces. Some of the traces revealed the overstretching transition of the hybrid handles, what confirms that only one molecular construct has been attached between the beads [330]. Rupture signatures correspond to unfolding events of different species, as judged by their force and extension values and also by the relaxation behaviour of the traces (see also insets in *Fig.3.9*).



**FIGURE 3.9.** Replicates of force-extension curves of the TERRA reference construct in which **(A)** a G-quadruplex/triplex or **(B)** single-stranded condensed conformations are probably unfolded and refolded (blue and green lines, respectively).

Three different types of rupture events can be distinguished and they can be assigned, *a priori*, to the unfolding of different structures by estimating the released ssRNA. Let  $L_{nt}$  be the mean, effective distance between successive nucleotides in ssRNA,  $d_G$  the solution-structure distance between first and last nucleotides in the G-structure (PDB ID: 2KBP) [163, 331],  $N_G$  the number of nucleotides in the folded structure and *x* the rupture extension measured in the OT experiments, then:

$$x = N_G \times L_{nt} - d_G \tag{1}$$

Applying equation (1), the estimated released extension upon the G4 and the G3 ruptures are: 21 (nt) ×  $L_{nt}$  – 1.5 (nm) = 9.2 nm and 15 (nt) ×  $L_{nt}$  – 1.9 (nm) = 5.8 nm, respectively, where we have used  $L_{nt}$  = 0.51 nm/nt following Yangyuoru et al. [331] (see *Fig. 3.10*).



**FIGURE 3.10.** *End-to-end distance measurements of non-canonical conformations.* **(A)** TERRA G-quadruplex and **(B)** putative G-triplex-like structure. *F* stands for the tension applied on the structure in each single-molecule experiment. Red arrows represent pulling direction and orange circles the first and last nucleotide of the sequence.

The interphosphate distance actually depends on the force at which the rupture takes place because it is related to the ssRNA elastic response. For ssDNA, Alemany et al. [332] found an effective length across forces of  $L_{nt} = 0.58$  nm/nt according to fittings of the force-extension curves of short ssDNA to inextensible elastic models. Since the inter-phosphate crystallographic distance in DNA (~0.7 nm/nt) is longer than in RNA (~0.59 nm/nt) [30, 313, 330], it is expected that the

effective distance between phosphates in the ssRNA under tension complies with the number used above. Then, we can conclude that the most probable rupture events arise from the unfolding of a G4/G3 (*Fig.3.9 (A)*) as well as of more general structures caused by ssRNA self-interactions (*Fig.3.9 (B)*). This interpretation agrees with previous studies: G3 and G4 rupture events were located primarily in the range of 20-30 pN [331], whereas single-stranded interactions were noticed between 4-10 pN, depending on ionic conditions and temperature [260, 332-335].

To *in situ* confirm this probabilistic association we recorded data over almost 300 rupture events of this reference molecular construct. In the assays, we sometimes subjected the same molecule to more than one stretch-relaxation cycle, each separated by a pause of 5-10 s in which the force had been relaxed to zero pN. During this pause, the molecule had extra-time to attempt refolding in case it had not happened during the relaxation pathway. The statistical analysis of the force-extension curves for single-rupture events is presented in *Figure 3.11*.



**FIGURE 3.11.** *Statistics of single rupture events in the TERRA reference construct.* **(A)** Rupture force and **(B)** unfolding extension distribution for TERRA molecules with five GGGUUA repeats flanked by four random ribonucleotides at both sides (*N*=299). Mean  $\pm$  standard error (s.e.) and standard deviation (s.d.) according to Gaussian fitting are: (A) 8.8  $\pm$  0.2 and 1.3 pN (pink line), 22.3  $\pm$  0.5 and 8.6 pN (light blue line) and (B) 8.2  $\pm$  0.2 and 2.4 nm. In (A), dark blue line corresponds to Gaussian double-peak fitting.

Panel (A) shows the rupture force distribution, where two major populations can be distinguished centred at nearly 9 and 23 pN. According to these forces, the proposed structures as majorly responsible for these two populations would be hairpin-like structure or condensed ssRNA (due to ssRNA self-interactions), for the lower and narrower force peak, and the G4/G3 unfolding, for the higher and wider force peak. Panel (B) shows the unfolding extension histogram.

To further confirm the interpretation of these two statistical populations, two filters in extension have been applied, based on the theoretical dimensions of G3 and G4 structures [331]: a low-pass filter with a cut-off extension of 12 nm, which favours rupture events corresponding to the unfolding of the G4/G3 conformations, and a high-pass filter above 12 nm, which favours the populations corresponding to the unfolding of hairpins and those with more random behaviour related with ssRNA self-interactions [326-328]. *Figure 3.12* shows the resulting force histograms. The populations correlate with the expected conformations. Namely, *Figure 3.12 (A)* shows a broad peak centred at approximately the same force as that shown for the wide population in *Figure 3.11 (A)*, being attributed to the G4/G3 unfolding. On the other hand, *Figure 3.12 (B)* populates majorly the low forces, including a peak at 7-8 pN similar to the narrow one in *Figure 3.11 (A)*, which has been associated mainly to the unfolding of generally condensed ssRNA structures.

The G3 has been described as an intermediate, possibly in-pathway conformation in the folding/unfolding of the G4 with similar stability to this fourstranded structure [318-320, 325, 336-339], although both conformations may not be resolved within the same single unfolding pathway [331, 340-342]. These forces for the G4/G3 are comparable within experimental error to those reported in [331], as expected, since we used a similar loading rate (see *Materials and Methods*), hence working similarly away from equilibrium [323, 343]. They are also higher than those found in [267] because in that study rupture events were not classified but considered within the same statistics hence giving rise to a mean force between those of the two populations distinguished in *Figure 3.11 (A)*.



**FIGURE 3.12.** *Rupture force distributions for the different conformations in the TERRA reference construct.* **(A)** G-Quadruplex and G-triplex (N= 241, 80.6%), data filtered in the range [0, 12] nm. Mean ± s.e. and s.d. according to Gaussian fitting is 22.0 ± 0.6 and 9.3 pN. **(B)** Duplex or other condensed conformations (N=58, 19.4%), data filtered in the range (12, 26] nm. Median ± s.d. according to a log-normal distribution is 8.7 ± 4.4 pN. Peak value (mode) ± s.e. is 7.2 ± 1.2 pN.

With regards to the peak at low force, apart from the randomly-condensed conformation [260, 332-335, 344], a very weakly stable hairpin may be formed (see insets in *Fig. 3.12 (B)* and *Fig. A.1* in Appendix A). This duplex possesses two G-bubbles and G-U and A-U pairings in the stem, according to *mfold* simulations with the 5-repeat reference sequence ( $\Delta G$ = -0.40 kcal/mol at 1 M NaCl) [345]. This duplex does not correspond to the G-hairpin, described as a possible G3 or G4 intermediate that results from two repetitions of the GGGUUA sequence [318-320]. It extends over almost the entire ssRNA sequence but its free energy does not overtake that of the thermal fluctuations (-0.7 *k*<sub>B</sub>*T*, being *k*<sub>B</sub> the Boltzmann constant and *T* the absolute temperature), which yields an equilibrium rupture

force much lower than 1 pN (see Appendix A). This makes the duplex very improbable according to the rupture force distribution of *Figure 3.11 (A)*.

#### 3.3.1.1. Comparison with the DNA analogous system

*Figure 3.13* shows the statistics of unfolding/refolding experiments with an equivalent molecular construct to the reference system entirely made of DNA, thus preserving sequence and length of both the central single-stranded segment and the double-stranded handles for attachment to beads (see *Fig. 3.4* and *Fig.3.7 (B)*).

Its rupture force distribution (*Fig.3.13 (A*)) exhibits a single broad peak centre at around the same force of that corresponding to the G4/G3. The peak force is shifted to the left by less than 2 pN with respect to the RNA distribution corresponding to the G4/G3 (*Fig.3.12 (A*)), in agreement with previous works that reported that DNA G4 unfold at lower force [270, 322-325].



**FIGURE 3.13.** Statistics of single rupture events in the DNA analogous system to the *TERRA reference one*. (A) Rupture force and (B) unfolding extension distribution for DNA molecules with five GGGTTA repeats flanked by four random nucleotides at both sides (N=163). Mean ± standard error (s.e.) and standard deviation (s.d.) according to Gaussian fitting are: (A) 20.4 ± 0.4 and 6.9 pN and (B) 8.6 ± 0.1 and 1.9 nm.

The most important feature of these control experiments is the absence of a peak near 9 pN, which were crosschecked with measurements in a high-resolution



instrument (see *Materials and Methods* section and compare curves from *Fig.3.9* and *Fig. 3.14*).

**FIGURE 3.14.** Replicates of force-extension curves of DNA molecules with five GGGTTA repeats in which a G-quadruplex/triplex is probably unfolded and refolded (blue and green lines, respectively).

It indicates that the potential of the RNA molecule to self-fold and thus block highly-organized structures is unique. In fact, stronger hybridization interactions in RNA than in DNA of equal sequence have been observed both in bulk and in single-molecule experiments previously [330]. Beyond electrostatic base-pairing considerations, RNA possesses higher tendency to self-interact due to the puckering and chemistry of the bases [313, 346], thus making G4-forming sequences in RNA more prone to exhibit conformational competition in the presence of extra single-stranded random sequences than in DNA. Moreover, the weakly stable hairpin that can be adopted by the TERRA reference tract is not possible in the analogous ssDNA ( $\Delta G > 0$ ), reflecting again the higher conformational potential of RNA relative to ssDNA [313].

#### 3.3.1.2. Comparison with the RNA control systems

To study the effect of both additional repetitions and the sequence of the neighbouring tracts in the formation of a single TERRA G4, three other RNA molecular constructs were synthesised: a 6-repeat and a 7-repeat molecules flanked by the same random sequences as previously used (44 and 50 nts, respectively) and a 5-repeat construction with more extra random RNA sequence (44 nts) (see *Materials and Methods* section). The qualitative behaviour is maintained for single rupture events, as reflected in their force-extension curves (*Fig. 3.15-3.17*): rupture events at both high forces, compatible with G4/G3, and low forces are observed.



**FIGURE 3.15.** Replicates of force-extension curves of TERRA molecules with six GGGUUA repeats in which **(A)** a G-quadruplex/triplex or **(B)** single-stranded RNA condensed conformations are probably unfolded and refolded (blue and green lines, respectively). Each panel corresponds to a different molecule.



**FIGURE 3.16.** Replicates of force-extension curves of TERRA molecules with seven GGGUUA repeats in which **(A)** a G-quadruplex/triplex or **(B)** single-stranded RNA condensed conformations are probably unfolded and refolded (blue and green lines, respectively). Each panel corresponds to a different molecule.



**FIGURE 3.17.** Replicates of force-extension curves of TERRA molecules with five GGGUUA repeats and with more extra random ssRNA in which **(A)** a G-quadruplex/triplex or **(B)** single-stranded RNA condensed conformations are probably unfolded and refolded (blue and green lines, respectively). Each panel corresponds to a different molecule.

The statistics over convergent populations of experiments are shown in *Figure* 3.18. For the 6-repeat molecule, *Figure* 3.18 (A1), two populations are distinguished, being the high-force distribution centred at a similar force (within experimental error) than that of the 5-repeat molecule previously studied (*Fig.3.11* (*A*)). However, the low force distribution presents this time a shift to higher forces, approaching the high-force peak compatible with the G4/G3 conformation. In agreement with this trend, this low-force distribution overlaps with that of the high-force for the 7-repeat molecule (see *Fig.3.18* (*B1*)). For the case of the 5-repeat molecule with extra random sequence, *Figure* 3.18 (*C1*), it is observed again the bimodal distribution shown in *Figures* 3.11 (*A*) and 3.18 (*A1*), with forces for the low and high force peak similar to those of the 6-repeat molecule (*Fig. 3.18* (*A1*)). This is expected since the number of nucleotides involved in these two molecules is identical (44 nt).

To further confirm these results, *Figure 3.19* presents the histograms resulting from the application of the filter previously used for the reference molecule. The three sequences display separate single-peak distributions: a high force peak at around 20 pN, which, as discussed, is attributed to the G4/G3 structures, and a low force peak. The latter increases with the number of nucleotides in the sequences: from ~9 pN for the reference sequence (*Fig. 3.12 (B)*), to ~15-18 pN for these longer molecules.

More importantly, from these histograms, we reveal that by increasing the number of repeats, the probability of stochastic blockage of G4/G3 formation monotonically increases (19.4%, 22.3% and 23.7% for the 5, 6 and 7-repeat molecules, respectively). The addition of more random sequence, as also shown in *Figure 3.18 (C1)*, does not follow this trend. In particular, the 5-repeat molecule decreases the probability of stochastic G4/G3 blockage from a 19.4% down to a 10.2% when 6 extra random nucleotides are added to the TERRA sequence. This is due to the fact that this molecule presents a lower number of Gs, which is a very promiscuous nucleobase thus tending to increase the probability of a condensate in RNA.



**FIGURE 3.18.** *Statistics of single rupture events in RNA control systems*. Rupture force (1) and unfolding extension distribution (2) for the RNA constructions of (A) 6 repeats (N=385), (B) 7 repeats (N= 317) and (C) 5 repeats with more random extra ssRNA (N= 205). Mean ± s.e. and s.d. according to Gaussian fitting are: (A1) 15.3 ± 1.0 and 5.9 pN (pink line), 26.8 ± 0.9 and 3.9 pN (light blue line); (A2) 8.9 ± 0.2 and 3.1 nml; (B1) 17.9 ± 0.5 and 8.5 pN; (B2) 8.5 ± 0.3 and 3.0 nm; (C1) 13.2 ± 1.9 and 4.1 pN (pink line), 23.5 ± 3.8 and 7.3 pN (light blue line) and (C2) 8.9 ± 0.1 and 1.9 nm. In (A1) and (C1) dark blue line corresponds to Gaussian double-peak fitting.



**FIGURE 3.19.** *Rupture force distributions for the different conformations in RNA control systems.* G-quadruplex and G-triplex formed by **(A1)** RNA molecules with six GGGUUA repeats (N= 299, 77.7%), **(B1)** RNA molecules with seven GGGUUA repeats (N= 242, 76.3%) and **(C1)** RNA molecules with five GGGUUA repeats and more extra random ssRNA (N= 184, 89.8%), data filtered in the range [0, 12] nm. Duplex or other condensed conformations formed by **(A2)** RNA molecules with six repeats (N= 86, 22.3 %), **(B2)** RNA molecules with seven repeats (N=75, 23.7%) and **(C2)** RNA molecules with five repeats and more extra random ssRNA (N= 21, 10.2%), data filtered in the range (12, 24] nm. Gaussian regressions have been used in A1 and C1 with mean and s.d.: (A1) 20.2 and 7.9 pN and (C1) 21.1 and 8.5 pN. Log-normal regressions have been applied to B1 and A2-C2 with median and s.d.: (A2) 15.1 and 14.7 pN, (B1) 19.7 and 8.7 pN, (B2) 14.0 and 15.3 pN, (C2) 18.3 and 9.9 pN. Peak values (mode) ± s.e. are: (A1) 20.2 ± 0.6 pN, (A2) 9.4 ± 1.3 pN, (B1) 16.9 ± 0.6 pN, (B2) 8.2 ± 0.9 pN, (C1) 21.1 ± 0.9 pN and (C2) 14.8 ± 0.9 pN.

#### 3.3.2. Consecutive rupture events

As discussed above, the formation of condensed ssRNA or the hairpin represented in *Figure 3.12 (B)* (insets) blocks the assembly of the G3 and the G4. Nevertheless, when one of these non-canonical structures assembles, it may interact with the excess ssRNA in our sequence. Then, the mechanical unfolding should provoke the separation of the extra ssRNA from the G4 or G3 before the full opening of the structure. According to the rupture forces and unfolding extensions obtained from force-extension curves as those represented in *Figures 3.20-3.23* for the TERRA molecules with 5-7 repeats and extra random sequence, this is the most probable scenario.



**FIGURE 3.20.** *Replicates of force-extension curves of TERRA molecules with five GGGUUA repeats in which sequential unfolding events take place*. Each panel corresponds to a different molecule. Blue traces correspond to stretching paths and green ones to relaxation paths.



**FIGURE 3.21.** *Replicates of force-extension curves of TERRA molecules with six GGGUUA repeats in which sequential unfolding events take place*. Each panel corresponds to a different molecule. Blue traces correspond to stretching paths and green ones to relaxation paths.



**FIGURE 3.22.** *Replicates of force-extension curves of TERRA molecules with seven GGGUUA repeats in which sequential unfolding events take place*. Each panel corresponds to a different molecule. Blue traces correspond to stretching paths and green ones to relaxation paths.



**FIGURE 3.23.** *Replicates of force-extension curves of TERRA molecules with five GGGUUA repeats and more extra random ssRNA in which sequential unfolding events take place.* Each panel corresponds to a different molecule. Blue traces correspond to stretching paths and green ones to relaxation paths.

The statistical analysis of the 5-repeat TERRA molecule when two consecutive rupture events are observed in the force-extension curves, *Figure 3.24*, confirms this interpretation: the first rupture peak complies with low forces and low extensions, *Figure 3.24 (A)*, and the second one with high forces and typical extensions for the G4 and the G3, *Figure 3.24 (B)*.

The statistics corresponding to the 6 and 7-repeat TERRA molecules and that with 5 repeats and extra random sequence, *Figure 3.25*, also follow this trend. Importantly, it is observed that the probability of interaction of the formed G4 with the excess single-stranded increases with the number of added nucleotides: 12.6%, 17.4% and 26.1% for 5, 6 and 7 repeats, and 27% for the 5-repeat molecule with

more extra random ssRNA, a trend which was expected for longer molecules since self-interactions allowed by the bending elasticity are more probable.



**FIGURE 3.24.** Statistics of single-molecule sequential unfolding events in the TERRA reference construct. (A) Rupture force and unfolding extension distributions for the first rupture event. Median and s.d. according to a log-normal fitting for this force distribution is 14.9 and 8.3 pN (peak value (mode)  $\pm$  s.e., 11.9  $\pm$  0.5 pN). Median and s.d. for the unfolding extension distribution is 7.3 and 4.4 nm (peak value (mode)  $\pm$  s.e., 5.7  $\pm$  0.4 nm). (B) Rupture force and unfolding extension distributions for the second rupture event. Mean  $\pm$  s.e. and s.d. for this force distribution, according to a Gaussian fitting (red line) is 27.4  $\pm$  0.5 and 7.3 pN. Median and s.d. according to log-normal fitting for the unfolding extension distribution are 7.6 and 2.8 nm (peak value (mode)  $\pm$  s.e., 6.7  $\pm$  0.2 nm).



**FIGURE 3.25.** *Statistics of single-molecule sequential unfolding events in control sequences.* Rupture force distributions for RNA molecules with **(A1)** six GGGUUA repeats (*N*=81), **(B1)** seven GGGUUA repeats (*N*= 112) and **(C1)** five GGGUUA repeats and more extra random ssRNA (*N*=76). Median and s.d. for the first peak (blue) according to lognormal fittings are (A1) 14.9 and 8.2 pN, (B1) 16.1 and 8.4 pN and (C1) 13.4 and 6.3 pN. Peak values (mode)  $\pm$  s.e. are: 12.0  $\pm$  0.6 pN, 13.2  $\pm$  1.5 pN and 11.2  $\pm$  0.2 pN, respectively. Mean  $\pm$  s.e. and s.d. for the second peak (red) according to a Gaussian fitting are (A1) 25.3  $\pm$  0.5 and 6.4 pN, (B1) 24.7  $\pm$  0.3 and 7.4 pN and (C1) 22.5  $\pm$  0.5 and 7.0 pN. Unfolding extension distributions for RNA molecules with **(A2)** six GGGUUA repeats, **(B2)** seven GGGUUA repeats and **(C2)** five GGGUUA repeats and more extra random ssRNA. Mean  $\pm$  s.e. and s.d. for the first peak (blue) according to Gaussian fittings are: (A2) 7.30  $\pm$  0.04 and 2.8 nm, (C2) 7.2  $\pm$  0.3 and 3.1 nm; median and s.d. according to a log-normal fitting is (B2) 7.2  $\pm$  4.2 nm (peak value (mode)  $\pm$  s.e., 5.6  $\pm$  0.4 nm). Mean  $\pm$  s.e. and s.d. for the second peak (red) according to a log-normal fitting is (B2) 7.2  $\pm$  4.2 nm (peak value (mode)  $\pm$  s.e., 5.6  $\pm$  0.4 nm). Mean  $\pm$  s.e. and s.d. for the second peak (red) according to a log-normal fitting is (B2) 7.2  $\pm$  4.2 nm (peak value (mode)  $\pm$  s.e., 5.6  $\pm$  0.4 nm). Mean  $\pm$  s.e. and s.d. for the second peak (red) according to a Gaussian fitting are (A2) 6.9  $\pm$  0.2 and 3.1 nm; (B2) 7.0  $\pm$  0.4 and 3.5 nm and (C2) 8.0  $\pm$  0.2 and 2.3 nm.

## **3.4. Discussion**

A strong conformational competition between condensed ssRNA and noncanonical structures has been found in the folding pathways of short TERRA molecules. In particular, it has been demonstrated that if a ssRNA condensed form is established in a RNA segment with *n* repetitions of the sequence GGGUUA, 4 < n< 8, it blocks the assembly of both the G4 and the G3, even considering that these latter conformations are more stable than the former for equal number of nucleotides. The number of TERRA repeats studied here is representative of the stochastic blockage phenomena since a number equal or greater than eight increases the configurational probability space and thus, the complexity of the simultaneous configurations.

The probability of G4/G3 formation, P(G-QT), can be calculated based on the next heuristic inference:

$$P(G-QT) = P(\text{consecutive RE}) + P(\text{single RE}) \times P(G-QT|\text{single RE})$$
(2)

where *G-QT* stands for G4 or G3, *RE* stands for rupture event and the vertical bar denotes conditional probability. The results, along with the previously presented experimental probabilities, are shown on *Table 3.5*. It is observed that the number of repeats hardly affects the probability of formation of the G4/G3, slightly above an 80%, and that the addition of random nucleotides flanking the hexanucleotide repeats attenuates the stochastic blockage, raising the probability of the non-canonical structures to a 90%.

	(GGGUUA) <sub>5</sub>	(GGGUUA) <sub>6</sub>	(GGGUUA)7	(GGGUUA) <sub>5</sub> + ssRNA
P(Consecutive RE)	0.126	0.174	0.261	0.270
P(Single RE)	0.874	0.826	0.739	0.730
P(G-QT   Single RE)	0.806	0.777	0.763	0.898
P(G-QT)	0.830	0.815	0.825	0.925
P(random coil)	0.170	0.185	0.175	0.075

**TABLE 3.5.** *Probability of G-quadruplex/triplex formation*. P, probability; RE, rupture event; G-QT, G-quadruplex or G-triplex; "|", conditional probability.

Single rupture events as those represented in *Figures 3.9* and *3.15-3.17* may correspond to separate events sequential but proximal in force, which would match to the release of one arm of the G4 rapidly followed by the rupture of the G3 into ssRNA. Neither these events nor the G-hairpin unfolding [318-320] could be distinguished within our experimental resolution. Although some force-extension curves with two rupture events in our experiments could be interpreted in this fashion, we have not found sufficient statistical significance. Then, we conclude that, if in-pathway, G3 and G-hairpin breakage takes place almost simultaneously after the G4 disassembly in our experiments; that is, the sequential release of the GGGUUA arms of the G4 are detected within the same rupture event in the force-extension curves at the herein used experimental velocity [325]. A triangular, three state kinetic model, as that proposed elsewhere [331, 340, 342, 347], is compatible with our results (*Fig.3.26*).



FIGURE 3.26. Schematic representation of the triangular, three state kinetic model of the mechanical unfolding of TERRA molecules.

The dynamic behaviour of the rupture events has another feature in the stretchrelaxation cycles: the reversibility. Refolding during relaxation is the normal behaviour but the mechanical hysteresis is typically lower for the duplexes with respect to the G4 or G3 conformations, as measured by the degree of overlapping of the stretch and relaxation traces (see *Fig. 3.9* and *3.15-3.17*). The refolding is more likely for the duplex than for the G3 or G4 since these latter conformations involve the ordered assembly of three and four strands, respectively, within similar time frames, which is less probable than the assembly of randomly-condensed ssRNA or duplexes, with higher entropy. Then, it is expected that the non-canonical conformations involve refolding further from equilibrium than simpler conformations at the same loading rate.

In these conditions, the free energy of unfolding of each species can be recovered by using *Jarzynski's equality* (2), which evaluates this equilibrium thermodynamic potential increment from non-equilibrium work measurements as follows [348]:

$$\Delta G_{unfold} = -k_B T \ln \sum_{i=1}^{N} \frac{1}{N} \exp\left(-\frac{W_i}{k_B T}\right)$$
(3)

where  $\Delta G_{unfold}$  is the free-energy difference between the folded and unfolded states of the corresponding conformation;  $k_BT$ , the product of the Boltzmann's constant and the temperature (0.592 kcal/mol at 298 K); *N*, the number of unfolding data compiled for the conformation under study and *W*, the amount of non-equilibrium work done to unfold the structure (G4/G3 or RNA random coils). The latter was calculated from the force and extension signatures in the stretching paths of the force-extension curves for the different conformations in each molecular preparation. *Tables 3.6* and *3.7* collect these data for the non-canonical structure and the condensates, respectively. *Tables 3.8* and *3.9* provide the resulting Gibbs free energy increments. Right columns show energies per nucleotide.
	W [kcal/mol] (mean ± s.e.)	W[kcal/(mol·nt)] (mean ± s.e.)
(GGGUUA) <sub>5</sub> ( <i>N</i> =241)	25.7 ± 0.8	$1.22 \pm 0.04$
(GGGTTA) <sub>5</sub> ( <i>N</i> =163)	28.4 ± 1.0	1.35 ± 0.05
(GGGUUA) <sub>5</sub> + random ssRNA ( <i>N</i> =184)	25.7 ± 0.8	$1.22 \pm 0.04$
(GGGUUA) <sub>6</sub> ( <i>N</i> =299)	$23.4 \pm 0.6$	$1.12 \pm 0.03$
(GGGUUA)7 (N=242)	22.5 ± 0.7	$1.07 \pm 0.03$

**TABLE 3.6.** *Non-equilibrium work done to unfold the G-quadruplex structure* (s.e., standard error).

	W [kcal/mol] (mean ± s.e.)	W[kcal/(mol·nt)] (mean ± s.e.)
<b>(GGGUUA)</b> <sub>5</sub> ( <i>N</i> =58)	37.6 ± 3.2	$0.99 \pm 0.08$
<b>(GGGTTA)</b> <sub>5</sub> ( <i>N</i> =21)	46.5 ± 4.4	$1.06 \pm 0.10$
(GGGUUA) <sub>5</sub> + random ssRNA (N=86)	37.1 ± 2.0	$0.84 \pm 0.05$
<b>(GGGUUA)</b> <sub>6</sub> ( <i>N</i> =75)	35.8 ± 2.2	$0.72 \pm 0.04$

**TABLE 3.7.** *Non-equilibrium work done to unfold condensed structures* (s.e., standard error).

	ΔG <sub>unfold</sub> [kcal/mol] (mean ± s.e.)	$\Delta G_{unfold}$ [kcal/(mol·nt)] (mean ± s.e.)
(GGGUUA) <sub>5</sub> ( <i>N</i> =241)	7.5 ± 0.4	$0.36 \pm 0.02$
(GGGTTA) <sub>5</sub> ( <i>N</i> =163)	6.9 ± 0.6	$0.33 \pm 0.03$
(GGGUUA) <sub>5</sub> + random ssRNA ( <i>N</i> =184)	8.1 ± 0.3	$0.39 \pm 0.02$
(GGGUUA) <sub>6</sub> ( <i>N</i> =299)	9.3 ± 0.3	$0.44 \pm 0.01$
(GGGUUA)7 (N=242)	$7.4 \pm 0.4$	$0.35 \pm 0.02$

**TABLE 3.8.** Change in free energy of G-quadruplex structure calculated by the Jarzinsky's equation (s.e., standard error).

	ΔG <sub>unfold</sub> [kcal/mol] (mean ± s.e.)	ΔG <sub>unfold</sub> [kcal/(mol·nt)] (mean ± s.e.)
(GGGUUA) <sub>5</sub> ( <i>N</i> =58)	12.3 ± 0.6	$0.32 \pm 0.02$
(GGGTTA) <sub>5</sub> ( <i>N</i> =21)	12.6 ± 0.6	$0.29 \pm 0.01$
(GGGUUA) <sub>5</sub> + random ssRNA (N=86)	$10.8 \pm 0.6$	$0.25 \pm 0.01$
(GGGUUA) <sub>6</sub> ( <i>N</i> =75)	11.1 ± 0.6	$0.22 \pm 0.01$

**TABLE 3.9.** Change in free energy of random coil structures calculated by the Jarzinsky's equation (s.e., standard error).

Results for the DNA analogue 5-repeat molecule yields  $\Delta G_{unfold} = -(6.9 \pm 0.6)$  kcal/mol, which is similar within experimental error with earlier bulk reports that yielded -7.1 kcal/mol for the parallel G4 assembled with a four-repeat molecule [349]. This consistency serves as a control check, both from a qualitative and a quantitave viewpoint, for our results at the single-molecule level. RNA G4 has always higher Gibbs free energy than the DNA analogue, also consistent with earlier literature [331].

### **Chapter 4**

### Collaborations

## 4.1. Mesoscopic model for DNA G-quadruplex force-induced unfolding

Single-molecule characterization is currently limited by the temporal and spatial resolutions of the experimental techniques. Resorting to theoretical modelling to obtain complementary information about the system under study is necessary. Focusing on the G4 structure, its dynamics has been modelled at different temporal and length scales, from quantum calculations [350, 351] and molecular dynamics simulations [325, 352-356] to mesoscopic approaches [357, 358]. However, only mesoscopic models allow the simulation of loading rates similar to the experimental ones and, until now, none has been applied to characterize the mechanical unfolding of a single G4.

From the force-induced unfolding experiments obtained in this thesis on a human DNA telomeric sequence capable of forming a G4 (see Chapter 3), a mesoscopic model to describe, for the first time, both its mechanical and thermal stability has been developed by the group led by Dr. A. Fiasconaro and Dr. F. Falo at Zaragoza University [359]. A previous atomistic study of telomeric DNA G4 performed in this theoretical research group [325] and the melting temperatures of DNA G4 [349] are also taken into account to define the model parameters.

As shown in *Figure 4.1*, the mesoscopic model proposed here considers each nucleotide of the ssDNA chain (depicted in green and blue), as well as each

monovalent cation located at the central channel of the G4 structure (in yellow), onto a single bead (see *Appendix B* for a detailed explanation of the theory behind the model). This simplification will allow studying the mechanical unfolding at lower velocities than permitted in atomistic simulations [325] and, then, perform direct comparison with experimental results.



**FIGURE 4.1.** *Scheme of a parallel G4 assembly.* Each nucleotide is represented by a single bead (not to scale). Green beads correspond to guanines; blue ones, to the nucleotides that form the loops; i.e. TTA, and yellow spheres to monovalent cations. Beads highlighted in red are those involved in the pulling simulations (see the text for details). For simplicity, the twist between successive G-tetrad planes is not represented.

Being *bead 1* and *bead 21*, the first and last nucleotides in the telomeric DNA chain (highlighted in red in *Fig. 4.1*), pulling simulations were performed by fixing the latter to the origin of coordinates, whereas the former is attached to a harmonic spring whose end is displaced at constant velocity, v (see *Fig.4.2*). The component of the force along the pulling direction was calculated as:

$$F = k_A \cdot \left( x_A(t) - x_1(t) \right) \tag{4.1}$$

where  $k_A$  is the elastic constant of the spring and  $x_A(t)$  and  $x_1(t)$  are the components of the distance along the pulling direction of the spring end (*point A*) and the pulled *bead 1*, respectively, as depicted in *Figure 4.2*.

The elastic constant,  $k_A$ , was set to 0.4 pN/nm, in accordance to the experimental value estimated from the slope of the force-extension curves in the enthalpic elasticity regime of the dsDNA handles before the rupture event

corresponding to the unfolding of the DNA G4 (N=20, mean ± s.d.: 0.4 ± 0.1 pN/nm; see *Fig. 4.3*); i.e. a sharp jump over the characteristic force-extension curve of the double-stranded handles.



**FIGURE 4.2.** *Scheme of the pulling simulations.* The first and last nucleotides in the telomeric DNA chain are represented by black beads and labelled as 1 and 21, respectively. *Bead 21* is fixed to the origin of coordinate, whereas *bead 1* is attached to a harmonic spring  $k_A$  whose end A is displaced at a constant velocity v.



**FIGURE 4.3.** *Calculation of the experimental loading rate,*  $k_A$ . (A) Typical forceextension curve of the system where the enthalpic elasticity regime of the dsDNA handles before the unfolding of the DNA G4 is highlighted in light blue. (B) Result of the linear regression applied to this region in order to calculate the loading rate in pN/nm (N=20 realizations, mean  $\pm$  s.d.: 0.4  $\pm$  0.1 pN/nm).

*Figure 4.4* shows a typical experimental force-extension curve (depicted in red) with the results obtained from two simulations at v=0.08 (in green and blue). Comparing qualitative behaviour of the curves and the values of the experimental and simulated unfolding force, a good performance of the theoretical model is confirmed.



FIGURE 4.4. Comparison of the force-extension curve between two simulations at v=0.08 and one experimental realization of the mechanical unfolding. The experimental curve (in red) shows superimposed the entropic elasticity regime of the double-stranded handles at low forces.

The next step was, then, expressing the velocity in real units or, what is the same, specifying the time scale. For that purpose, the mean value of the unfolding force,  $F_u$ , was considered as a function of the velocity, or equivalently, as a function of the pulling rate, r; i.e. the rate of variation of the applied pulling force. Being defined  $F_u$  from equation (4.1) and taking into account that the pulling experiment is performed with a harmonic spring as mentioned before—with one end fixed to the origin of coordinates  $(d(x_1(t))/t = 0)$  and the other end moved at constant velocity  $(d(x_A(t))/t = v)$ —, then:

$$r = \frac{dF(t)}{dt} = k_A \frac{d}{dt} \left( x_A(t) - x_1(t) \right) \approx k_A \cdot v \tag{4.2}$$

*Figure 4.5* shows the resulting  $F_u$  vs. r curve (represented by cross symbols). It nicely reproduces the non-linear increasing behaviour observed in dynamic force spectroscopy experiments at high velocities —well-fitted by the *Dudko-Hummer-Szabo* (DHS) model [360]— and the almost force independent regime at low velocities —well-predicted by the *Yoreo* model [361]—. Furthermore, two experimental force values are also included in the graph to show the correct prediction of the rupture-force kinetics by the designed mesoscopic model: (i) the unfolding force  $F_u = 20.4$  pN obtained from the DNA experiments carried out in this thesis (see *Fig.3.13 (A)* in Chapter 3) at a pulling velocity v = 11.8 nm/s ± 1.4 nm/s (red square) —calculated by applying a linear fitting to extension vs time data, *N*=20, near the rupture event; see *Fig. 4.6*— and (ii) the equilibrium force  $F_u = 2.5$  pN from the constant force experiments of Long et al. [341] (black arrow in the left axis).



**FIGURE 4.5.** *Mean value of the unfolding force,*  $\langle F_u \rangle$ *, as a function of the pulling rate, r, and fitting to the Dudko-Hummer-Szabo (DHS) and the Yoreo models.* Note that the three lowest values of the pulling rates are excluded from the DHS fittings because in this region some rebinding events are observed during the unfolding, and, then, the kinetic models are not valid. Two experimental force values are included in this figure:  $F_u = 20.4$  pN from the unfolding experiments carried out in this thesis (red square; see *Fig.3.13 (A)* in Chapter 3) and  $F_u = 2.5$  pN from the constant force experiments of Long et al. [341] (arrow in the left axis). The pulling rate is expressed in dimensionless units.



**FIGURE 4.6.** *Calculation of the experimental pulling rate.* **(A)** Typical extension-time curve of the system where the enthalpic elasticity regime of the dsDNA handles before the unfolding of the DNA G4 is highlighted in green. **(B)** Result of the linear regression applied to this region in order to calculate the pulling rate in nm/s (N=20 realizations, mean ± s.d.: 11.8 ± 1.4 nm/s).

From the experimental data obtained in this thesis ( $F_u = 20.4 \text{ pN}$ ,  $v = 11.8 \pm 1.4 \text{ nm/s}$  and  $k_A = 0.4 \pm 0.1 \text{ pN/nm}$ ), a pulling rate r = 0.0014 was obtained, giving rise to a time unit of 0.043 s [359].

Finally, after being well-characterized the time and force scales of the mesoscopic model, a simulated unfolding force distribution similar to the experimental one shown in Chapter 3 (*Fig. 3.13*) was obtained (*Fig.4.7*). The good agreement between the simulated and experimental distributions is again a clear feature of the suitability of the model to predict real force-induced unfolding experiments.

Therefore, the mesoscopic model proposed here is able to both reproduce unfolding forces in the same order of magnitude as in experimental studies —which are impossible with atomistic calculations [325]— and make predictions. In particular, it is possible to explore wide time scales and study the unfolding pathways by using loading rates up to five orders of magnitude lower than those allowed in microscopic simulations. Because of that, this theoretical model could be used as a tool for performing systematic studies on mechanical stability of different G4 conformations, either G4 geometries according to loop orientation (parallel vs antiparallel) or G4 tandems repeats [267], like those existing at the end of human telomeres.



**FIGURE.4.7.** *Comparison between simulated and experimental unfolding force distribution*. Blue histogram correspond to the simulated unfolding force distribution and the red line to the Gaussian curve drawn with the experimental values of the unfolding mean force and its standard deviation ( $20.4 \pm 6.9$  pN).

#### 4.2. Stabilization of DNA G-quadruplexes by ruthenium complexes

G-rich single-stranded DNA telomeric ends have been found to adopt tandems of G4s, inactive conformations that are no longer recognized, nor elongated, by telomerase [166, 167]. Because of the overexpression of this enzyme in above 85% of human cancer cells [128], the design of small molecules that favours G4 formation and stabilization have become a promising approach for cancer chemotherapy [179, 186, 362-368]. To date, a number of G4 stabilizing drugs have been synthesised and demonstrated telomerase inhibition, including organic heteroaromatic compounds [180, 187, 188, 192, 369-377] and metal complexes with aromatic ligands [378-386]. Basically, they consist of large planar aromatic rings, which are capable of stacking onto a G4 structure through  $\pi$ - $\pi$  interactions with its G-tetrads, and a positively charge substituent, to interact with the grooves, loops and negatively charged phosphate backbone of the G4. However, most of them present relatively weak selectivity for G4 structures over duplex ones, resulting in acute toxic and intolerable side effects on normal tissues [387]. In addition, although telomeric sequences can exhibit a wide range of topologies, conformational selectivity has hardly been considered in almost all of the compounds recently reported with a good G4 affinity and stabilization ability [388, 389]. Hence, improving G4 selectivity is an essential goal for these G4 drugs, not only to decrease the inherent cytotoxicity but also to lead to the development of disease-specific therapeutic treatments.

Ruthenium(II) complexes, especially dinuclear species, have been reported to be G4 probes [390-396] and stabilizers [382, 397-409], exhibiting a good antitumoral response. These transition metal compounds offer significant advantages over other G4-binding drugs because of their extraordinary versatility in shape, charge and size. In particular, the octahedral coordination geometry on the Ru<sup>II</sup> core can give rise to diverse three-dimensional structures with prominent DNA binding properties [397, 408].

Until now, several studies about the mechanisms of interaction between small molecules and DNA duplex structures have been reported by single-molecule force spectroscopy [410-415], including ruthenium compounds [416, 417]. With regards G4s, only Koirala et al. [418] have provided detailed insights into the mechanical, kinetic and thermodynamic properties of these structures when they interact with G4 stabilizers.

Here, a new dinuclear ruthenium complex has been synthesised with the aim of studying its influence on the stability of human telomeric G4 structures by OT at the single-molecule level.

#### 4.2.1. Materials and methods

#### 4.2.1.1. Synthesis of the ruthenium complex

The ruthenium dimer  $[(\text{Ru}(\eta^6:\kappa^{1}-(1-\text{methyl-3-phenylpropylamine}))_2(\text{tpphz})]\text{Cl}_4$  was prepared by the reaction of the tether complex  $[\text{RuCl}_2(\eta^6:\kappa^{1}-\text{C}_6\text{H}_5(\text{C}_4\text{H}_8)\text{NH}_2)]$  with the tpphz ligand (tetrapyrido[3,2-a:2',3'-c:3'',2''-h:2''',3'''-j]phenazine) in a mixture of water and methanol (F*ig. 4.8*). It was obtained as the chloride salt and exhibited good solubility in aqueous solution. Its structure was characterized by <sup>1</sup>H NMR spectroscopy and elemental analysis. Its stability under the buffer conditions used in the OT experiments –10 mM Tris-Cl (pH=7.8), 100 mM KCl 1 mM EDTA– was also checked by <sup>1</sup>H NMR spectroscopy over 24 hours. All this work was performed by the research group "*Metallodrugs in cancer research*" led by Dr. Ana M. Pizarro at IMDEA Nanoscience.



**FIGURE 4.8.** Synthesis of ruthenium(II) complex [( $Ru(\eta^6:\kappa^1-(1-methyl-3-phenylpropylamine)$ )<sub>2</sub>(tpphz)]Cl<sub>4</sub>. Scheme of the reaction between the tether complex [ $RuCl_2(\eta^6:\kappa^1-C_6H_5(C_4H_8)NH_2$ )] and the tpphz ligand (tetrapyrido[3,2-a:2',3'-c:3'',2''-h:2''',3'''-j]phenazine) in a mixture of water and methanol (MeOH/H<sub>2</sub>O). The tpphz ligand and the tether complex were synthesised according to literature [419, 420].

#### 4.2.1.2. Synthesis of human telomeric DNA G-quadruplexes

The molecular construction used here to test the efficiency of the ruthenium dimer (RuD) in the stabilization of DNA G4-forming sequence by OT is the same as described in *section 3.2.2* in Chapter 3.

#### 4.2.1.3. Synthesis of dsDNA molecules for control experiments

Double-stranded DNA molecules were also synthesised to test the affinity of the RuD with duplex structures by OT. The molecular construction consists of a 2656-bp dsDNA main molecule and two DNA handles, one of 230 bp modified with digoxigenins and the other of 17 nt biotinylated at its 3' end.

The main molecule was obtained by digestion of the pUC19 DNA vector (New England Biolabs, NEB; Ref. N3041S) with HindIII (NEB; High Fidelity Enzyme, Ref. R3104S) and BamHI (NEB; High Fidelity Enzyme, Ref. R3136S) restriction endonucleases at 37°C during 2 hours. After that, DNA was purified using the *Wizard*® *SV Gel and PCR Clean-Up System* (Promega, Ref. A9281). Its size was checked by 1% agarose gel electrophoresis.

Digoxigenin (DIG)-labelled DNA handles were generated by PCR amplification of the pUC19 fragment [196, 647] —which contains the multiple cloning site—, following the protocol described in [421]. PCR products were digested with BamHI (NEB; High Fidelity Enzyme, Ref. R3136S) and gel-purified to recover a 230-bp DIG-modified DNA handle.

The biotinylated DNA handle was also synthesised according to Ref. [421]: a 3' biotin-labelled DNA oligonucleotide with its 5' end complementary to the HindIII sequence (5'-AGCTATCAGGCTACAAA-3').

Finally, DNA handles were ligated using the T4 DNA ligase (Roche, Cat. No. 10 481 220 001) to their complementary positions at the DNA main molecule following the manufacturer's reaction conditions (16°C, overnight). The resulting

product was purified using the *Wizard*® *SV Gel and PCR Clean-Up System* (Promega, Ref. A9281).

#### 4.2.1.4. Optical-tweezers experiments

To test the effect of the synthesised RuD over the stability of G4 structures, single-molecule experiments were performed by optical tweezers. First, a single molecular construction containing a human telomeric G4-forming sequence is stretched and relaxed several times to check the correct formation of the G4 structure; i.e. the visualization of a rupture event by the force region  $20.4 \pm 6.9$  pN (mean  $\pm$  s.d.; see *section* 3.3.1.1 in Chapter 3). Then, maintaining the molecular construct at a constant force (usually 10 pN to favour the binding of the complex) the RuD is flowed during a few minutes. After that, the molecule is stretched again and both the unfolding force and the shape of curve are compared with the data in absence of the RuD (see graphical description in *Fig.* 4.9). Control experiments with dsDNA molecules were also carried out following the same methodology to check if this RuD presents affinity for duplexes.



**FIGURE 4.9.** *RuD single-molecule experiments by OT.* A human telomeric G4-forming sequence (A) is stretched and relaxed several times (B) to check the correct formation of the G4 structure. Then, maintaining the molecular construct at a constant force (usually 10 pN to favour the binding of the complex) the RuD (depicted as orange ovals) is flowed during 30 s approximately (C). After that, the molecule is stretched again and both the unfolding force and the shape of curve are compared with the data in absence of RuD (D).

#### 4.2.2. Preliminary results and discussion

The stabilization of telomeric G4 species by a new RuD has been started to study at the single-molecule level by OT in this thesis. Different RuD concentrations (from 2  $\mu$ M to 0.1  $\mu$ M) and incubation times (from 5 min to 30 s) —i.e. the time at which RuD is flowed inside the microfluidics chamber while the molecule is subjected to constant force to favour the binding— have been tested. Our preliminary results show that, after flowing the RuD, the rupture event due to the unfolding of the DNA G4 structure occurs at higher forces (see *Fig.4.9*), an effect that could be associated to the binding of the metal drug as reported in [418].



FIGURE 4.9. Comparison between the force-extension curves of a molecular construct containing a human telomeric G4-forming sequence before and after flowing the RuD. Before flowing the RuD inside the microfluidic chamber, a typical rupture force for a DNA G4 is obtained (mean  $\pm$  s.d: 20.4  $\pm$  6.9 pN; see section 3.3.1.1 in Chapter 3). After flowing the metal complex ([RuD] = 0.1  $\mu$ M, 3 min), the unfolding of the G4 happens at higher forces. This effect could be considered as a result of the RuD binding, but more experiments need to be done to confirm its G4 stabilization ability. In the legend, "w/o" and "w/" mean "without" and "with RuD", respectively. Measurements were performed using a miniTweezers set-up.

In addition to this effect, a longer entropic plateau is also observed. To confirm if it is due to the interaction of the RuD with the dsDNA handles, control experiments with dsDNA molecules have been performed. As shown in *Figure 4.10*, the same deformation of the curve at low forces is obtained, suggesting that the RuD can also present affinity for duplex structures through intercalation between DNA base pairs.



**FIGURE 4.10.** *Comparison between the force-extension curves of a dsDNA molecule before and after flowing the RuD.* After flowing the metal complex ([RuD] =  $0.5 \mu$ M, 3 min), the unfolding of the G4 happens at higher forces. This effect could be considered as a result of the RuD intercalation between the DNA base pairs. Measurements were performed using a miniTweezers set-up.

According to our preliminary results, the RuD synthesised by Pizarro's group seems to stabilise DNA G4 conformations, but also to interact with duplex structures. For future therapeutic purposes, it will be necessary to find the proper RuD concentration and incubation time in which its intercalation between DNA base pairs is avoided. Nevertheless, this research work lays the foundation for future projects about studying induced mechanical changes in DNA and RNA G4 structures by G4-selective compounds at the single-molecule level.

## Conclusions

TElomeric Repeat containing RNA (TERRA) molecules have been studied by single-molecule microscopic and force-spectroscopic techniques. The interest in these non-coding transcripts with subtelomere-derived sequence lies in their potential biological roles, such as telomeric heterochromatin formation and telomere protection, which are probably associated to their capacity to fold into multiple G-quadruplex blocks. Therefore, to better understand these RNA species in terms of structure and mechano-chemical stability, their analysis by atomic force microscopy (AFM) and optical tweezers (OT) has been herein performed.

According to the results obtained in this thesis, it can be concluded that:

- 1. In contrast to the expected scenario in which the TERRA sequence folds into a parallel G-quadruplex or a G-triplex, a remarkable conformational competition between unspecific self-interacting ssRNA species and these non-canonical conformations associated with the TERRA repeats, (GGGUUA)<sub>n</sub>, has been observed for the first time and at the singlemolecule level.
- **2.** This interference represents a stochastic blockage in the G-quadruplex folding pathways of up to 20% due to the condensation of TERRA single strands. The formation of a condensed ssRNA 'knot' instead of a G-quadruplex/triplex can largely facilitate its denaturation by specialized protein machinery; specifically, decondensing the ssRNA involves an average force barrier near 10 pN, whereas that of the G-triplex or the G-quadruplex is on average above 20 pN.

- **3.** This behaviour is unique to the RNA G-quadruplex species since we have not found an analogue behaviour in their DNA counterparts herein synthesised.
- **4.** Application of Jarzynski's equality reveals that the TERRA G-quadruplex free-energy per nucleotide is larger than that of the condensates. A good agreement has been found with earlier literature in the use of this analysis to DNA analogue molecules, in which equilibrium free energies were obtained by ensemble-averaged experiments, thus supporting the validity of our mechanical stability analysis on the TERRA molecules. In tune with the profile of the stretch-relaxation cycles, this analysis also confirms that our single-molecule force-induced unfolding protocol takes place away from equilibrium.
- **5.** G-quadruplex unfolding is a halfway problem in complexity between protein and duplex nucleic acid unfolding.
- **6.** With regards to AFM results, a strategy to analyse RNA Gquadruplex/triplex structures on mica surface at the single-molecule level has been determined. We believe our methodology is useful not only to cross-check G-quadruplex molecular preparations but also for applications in nanotechnology that make use of non-canonical conformations as building blocks over a surface under controlled conditions.
- 7. The force-induced unfolding experiments obtained for the human DNA telomeric sequence capable of forming a G-quadruplex has been also used to develop a mesoscopic model capable to describe, for the first time, both the mechanical and thermal stability of this non-canonical structure. This model allows not only to reproduce unfolding forces in the same order of magnitude as in experimental studies, but also to make predictions.

**8.** According to our preliminary results, the dinuclear ruthenium complex proposed here seems to stabilise DNA G-quadruplex conformations, but also to interact with duplex structures. For therapeutic purposes, further experiments need to be done to find non-DNA duplex intercalating conditions. Nonetheless, this research work lays the foundation for future projects about studying induced mechanical changes in DNA and RNA G-quadruplexes by G4-selective compounds at the single-molecule level.

## Conclusiones

En esta tesis se han estudiado moléculas de ARN con diversas repeticiones de la secuencia telómérica GGGUUA, comúnmente denominadas *TERRA* (del inglés, *TElomeric Repeat-containing RNA*), mediante la aplicación de técnicas de moléculas individuales, tanto de microscopia como de espectroscopia de fuerza. El interés en estos ARN no codificantes reside en sus potenciales funciones biológicas, tales como la formación de heterocromatina en los telómeros o la protección de estos últimos, las cuales probablemente estén asociadas con la capacidad de plegamiento de dichos ARNs en tándems de *G-quadruplexes*. Por tanto, para un mayor entendimiento de estas moléculas en términos de estructura y estabilidad mecano-química, se ha procedido a su análisis mediante microscopía de fuerza atómicas (AFM; del inglés, *Atomic Force Microscopy*) y pinzas ópticas (OT; del inglés, *Optical Tweezers*).

De acuerdo a los resultados obtenidos en esta tesis, se concluye que:

- 1. En contra del patrón de plegamiento esperado para una secuencia *TERRA*; esto es, bien en forma de *G-quadruplex* paralelo o como *G-triplex*, se ha observado, por primera vez y a nivel de molécula individual, una notable competencia conformacional entre dichas estructuras no canónicas y aquellas resultantes de la interacción inespecífica inherente a hebras sencillas de ARN.
- Esta interferencia se traduce en un bloqueo estocástico del plegamiento de las moléculas *TERRA* en forma de estructuras tipo *G-quadruplex/triplex* de hasta un 20%. La formación de un condensado de ARN, en lugar de un *G-quadruplex/triplex*, podría facilitar, en gran

medida, su desnaturalización mediante complejos proteicos específicos, ya que su barrera de fuerza se encuentra próxima a los 10 pN frente a los más de 20 pN del *G-triplex* o del *G-quadruplex*.

- Este comportamiento es inherente a la naturaleza ribonucleica de las moléculas estudiadas, pues no se ha encontrado el mismo patrón en sus análogos químicos de ADN.
- 4. Mediante la aplicación del teorema de *Jarzynski*, se ha obtenido que la energía libre por nucleótido de un *G-quadruplex* de *TERRA* es mayor que la de los condensados de ARN. Paralelamente, se ha encontrado un gran nivel de concordancia entre la energía libre del sistema análogo de ADN resultante de aplicar dicho teorema y los datos existentes en bibliografía procedentes de experimentos de espectroscopía visible. Esto respalda la validez de nuestro análisis de estabilidad mecánica en moléculas *TERRA*. Por último, este análisis termodinámico también confirma que nuestro protocolo de despliegue inducido por la aplicación de fuerza a una sola molécula tiene lugar fuera del equilibrio, lo cual se refleja experimentalmente en la histéresis de las curvas fuerza-extensión.
- **5.** La complejidad de la dinámica de apertura de las estructuras tipo *Gquadruplex* se encuentra en un punto intermedio entre la de las proteínas y la de los ácidos nucleicos de doble cadena.
- **6.** Con respecto a los resultados de AFM, se ha desarrollado una estrategia exitosa para comprobar la formación de estructuras tipo *G-quadruplex/triplex* en moléculas *TERRA*. Esta metodología podría ser útil también para aplicaciones en nanotecnología, en las que dichas estructuras serían usadas para la construcción de sistemas modulares bajo condiciones controladas.

- 7. A partir de los resultados obtenidos mediante pinzas ópticas para los *Gquadruplexes* de ADN, se ha podido desarrollar un modelo mesoscópico que describe, por primera vez, tanto la estabilidad mecánica como térmica de este tipo de estructuras. Dicho modelo no sólo reproduce las fuerzas de despliegue en el mismo orden de magnitud que los estudios experimentales, sino que también permite hacer predicciones.
- 8. Conforme a los resultados preliminares obtenidos, el complejo de rutenio propuesto en esta tesis parece estabilizar *G-quadruplexes* de ADN, aunque también se ha encontrado interacción con estructuras de doble cadena. Para su utilización con fines terapéuticos, es necesario llevar a cabo más experimentos para encontrar condiciones no intercalantes con ADN de doble cadena. No obstante, este trabajo ha servido para sentar las bases a futuros proyectos acerca de los cambios mecánicos inducidos por compuestos selectivos de *G-quadruplexes* en estructuras de ADN y ARN a nivel de una sola molécula.

## Appendix A Simulations of TERRA sequences by *mfold*

In order to predict the formation of stable hairpin-like structures in the DNA and RNA G4-forming sequences studied in Chapter 3, *mfold* simulations [345] were carried out.

For the TERRA reference sequence, UAUA(GGGUUA)<sub>5</sub>UAGU, only one stable structure of this type was found (*Fig. A.1*). It possesses two G-bubbles and G-U and A-U pairings in the stem. It extends over almost the entire ssRNA sequence but its free energy ( $\Delta G$ = –0.40 kcal/mol at 1 M NaCl) does not overtake that of the thermal fluctuations (-0.7  $k_BT$ , being  $k_B$  the Boltzmann constant and T the temperature in the absolute scale), which yields an equilibrium rupture force much lower than 1 pN, calculated as follows:

$$F = \frac{0.40 \left(\frac{kcal}{mol}\right) \cdot 10^3 \left(\frac{cal}{kcal}\right) \cdot 4.184 \left(\frac{J}{cal}\right) \cdot 1 \left(\frac{N \cdot m}{J}\right)}{6.022 \cdot 10^{23} (mol^{-1}) \cdot 38 (nt) \cdot 0.51 \left(\frac{nm}{nt}\right) \cdot 10^{-9} \left(\frac{m}{nm}\right)} \cdot 10^{12} \left(\frac{pN}{N}\right) \approx 0.14 \ pN$$

This makes this structure very improbable according to the rupture force distribution of *Figure 3.11 (A)* in Chapter 3. In addition, non-stable hairpin-like structures were obtained for its DNA analogous sequence, TATA(GGGTTA)<sub>5</sub>TAGT ( $\Delta G > 0$ ), which is a consequence of the higher conformational potential of RNA relative to ssDNA.

The RNA control systems with more human telomeric repeats as well as that with extra random sequence have also the potential to form hairpin-like structures (see *Fig. A.2* and *Fig. A.3*) with stabilities ranging from -1.50 to -2.60 kcal/mol, which similarly correspond to forces well below 1 pN and thus highly improbable.



FIGURE A.1. Stable hairpin-like structure of the TERRA reference sequence,  $UAUA(GGGUUA)_5UAGU$ , obtained by mfold simulations. It possesses two G-bubbles (highlighted with light blue dashed circles) and G-U and A-U pairings in the stem ( $\Delta G$ = – 0.40 kcal/mol at 1 M NaCl).



FIGURE A.2. Stable hairpin-like structure of the RNA control sequences with more telomeric repeats obtained by mfold simulations. (A) Secondary structure of the RNA control sequence UAUA(GGGUUA)<sub>6</sub>UAGU with two G-bubbles and G-U and A-U pairings in the stem ( $\Delta G$ = –1.50 kcal/mol at 1 M NaCl). (B) Secondary structure of the RNA control sequence UAUA(GGGUUA)<sub>7</sub>UAGU, similar to the previous one, but with three G-bubbles ( $\Delta G$ = –1.50 kcal/mol at 1 M NaCl). G-bubbles are highlighted with light blue dashed circles.



FIGURE A.3. Stable hairpin-like structures of the RNA control sequence with more extra random ribonucleotides, UAUA (GGGUUA)<sub>5</sub> UAGUCUUCGA, obtained by mfold simulations. Free energies at 1 M NaCl ranging from -2.60 kcal/mol (A) to -1.60 kcal/mol (B, C).

# Appendix B

# DNA G-quadruplex mesoscopic model: theory

The mesoscopic model proposed in Chapter 4 to simulate the force-induced unfolding of a DNA G4 consists of considering each nucleotide of the single strand and the two central cations as single beads. A schematic representation of a parallel G4 assembly is shown in *Figure B.1*, including the structural parameters considered in the calculations. In particular,  $l_i$  is the distance between two consecutive beads in the chain;  $d_{GG}$ , the distance between two guanines of the same plane;  $d_{IG}$ , the distance between each central cation and its neighbour guanines and  $\theta_i$ , the angle between three consecutive guanines.



**FIGURE B.1.** *Scheme of a parallel G4 structure with its main structural parameters.* Each nucleotide is represented by a single bead (not to scale). Green beads correspond to guanines; blue ones, to the nucleotides that form the loops; i.e. TTA, and yellow spheres to monovalent cations. Beads highlighted in red are those involved in the pulling simulations (see Chapter 4). For simplicity, the twist between successive G-tetrad planes is not represented.

The Hamiltonian of the system under study is composed by the following terms:

#### > A harmonic interaction between two consecutive beads (*U*<sub>str</sub>):

$$U_{str} = \sum_{i} \frac{1}{2} k_s (l_i - l_0)^2$$
(B.1)

where  $l_i = |r_{i+1} - r_i|$ ,  $r_i$  is the position vector of the *i*-th particle,  $l_0$  is the equilibrium separation between beads and  $k_s$  the elastic constant of the chain. We take  $l_0 = 0.65$  nm, which is, approximately, the distance between consecutive phosphor atoms of the backbone according to the crystal structure of the human telomeric parallel G4 [88].

A Morse potential between consecutive guanines (G) of the same Gquartet (U<sub>GG</sub>), to simulate the Hoogsten hydrogen bonds between guanines:

$$U_{GG} = \sum_{p=1}^{3} \sum_{g=1}^{4} D_g \left( e^{-\alpha_G \left( d_{GG_g} - r_0 \right)} - 1 \right)^2$$
(B.2)

where the sums go over the number of G-quartets that constitutes a parallel DNA G4 structure, p, and the number of guanines involved in each of these planes, g.  $d_{GG_g}$  is the distance between two consecutive guanines in a G-quartet;  $r_0$ , the equilibrium length of the side of each plane and  $D_G$ , the dissociation energy corresponding to the depth of the Morse potential well. To consider the fact that hydrogen bonds between guanines are stronger when all the bonds of the same G-tetrad are formed [350],  $D_G$  is defined to be dependent on the distances between the guanines of the same plane; i.e.  $d_{GG_g}$  and  $r_0$ , as follows:

$$D_G = D_0 \left\{ 1 + 2e^{-\delta \left( \sum_g d_{GG_g} - 4r_0 \right)} \right\}$$
(B.3)

where  $\delta$  sets the length scale for the decay of  $D_G$  and  $D_0$  is the true dissociation energy. Thus,  $D_G$  varies from  $3D_0$ , when the G4 is folded and the side of the plane is  $r_0$ , to  $D_0$  when any of the distances  $d_{GG_g}$  increases beyond  $\delta^{-1}$ .



FIGURE B.2. Potential between guanines of the same plane ( $U_{GG}$ ) and between one cation and its neighbour guanine ( $U_{IG}$ ). Red lines mark the threshold distances used to define the coordination numbers between guanines as well as between the central ions and the neighbour guanines. Black lines in (B) delimit the interval [a,b] where the Coulomb force attraction tends linearly to zero.

An interaction potential between each central cation (I) and its neighbour guanines (U<sub>IG</sub>):

$$U_{IG} = \sum_{c=1}^{2} \sum_{n=1}^{8} A/d_{cn}^{12} - Q_{cn}/d_{cn}$$
(B.4)

where the sums go over the number of cations, *c*, and neighbour guanines, *n*.  $A/d_{cn}^{12}$  is a repulsive term related to the excluded volume effect and  $Q_{cn}$ , to the strength of the effective attraction between the metal ion and the guanine. The value of constant *A* is selected so that the minimum of  $U_{IG}$  is located at  $\sqrt{2(r_0/2)^2 + (p_0/2)^2}$  ( $p_0$  is the distance between the centres of two consecutive planes). The Coulomb contribution  $Q_{cn}/d_{cn}$  tends linearly to zero in the interval (*a,b*), indicated in *Figure B.2 (B*) with two black vertical lines. A purely repulsive interaction between the two ions  $U_{I_1I_2} = Q_{I_1I_2}/d_{I_1I_2}$  is also included. We set  $Q_{cn} = 2.5Q_{I_1I_2}$ . It is worth noting that, even if the interactions between the two ions and between each ion and the guanine, have an electrostatic origin, their features are different. In fact, the interaction *metal ion-guanine* is due to the metal-ion coordination between the ion and the oxygen *O6* of the guanine. Therefore, it is more sensitive to their distance separation than if it were a pure electrostatic interaction [422]. For this reason, we introduce, as a typical procedure adopted in these cases, a cut-off distance for the interaction between the cation and the guanines.

A bending energy interaction between the three consecutive beads that belong to each side of the G4 stem (U<sub>ben</sub>):

$$U_{ben} = \sum_{i} \frac{k_b}{2} \left[ 1 - \cos(\theta_i - \theta_0) \right]$$
(B.5)

where  $k_b$  is a bending elastic constant,  $\theta_0 \approx 150^\circ$  due to the twist between planes (not represented in *Fig. B1*) and  $\theta_i$  is the angle between vectors  $l_{i+1}$ and  $l_i$ . This term accounts for the stacking interactions between consecutive guanines and confers stability to the G4 stem. For the beads of the loops the bending can be neglected.

A repulsive Lennard Jones interaction (V<sub>LJ</sub>), to simulate the excluded volume effect:

$$V_{LJ} = \sum_{i} \sum_{j>i} 4\epsilon \left[ \left( \frac{\sigma}{r_{ij}} \right)^{12} - \left( \frac{\sigma}{r_{ij}} \right)^{6} \right]$$
(B.6)

if  $r_{ij} < 1.122\sigma$  between all the beads of the G4.  $\varepsilon$  is the depth of the potential well;  $\sigma$ , the finite distance at which the potential is zero and  $r_{ij}$ , the distance between particles *i* and *j*.

The dynamics of the system is, then, obtained from the over-damped equations of motion of the j -th bead of the G4 (j = 1, ..., 21):

$$m_j \dot{\boldsymbol{r}}_j = -\nabla_{\boldsymbol{r}_j} U_{str} - \nabla_{\boldsymbol{r}_j} U_{GG} - \nabla_{\boldsymbol{r}_j} U_{IG} - \nabla_{\boldsymbol{r}_j} U_{ben} - \nabla_{\boldsymbol{r}_j} V_{LJ} + \sqrt{2k_b T m_j} \eta_j(t) \quad (B.7)$$

and for the *i* -th bead representing the cations (i = 1, 2):

$$m_i \dot{\boldsymbol{I}}_i = -\nabla_{\boldsymbol{I}_i} U_{IG} - \nabla_{\boldsymbol{I}_i} U_{I_1 I_2} + \sqrt{2k_b T m_j} \eta_j(t)$$
(B.8)

The last terms in equations (B.7) and (B.8) represent the thermal contribution as a Gaussian uncorrelated noise. The damping is taken implicit into the time units. The mass of the ion  $m_i$  and of the nucleotide  $m_j$  are taken as  $m_j = 3.85 m_i$ . Length is given in units of  $l_0 = 0.65$  nm. Energy and time units were derived from experimental data —as explained in *Section 4.1* in Chapter 4— to be expressed in real terms. The rest of the parameters, in dimensionless units, are:  $D_0 = 1$ ,  $k_s = 100$ ,  $k_b = 2$ ,  $\alpha_G = 10$  and  $\delta = 0.3$ .

Systems of equations (B.7) and (B.8) have been integrated with the stochastic *Euler* algorithm with  $dt = 10^{-4}$  and pulling simulations were conducted at a temperature lower than the melting until the G4 is unfolded.

The resulting unfolding force,  $F_u$ , as a function of the pulling rate, r (see *Fig. 4.5* in Chapter 4) was fitted to the *Dudko-Hummer-Szabo* (DHS) and *Yoreo* kinetic models, defined as follows:

DHS model: 
$$\langle F_{u,D} \rangle = \frac{G^+}{\nu x_u} \left\{ 1 - \left[ \frac{1}{\beta G^+} ln \frac{k_0 e^{\beta G^+ + \gamma}}{\beta x_u r} \right]^{\nu} \right\}$$
 (B.9)

where  $G^+$  is the barrier separating the folded and unfolded states of a G4;  $x_u$  is the position of the barrier;  $k_0$  is the unfolding rate constant in the absence of force;  $\gamma \approx 0.577$  is the Euler-Marchesoni constant;  $\beta$  is the reciprocal of the

thermodynamic temperature of a system,  $\beta = 1/k_BT$ , being  $k_B$  the Boltzmann constant; and  $\nu$  is a parameter that sets the shape of the free energy potential to a cusp potential (if  $\nu = 1/2$ ) or to a linear-cubic potential (if  $\nu = 2/3$ ).

Yoreo model: 
$$\langle F_{u,Y} \rangle = f_{eq} + e^{\left(\frac{k_u(f_{eq}k_BT)}{rx_u}\right)} E_1\left(\frac{k_u(f_{eq}k_BT)}{rx_u}\right)$$
 (B.10)

where:

-  $f_{eq}$  is the force at which the force dependent folding and unfolding rates are equal and depends on the elastic constant  $k_A$  of the pulling spring:

$$f_{eq} = \sqrt{2k_A G^+} \tag{B.11}$$

-  $E_1(z)$  is the exponential integral defined in equation (B.12):

$$E_1(z) = \int_z^\infty \frac{e^{-s}}{s} ds \tag{B.12}$$

which is approximated by equation (B.13):

$$e^{z}E_{1}(z) \approx \ln(1 + e^{-\gamma}/z) \tag{B.13}$$

-  $k_u(f)$  is the force dependent unfolding rate from the Bell-Evans-Ritchie model [423], defined as follows:

$$k_u(f) = k_0 exp(\beta(fx_u - 0.5k_A x_u^2))$$
(B.14)

## Appendix C List of publications

- <u>Gutiérrez, I.</u>; Garavís, M.; de Lorenzo, S.; Villasante, A.; González, C. and Arias-Gonzalez, J.R. Single-stranded condensation stochastically blocks Gquadruplex assembly in human telomeric RNA. 2017. Pre-print.
- Bergues-Pupo, A.E; <u>Gutiérrez, I.</u>; Arias-Gonzalez, J.R.; Falo, F. and Fiasconaro, A. *Mesoscopic model for DNA G-quadruplex unfolding*. 2017. Accepted for publication in Scientific Reports.
## **Bibliography**

- Fire, A., S. Xu, M.K. Montgomery, S.A. Kostas, S.E. Driver, and C.C. Mello, *Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans*. Nature, **1998**. 391(6669): p. 806-811.
- Montgomery, M.K., S. Xu, and A. Fire, RNA as a target of double-stranded RNA-mediated genetic interference in Caenorhabditis elegans. Proceedings of the National Academy of Sciences, 1998. 95(26): p. 15502-15507.
- 3. Timmons, L. and A. Fire, *Specific interference by ingested dsRNA*. Nature, **1998**. 395(6705): p. 854-854.
- 4. Um, S.H., J.B. Lee, N. Park, S.Y. Kwon, C.C. Umbach, and D. Luo, *Enzyme-catalysed assembly of DNA hydrogel.* Nature materials, **2006**. 5(10): p. 797-801.
- 5. Park, N., S.H. Um, H. Funabashi, J. Xu, and D. Luo, *A cell-free protein-producing gel.* Nature materials, **2009**. 8(5): p. 432-437.
- Chhabra, R., J. Sharma, Y. Liu, S. Rinker, and H. Yan, DNA self-assembly for nanomedicine. Adv. Drug Del. Rev., 2010. 62(6): p. 617-625.
- 7. Lee, J.B., M.J. Campolongo, J.S. Kahn, Y.H. Roh, M.R. Hartman, and D. Luo, *DNA-based nanostructures for molecular sensing.* Nanoscale, **2010**. 2(2): p. 188-197.
- 8. Seeman, N.C., *Nanomaterials based on DNA*. Annu. Rev. Biochem., **2010**. 79: p. 65-87.
- 9. Roh, Y.H., J.B. Lee, P. Kiatwuthinon, M.R. Hartman, J.J. Cha, S.H. Um, D.A. Muller, and D. Luo, *DNAsomes: multifunctional DNA-based nanocarriers.* Small, **2011**. 7(1): p. 74-78.
- Peng, S., T.L. Derrien, J. Cui, C. Xu, and D. Luo, *From cells to DNA materials*. Mater. Today, 2012. 15(5): p. 190-194.
- 11. Church, G.M., Y. Gao, and S. Kosuri, *Next-generation digital information storage in DNA*. Science, **2012**. 337(6102): p. 1628-1628.
- Goldman, N., P. Bertone, S. Chen, C. Dessimoz, E.M. LeProust, B. Sipos, and E. Birney, *Towards practical, high-capacity, low-maintenance information storage in synthesized DNA.* Nature, **2013**. 494(7435): p. 77-80.
- Guo, P., *The emerging field of RNA nanotechnology*. Nature nanotechnology, **2010**. 5(12): p. 833-842.
- 14. Avery, O.T., C.M. MacLeod, and M. McCarty, *Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a*

*desoxyribonucleic acid fraction isolated from pneumococcus type III.* The Journal of experimental medicine, **1944**. 79(2): p. 137.

- 15. Hershey, A.D. and M. Chase, *Independent functions of viral protein and nucleic acid in growth of bacteriophage.* The Journal of general physiology, **1952**. 36(1): p. 39-56.
- 16. Zamenhof, S., G. Brawerman, and E. Chargaff, *On the desoxypentose nucleic acids from several microorganisms.* Biochim. Biophys. Acta, **1952**. 9: p. 402-405.
- 17. Wilkins, M. and J. Randall, *Crystallinity in sperm heads: molecular structure of nucleoprotein in vivo.* Biochim. Biophys. Acta, **1953**. 10: p. 192-193.
- 18. Watson, J. and F. Crick, *A structure for ADN*. Nature, **1953**. 171(4356): p. 737-8.
- 19. Wang, J.C., *Helical repeat of DNA in solution*. Proceedings of the National Academy of Sciences, **1979**. 76(1): p. 200-203.
- 20. Franklin, R.E. and R.G. Gosling, *The structure of sodium thymonucleate fibres. I. The influence of water content.* Acta Crystallogr., **1953**. 6(8-9): p. 673-677.
- 21. Franklin, R.E. and R.G. Gosling, *The structure of sodium thymonucleate fibres. II. The cylindrically symmetrical Patterson function.* Acta Crystallogr., **1953**. 6(8-9): p. 678-685.
- 22. Franklin, R.E. and R.G. Gosling, *Molecular configuration in sodium thymonucleate*. Nature, **1953**. 171(4356): p. 740-741.
- Wahl, M.C. and M. Sundaralingam, *Crystal structures of A-DNA duplexes*. Biopolymers, **1997**.
  44(1): p. 45-63.
- 24. Nekludova, L. and C.O. Pabo, *Distinctive DNA conformation with enlarged major groove is found in Zn-finger-DNA and other protein-DNA complexes.* Proceedings of the National Academy of Sciences, **1994**. 91(15): p. 6948-6952.
- 25. Timsit, Y., DNA structure and polymerase fidelity. J. Mol. Biol., **1999**. 293(4): p. 835-853.
- 26. Lu, X.-J., Z. Shakked, and W.K. Olson, *A-form conformational motifs in ligand-bound DNA structures.* J. Mol. Biol., **2000**. 300(4): p. 819-840.
- Hormeño, S., F. Moreno-Herrero, B. Ibarra, J.L. Carrascosa, J.M. Valpuesta, and J.R. Arias-Gonzalez, *Condensation prevails over BA transition in the structure of DNA at low humidity*. Biophys. J., **2011**. 100(8): p. 2006-2015.
- Hormeño, S., B. Ibarra, J.L. Carrascosa, J.M. Valpuesta, F. Moreno-Herrero, and J.R. Arias-Gonzalez, *Mechanical Properties of High-G-C Content DNA with A-Type Base-Stacking*. Biophys. J., **2011**. 100(8): p. 1996-2005.
- 29. Hormeño, S., B. Ibarra, J.M. Valpuesta, J.L. Carrascosa, and J. Ricardo Arias-Gonzalez, *Mechanical stability of low-humidity single DNA molecules.* Biopolymers, **2012**. 97(4): p. 199-208.
- 30. Arias-Gonzalez, J.R., *Single-molecule portrait of DNA and RNA double helices*. Integrative Biology, **2014**. 6(10): p. 904-925.
- Wang, A., G.J. Quigley, F.J. Kolpak, J.L. Crawford, J.H. Van Boom, G. van der Marel, and A. Rich, *Molecular structure of a left-handed double helical DNA fragment at atomic resolution*. Nature, **1979**. 282(5740): p. 680.

- Hardwick, J.S., D. Ptchelkine, A.H. El-Sagheer, I. Tear, D. Singleton, S.E.V. Phillips, A.N. Lane, and T. Brown, *5-Formylcytosine does not change the global structure of DNA*. Nat. Struct. Mol. Biol., **2017**. advance online publication.
- 33. Drew, H.R., R.M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura, and R.E. Dickerson, *Structure of a B-DNA dodecamer: conformation and dynamics*. Proceedings of the National Academy of Sciences, **1981**. 78(4): p. 2179-2183.
- 34. https://www.liverpool.ac.uk/~bates/Z-DNAframe.html.
- 35. Bikard, D., C. Loot, Z. Baharoglu, and D. Mazel, *Folded DNA in action: hairpin formation and biological functions in prokaryotes.* Microbiol. Mol. Biol. Rev., **2010**. 74(4): p. 570-588.
- Brázda, V., R.C. Laister, E.B. Jagelská, and C. Arrowsmith, *Cruciform structures are a common DNA feature important for regulating biological processes.* BMC Mol. Biol., **2011**. 12(1): p.
   33.
- Wells, R., D. Collier, J. Hanvey, M. Shimizu, and F. Wohlrab, *The chemistry and biology of unusual DNA structures adopted by oligopurine. oligopyrimidine sequences.* The FASEB Journal, **1988**. 2(14): p. 2939-2949.
- 38. Htun, H. and J.E. Dahlberg, *Topology and formation of triple-stranded H-DNA*. Science, **1989**.
  243(4898): p. 1571-1577.
- Frank-Kamenetskii, M.D. and S.M. Mirkin, *Triplex DNA structures*. Annu. Rev. Biochem., 1995. 64(1): p. 65-95.
- 40. Soyfer, V.N. and V.N. Potaman, *Triplex Recognition*, in *Triple-helical Nucleic Acids*. 1996, Springer. p. 151-193.
- 41. Gehring, K., J.-L. Leroy, and M. GuÃ, *A tetrameric DNA structure with protonated cytosine. cytosine base pairs.* Nature, **1993**. 363(6429): p. 561.
- 42. Leroy, J.L., K. Gehring, A. Kettani, and M. Gueron, *Acid multimers of oligodeoxycytidine strands: stoichiometry, base-pair characterization, and proton exchange properties.* Biochemistry, **1993**. 32(23): p. 6019-6031.
- 43. Nelson, D.L., A.L. Lehninger, and M.M. Cox, *Lehninger principles of biochemistry*. 2008: Macmillan.
- 44. Batey, R.T., R.P. Rambo, and J.A. Doudna, *Tertiary motifs in RNA structure and folding.* Angew. Chem. Int. Ed., **1999**. 38(16): p. 2326-2343.
- 45. Moore, P.B., *Structural motifs in RNA*. Annu. Rev. Biochem., **1999**. 68(1): p. 287-300.
- 46. Leontis, N.B., A. Lescoute, and E. Westhof, *The building blocks and motifs of RNA architecture*. Curr. Opin. Struct. Biol., **2006**. 16(3): p. 279-287.
- 47. Holbrook, S.R., *Structural principles from large RNAs.* Annu. Rev. Biophys., **2008**. 37: p. 445-464.
- 48. Rich, A. and D.R. Davies, *A new two stranded helical structure: polyadenylic acid and polyuridylic acid.* J. Am. Chem. Soc., **1956**. 78(14): p. 3548-3549.

- 49. Rich, A., A hybrid helix containing both deoxyribose and ribose polynucleotides and its relation to the transfer of information between the nucleic acids. Proceedings of the National Academy of Sciences, **1960**. 46(8): p. 1044-1053.
- 50. Herrero-Galán, E., M.E. Fuentes-Perez, C. Carrasco, J.M. Valpuesta, J.L. Carrascosa, F. Moreno-Herrero, and J.R. Arias-Gonzalez, *Mechanical identities of RNA and DNA double helices unveiled at the single-molecule level.* J. Am. Chem. Soc., **2012**. 135(1): p. 122-131.
- 51. Devi, G., Y. Zhou, Z. Zhong, D.F.K. Toh, and G. Chen, *RNA triplexes: from structural principles to biological and biotech applications.* Wiley Interdisciplinary Reviews: RNA, **2015**. 6(1): p. 111-128.
- 52. Theimer, C.A., C.A. Blois, and J. Feigon, *Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function.* Mol. Cell, **2005**. 17(5): p. 671-682.
- 53. Shefer, K., Y. Brown, V. Gorkovoy, T. Nussbaum, N.B. Ulyanov, and Y. Tzfati, *A triple helix within a pseudoknot is a conserved and essential element of telomerase RNA.* Mol. Cell. Biol., **2007**. 27(6): p. 2130-2143.
- 54. Ulyanov, N.B., K. Shefer, T.L. James, and Y. Tzfati, *Pseudoknot structures with conserved base triples in telomerase RNAs of ciliates.* Nucleic Acids Res., **2007**. 35(18): p. 6150-6160.
- 55. Kim, N.-K., Q. Zhang, J. Zhou, C.A. Theimer, R.D. Peterson, and J. Feigon, *Solution structure and dynamics of the wild-type pseudoknot of human telomerase RNA.* J. Mol. Biol., **2008**. 384(5): p. 1249-1261.
- 56. Qiao, F. and T.R. Cech, *Triple-helix structure in telomerase RNA contributes to catalysis.* Nat. Struct. Mol. Biol., **2008**. 15(6): p. 634-640.
- 57. Cash, D.D., O. Cohen-Zontag, N.-K. Kim, K. Shefer, Y. Brown, N.B. Ulyanov, Y. Tzfati, and J. Feigon, *Pyrimidine motif triple helix in the Kluyveromyces lactis telomerase RNA pseudoknot is essential for function in vivo.* Proceedings of the National Academy of Sciences, **2013**. 110(27): p. 10970-10975.
- 58. Mitton-Fry, R.M., S.J. DeGregorio, J. Wang, T.A. Steitz, and J.A. Steitz, *Poly (A) tail recognition by a viral RNA element through assembly of a triple helix.* Science, **2010**. 330(6008): p. 1244-1247.
- 59. Tycowski, K.T., M.-D. Shu, S. Borah, M. Shi, and J.A. Steitz, *Conservation of a triple-helixforming RNA stability element in noncoding and genomic RNAs of diverse viruses.* Cell reports, **2012**. 2(1): p. 26-32.
- 60. Wilusz, J.E., C.K. JnBaptiste, L.Y. Lu, C.-D. Kuhn, L. Joshua-Tor, and P.A. Sharp, *A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly (A) tails.* Genes Dev., **2012**. 26(21): p. 2392-2407.
- 61. Appasamy, S.D., E.I. Ramlan, and M. Firdaus-Raih, *Comparative sequence and structure analysis reveals the conservation and diversity of nucleotide positions and their associated tertiary interactions in the riboswitches.* PLoS One, **2013**. 8(9): p. e73984.

- 62. Liberman, J.A., M. Salim, J. Krucinska, and J.E. Wedekind, *Structure of a class II preQ1 riboswitch reveals ligand recognition by a new fold.* Nat. Chem. Biol., **2013**. 9(6): p. 353-355.
- Batey, R.T., *Structure and mechanism of purine-binding riboswitches.* Q. Rev. Biophys., **2012**.
  45(3): p. 345-381.
- 64. Buske, F.A., J.S. Mattick, and T.L. Bailey, *Potential in vivo roles of nucleic acid triple-helices*. RNA Biol., **2011**. 8(3): p. 427-439.
- 65. Butcher, S.E. and A.M. Pyle, *The molecular interactions that stabilize RNA tertiary structure: RNA motifs, patterns, and networks.* Acc. Chem. Res., **2011**. 44(12): p. 1302-1311.
- 66. Conrad, N.K., *The emerging role of triple helices in RNA biology.* Wiley Interdisciplinary Reviews: RNA, **2014**. 5(1): p. 15-29.
- 67. Kumari, S., A. Bugaut, J.L. Huppert, and S. Balasubramanian, *An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation.* Nat. Chem. Biol., **2007**. 3(4): p. 218-221.
- 68. Kumari, S., A. Bugaut, and S. Balasubramanian, *Position and stability are determining factors* for translation repression by an RNA G-quadruplex-forming sequence within the 5' UTR of the NRAS proto-oncogene. Biochemistry, **2008**. 47(48): p. 12664.
- 69. Beaudoin, J.-D. and J.-P. Perreault, *5'-UTR G-quadruplex structures acting as translational repressors.* Nucleic Acids Res., **2010**. 38(20): p. 7022-7036.
- Xu, Y., Y. Suzuki, K. Ito, and M. Komiyama, *Telomeric repeat-containing RNA structure in living cells.* Proceedings of the National Academy of Sciences, **2010**. 107(33): p. 14579-14584.
- 71. Bugaut, A. and S. Balasubramanian, *5'-UTR RNA G-quadruplexes: translation regulation and targeting.* Nucleic Acids Res., **2012**. 40(11): p. 4727-4741.
- Beaudoin, J.-D. and J.-P. Perreault, *Exploring mRNA 3'-UTR G-quadruplexes: evidence of roles in both alternative polyadenylation and mRNA shortening.* Nucleic Acids Res., 2013. 41(11): p. 5898-5911.
- 73. Biffi, G., M. Di Antonio, D. Tannahill, and S. Balasubramanian, *Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells.* Nat. Chem., **2014**. 6(1): p. 75-80.
- 74. Dykxhoorn, D.M., C.D. Novina, and P.A. Sharp, *Killing the messenger: short RNAs that silence gene expression.* Nature reviews Molecular cell biology, **2003**. 4(6): p. 457-467.
- Lu, J., G. Getz, E.A. Miska, E. Alvarez-Saavedra, J. Lamb, D. Peck, A. Sweet-Cordero, B.L. Ebert, R.H. Mak, and A.A. Ferrando, *MicroRNA expression profiles classify human cancers*. Nature, 2005. 435(7043): p. 834-838.
- 76. Bumcrot, D., M. Manoharan, V. Koteliansky, and D.W. Sah, *RNAi therapeutics: a potential new class of pharmaceutical drugs.* Nat. Chem. Biol., **2006**. 2(12): p. 711-719.
- 77. Calin, G.A. and C.M. Croce, *MicroRNA signatures in human cancers*. Nature reviews cancer, **2006**. 6(11): p. 857-866.

- 78. Petersen, C.P., M.-E. Bordeleau, J. Pelletier, and P.A. Sharp, *Short RNAs repress translation after initiation in mammalian cells.* Mol. Cell, **2006**. 21(4): p. 533-542.
- 79. Ventura, A. and T. Jacks, *MicroRNAs and cancer: short RNAs go a long way.* Cell, **2009**. 136(4): p. 586-591.
- 80. Siomi, M.C., K. Sato, D. Pezic, and A.A. Aravin, *PIWI-interacting small RNAs: the vanguard of genome defence.* Nature reviews Molecular cell biology, **2011**. 12(4): p. 246-258.
- 81. Gellert, M., M.N. Lipsett, and D.R. Davies, *Helix formation by guanylic acid.* Proceedings of the National Academy of Sciences, **1962**. 48(12): p. 2013-2018.
- 82. Lipps, H.J., W. Gruissem, and D.M. Prescott, *Higher order DNA structure in macronuclear chromatin of the hypotrichous ciliate Oxytricha nova.* Proceedings of the National Academy of Sciences, **1982**. 79(8): p. 2495-2499.
- 83. Sundquist, W.I. and A. Klug, *Telomeric DNA dimerizes by formation of guanine tetrads* between hairpin loops. **1989**.
- Lipps, H.J. and D. Rhodes, *G-quadruplex structures: in vivo evidence and function*. Trends Cell Biol., **2009**. 19(8): p. 414-422.
- 85. Bochman, M.L., K. Paeschke, and V.A. Zakian, *DNA secondary structures: stability and function of G-quadruplex structures.* Nature Reviews Genetics, **2012**. 13(11): p. 770-780.
- 86. Rhodes, D. and H.J. Lipps, *G-quadruplexes and their regulatory roles in biology*. Nucleic Acids Res., **2015**: p. gkv862.
- Hänsel-Hertsch, R., M. Di Antonio, and S. Balasubramanian, DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential. Nature Reviews Molecular Cell Biology, 2017. 18(5): p. 279-284.
- 88. Parkinson, G.N., M.P. Lee, and S. Neidle, *Crystal structure of parallel quadruplexes from human telomeric DNA*. Nature, **2002**. 417(6891): p. 876-880.
- 89. Schultze, P., N.V. Hud, F.W. Smith, and J. Feigon, *The effect of sodium, potassium and ammonium ions on the conformation of the dimeric quadruplex formed by the Oxytricha nova telomere repeat oligonucleotide d (G4T4G4).* Nucleic Acids Res., **1999**. 27(15): p. 3018-3028.
- 90. Cevec, M. and J. Plavec, *Role of loop residues and cations on the formation and stability of dimeric DNA G-quadruplexes.* Biochemistry, **2005**. 44(46): p. 15238-15246.
- 91. Clay, E.H. and I.R. Gould, *A combined QM and MM investigation into guanine quadruplexes.* J.
   Mol. Graphics Model., 2005. 24(2): p. 138-146.
- 92. Dingley, A.J., R.D. Peterson, S. Grzesiek, and J. Feigon, *Characterization of the cation and temperature dependence of DNA quadruplex hydrogen bond properties using high-resolution NMR*. J. Am. Chem. Soc., **2005**. 127(41): p. 14466-14472.
- 93. van Mourik, T. and A.J. Dingley, Characterization of the Monovalent Ion Position and Hydrogen-Bond Network in Guanine Quartets by DFT Calculations of NMR Parameters. Chemistry-A European Journal, 2005. 11(20): p. 6064-6079.

- 94. Burge, S., G.N. Parkinson, P. Hazel, A.K. Todd, and S. Neidle, *Quadruplex DNA: sequence, topology and structure.* Nucleic Acids Res., **2006**. 34(19): p. 5402-5415.
- 95. Neidle, S., *4 Nonstandard and Higher-Order DNA Structures: DNA–DNA Recognition*, in *Principles of Nucleic Acid Structure*. 2008, Academic Press: New York. p. 81-131.
- 96. Aboul-Ela, F., A.I. Murchie, and D.M. Lilley, *NMR study of parallel-stranded tetraplex formation by the hexadeoxynucleotide d (TG4T).* Nature, **1992**. 360(6401): p. 280-282.
- 97. Wang, Y. and D.J. Patel, *Guanine residues in d (T2AG3) and d (T2G4) form parallel-stranded potassium cation stabilized G-quadruplexes with anti glycosidic torsion angles in solution.* Biochemistry, **1992**. 31(35): p. 8112-8119.
- 98. Wang, Y. and D.J. Patel, *Solution structure of the human telomeric repeat d [AG3 (T2AG3) 3] G-tetraplex.* Structure, **1993**. 1(4): p. 263-282.
- 99. Phillips, K., Z. Dauter, A.I. Murchie, D.M. Lilley, and B. Luisi, *The crystal structure of a parallel-stranded guanine tetraplex at 0.95 Å resolution.* J. Mol. Biol., **1997**. 273(1): p. 171-182.
- 100. Horvath, M.P. and S.C. Schultz, *DNA G-quartets in a 1.86 Å resolution structure of an Oxytricha nova telomeric protein-DNA complex.* J. Mol. Biol., **2001**. 310(2): p. 367-377.
- 101. Haider, S., G.N. Parkinson, and S. Neidle, *Crystal structure of the potassium form of an Oxytricha nova G-quadruplex*. J. Mol. Biol., **2002**. 320(2): p. 189-200.
- 102. Črnugelj, M., P. Šket, and J. Plavec, Small change in a G-rich sequence, a dramatic change in topology: new dimeric G-quadruplex folding motif with unique loop orientations. J. Am. Chem. Soc., 2003. 125(26): p. 7866-7871.
- 103. Hazel, P., J. Huppert, S. Balasubramanian, and S. Neidle, *Loop-length-dependent folding of G-quadruplexes*. J. Am. Chem. Soc., **2004**. 126(50): p. 16405-16415.
- 104. Luu, K.N., A.T. Phan, V. Kuryavyi, L. Lacroix, and D.J. Patel, *Structure of the human telomere in K+ solution: an intramolecular (3+ 1) G-quadruplex scaffold.* J. Am. Chem. Soc., **2006**. 128(30): p. 9963-9970.
- 105. Rachwal, P.A., I.S. Findlow, J.M. Werner, T. Brown, and K.R. Fox, *Intramolecular DNA quadruplexes with different arrangements of short and long loops.* Nucleic Acids Res., 2007. 35(12): p. 4214-4222.
- 106. Rachwal, P.A., T. Brown, and K.R. Fox, *Sequence effects of single base loops in intramolecular quadruplex DNA*. FEBS Lett., **2007**. 581(8): p. 1657-1660.
- 107. Phan, A.T., V. Kuryavyi, S. Burge, S. Neidle, and D.J. Patel, *Structure of an unprecedented G-quadruplex scaffold in the human c-kit promoter.* J. Am. Chem. Soc., **2007**. 129(14): p. 4386-4392.
- 108. Karsisiotis, A.I., N.M.a. Hessari, E. Novellino, G.P. Spada, A. Randazzo, and M. Webba da Silva, *Topological Characterization of Nucleic Acid G-Quadruplexes by UV Absorption and Circular Dichroism.* Angew. Chem. Int. Ed., **2011**. 50(45): p. 10645-10648.

- 109. Cang, X., J. Šponer, and T.E. Cheatham, *Explaining the varied glycosidic conformational, G-tract length and sequence preferences for anti-parallel G-quadruplexes.* Nucleic Acids Res., 2011. 39(10): p. 4499-4512.
- Rawal, P., V.B.R. Kummarasetti, J. Ravindran, N. Kumar, K. Halder, R. Sharma, M. Mukerji,
   S.K. Das, and S. Chowdhury, *Genome-wide prediction of G4 DNA as regulatory motifs: role in Escherichia coli global regulation.* Genome Res., 2006. 16(5): p. 644-655.
- 111. Beaume, N., R. Pathak, V.K. Yadav, S. Kota, H.S. Misra, H.K. Gautam, and S. Chowdhury, Genome-wide study predicts promoter-G4 DNA motifs regulate selective functions in bacteria: radioresistance of D. radiodurans involves G4 DNA-mediated regulation. Nucleic Acids Res., 2013. 41(1): p. 76-89.
- 112. Johnson, J.E., J.S. Smith, M.L. Kozak, and F.B. Johnson, *In vivo veritas: using yeast to probe the biological functions of G-quadruplexes.* Biochimie, **2008**. 90(8): p. 1250-1263.
- 113. Huppert, J., *Hunting G-quadruplexes*. Biochimie, **2008**. 90(8): p. 1140-1148.
- Maizels, N., *Dynamic roles for G4 DNA in the biology of eukaryotic cells*. Nat. Struct. Mol. Biol.,
  2006. 13(12): p. 1055-1059.
- 115. Sundquist, W.I. and S. Heaphy, *Evidence for interstrand quadruplex formation in the dimerization of human immunodeficiency virus 1 genomic RNA.* Proceedings of the National Academy of Sciences, **1993**. 90(8): p. 3393-3397.
- 116. Norseen, J., F.B. Johnson, and P.M. Lieberman, *Role for G-quadruplex RNA binding by Epstein-Barr virus nuclear antigen 1 in DNA replication and metaphase chromosome attachment.* J. Virol., **2009**. 83(20): p. 10336-10346.
- 117. Tuesuwan, B., J.T. Kern, P.W. Thomas, M. Rodriguez, J. Li, W.M. David, and S.M. Kerwin, Simian virus 40 large T-antigen G-quadruplex DNA helicase inhibition by G-quadruplex DNAinteractive agents. Biochemistry, **2008**. 47(7): p. 1896-1909.
- 118. Métifiot, M., S. Amrane, S. Litvak, and M.-L. Andreola, *G*-quadruplexes in viruses: function and potential therapeutic applications. Nucleic Acids Res., **2014**. 42(20): p. 12352-12366.
- 119. Capra, J.A., K. Paeschke, M. Singh, and V.A. Zakian, *G-quadruplex DNA sequences are* evolutionarily conserved and associated with distinct genomic features in Saccharomyces cerevisiae. PLoS Comput. Biol., **2010**. 6(7): p. e1000861.
- 120. Collie, G.W., G.N. Parkinson, S. Neidle, F. Rosu, E. De Pauw, and V. Gabelica, *Electrospray* mass spectrometry of telomeric RNA (TERRA) reveals the formation of stable multimeric *G*-quadruplex structures. J. Am. Chem. Soc., **2010**. 132(27): p. 9328-9334.
- 121. Martadinata, H., B. Heddi, K.W. Lim, and A.T.n. Phan, *Structure of long human telomeric RNA* (*TERRA*): *G-quadruplexes formed by four and eight UUAGGG repeats are stable building blocks.* Biochemistry, **2011**. 50(29): p. 6455-6461.
- 122. de Lange, T., *Protection of mammalian telomeres.* Oncogene, **2002**. 21(4): p. 532.
- 123. Moyzis, R.K., J.M. Buckingham, L.S. Cram, M. Dani, L.L. Deaven, M.D. Jones, J. Meyne, R.L. Ratliff, and J.-R. Wu, *A highly conserved repetitive DNA sequence,(TTAGGG) n, present at the*

*telomeres of human chromosomes.* Proceedings of the National Academy of Sciences, **1988**. 85(18): p. 6622-6626.

- 124. De Lange, T., *Shelterin: the protein complex that shapes and safeguards human telomeres.* Genes Dev., **2005**. 19(18): p. 2100-2110.
- 125. Liu, D., M.S. O'Connor, J. Qin, and Z. Songyang, *Telosome, a mammalian telomere-associated complex formed by multiple telomeric proteins.* J. Biol. Chem., **2004**. 279(49): p. 51338-51342.
- 126. Blackburn, E.H., *Switching and signaling at the telomere*. Cell, **2001**. 106(6): p. 661-673.
- 127. Collins, K. and J.R. Mitchell, *Telomerase in the human organism.* Oncogene, **2002**. 21(4): p. 564.
- 128. Kim, N.W., M.A. Piatyszek, K.R. Prowse, and C.B. Harley, *Specific association of human telomerase activity with immortal cells and cancer.* Science, **1994**. 266(5193): p. 2011.
- 129. Meyerson, M., *Role of telomerase in normal and cancer cells.* J. Clin. Oncol., **2000**. 18(13): p. 2626-2634.
- Hemann, M.T., M.A. Strong, L.-Y. Hao, and C.W. Greider, *The shortest telomere, not average telomere length, is critical for cell viability and chromosome stability.* Cell, **2001**. 107(1): p. 67-77.
- 131. Blasco, M.A., *Telomeres and human disease: ageing, cancer and beyond.* Nature Reviews Genetics, **2005**. 6(8): p. 611-622.
- 132. Collado, M., M.A. Blasco, and M. Serrano, *Cellular senescence in cancer and aging.* Cell, 2007. 130(2): p. 223-233.
- 133. Shay, J.W. and W.E. Wright. *Role of telomeres and telomerase in cancer*. in *Semin. Cancer Biol.*2011. Elsevier.
- 134. Hahn, W.C., S.A. Stewart, M.W. Brooks, S.G. York, E. Eaton, A. Kurachi, R.L. Beijersbergen, J.H. Knoll, M. Meyerson, and R.A. Weinberg, *Inhibition of telomerase limits the growth of human cancer cells.* Nat. Med., **1999**. 5(10): p. 1164-1170.
- 135. Oeseburg, H., R.A. de Boer, W.H. van Gilst, and P. van der Harst, *Telomere biology in healthy aging and disease.* Pflügers Archiv-European Journal of Physiology, **2010**. 459(2): p. 259-268.
- 136. Griffith, J.D., L. Comeau, S. Rosenfield, R.M. Stansel, A. Bianchi, H. Moss, and T. De Lange, *Mammalian telomeres end in a large duplex loop.* Cell, **1999**. 97(4): p. 503-514.
- De Lange, T., *T-loops and the origin of telomeres*. Nature Reviews Molecular Cell Biology, 2004. 5(4): p. 323-329.
- Paeschke, K., T. Simonsson, J. Postberg, D. Rhodes, and H.J. Lipps, *Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo.* Nat. Struct. Mol. Biol., 2005. 12(10): p. 847-854.
- 139. Tang, J., Z.-y. Kan, Y. Yao, Q. Wang, Y.-h. Hao, and Z. Tan, *G-quadruplex preferentially forms at the very 3' end of vertebrate telomeric DNA.* Nucleic Acids Res., **2008**. 36(4): p. 1200-1208.

- 140. Biffi, G., D. Tannahill, J. McCafferty, and S. Balasubramanian, *Quantitative visualization of DNA G-quadruplex structures in human cells.* Nat. Chem., **2013**. 5(3): p. 182-186.
- 141. Yang, Q., J. Xiang, S. Yang, Q. Zhou, Q. Li, Y. Tang, and G. Xu, *Verification of specific G-quadruplex structure by using a novel cyanine dye supramolecular assembly: I. Recognizing mixed G-quadruplex in human telomeres.* Chem. Commun., **2009**(9): p. 1103-1105.
- 142. Azzalin, C.M., P. Reichenbach, L. Khoriauli, E. Giulotto, and J. Lingner, *Telomeric repeat-containing RNA and RNA surveillance factors at mammalian chromosome ends.* Science, 2007. 318(5851): p. 798-801.
- 143. Schoeftner, S. and M.A. Blasco, *Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II.* Nat. Cell Biol., **2008**. 10(2): p. 228-236.
- 144. Azzalin, C.M. and J. Lingner, *Telomeres: the silence is broken.* Cell Cycle, **2008**. 7(9): p. 1161-1165.
- 145. Horard, B. and E. Gilson, *Telomeric RNA enters the game*. Nat. Cell Biol., **2008**. 10(2): p. 113-114.
- 146. Luke, B., A. Panza, S. Redon, N. Iglesias, Z. Li, and J. Lingner, *The Rat1p 5' to 3' exonuclease degrades telomeric repeat-containing RNA and promotes telomere elongation in Saccharomyces cerevisiae.* Mol. Cell, **2008**. 32(4): p. 465-477.
- 147. Caslini, C., J.A. Connelly, A. Serna, D. Broccoli, and J.L. Hess, *MLL associates with telomeres and regulates telomeric repeat-containing RNA transcription.* Mol. Cell. Biol., 2009. 29(16): p. 4519-4526.
- Deng, Z., J. Norseen, A. Wiedmer, H. Riethman, and P.M. Lieberman, *TERRA RNA binding to TRF2 facilitates heterochromatin formation and ORC recruitment at telomeres.* Mol. Cell, 2009. 35(4): p. 403-413.
- 149. Luke, B. and J. Lingner, *TERRA: telomeric repeat-containing RNA*. The EMBO journal, **2009**. 28(17): p. 2503-2510.
- 150. Deng, Z., A.E. Campbell, and P.M. Lieberman, *TERRA, CpG methylation, and telomere heterochromatin: lessons from ICF syndrome cells.* Cell Cycle, **2010**. 9(1): p. 69-74.
- 151. De Silanes, I.L., M.S. d'Alcontres, and M.A. Blasco, *TERRA transcripts are bound by a complex array of RNA-binding proteins.* Nature communications, **2010**. 1: p. 33.
- 152. Feuerhahn, S., N. Iglesias, A. Panza, A. Porro, and J. Lingner, *TERRA biogenesis, turnover and implications for function.* FEBS Lett., **2010**. 584(17): p. 3812-3818.
- 153. Redon, S., P. Reichenbach, and J. Lingner, *The non-coding RNA TERRA is a natural ligand and direct inhibitor of human telomerase.* Nucleic Acids Res., **2010**: p. gkq296.
- Biffi, G., D. Tannahill, and S. Balasubramanian, An Intramolecular G-Quadruplex Structure Is Required for Binding of Telomeric Repeat-Containing RNA to the Telomeric Protein TRF2. J. Am. Chem. Soc., 2012. 134(29): p. 11974-11976.
- 155. Dai, J., M. Carver, and D. Yang, *Polymorphism of human telomeric quadruplex structures.* Biochimie, **2008**. 90(8): p. 1172-1183.

- 156. Xu, Y., K. Kaminaga, and M. Komiyama, *G*-quadruplex formation by human telomeric repeatscontaining RNA in Na+ solution. J. Am. Chem. Soc., **2008**. 130(33): p. 11179-11184.
- 157. Martadinata, H. and A.T. Phan, Structure of propeller-type parallel-stranded RNA Gquadruplexes, formed by human telomeric RNA sequences in K+ solution. J. Am. Chem. Soc., 2009. 131(7): p. 2570-2578.
- 158. Zhang, D.-H., T. Fujimoto, S. Saxena, H.-Q. Yu, D. Miyoshi, and N. Sugimoto, *Monomorphic RNA G-quadruplex and polymorphic DNA G-quadruplex structures responding to cellular environmental factors*. Biochemistry, **2010**. 49(21): p. 4554-4563.
- 159. Xu, Y., *Chemistry in human telomere biology: structure, function and targeting of telomere DNA/RNA.* Chem. Soc. Rev., **2011**. 40(5): p. 2719-2740.
- 160. Xu, Y., Human telomere RNA: a potential target for ligand recognition. Curr. Pharm. Des.,
   2012. 18(14): p. 2096-2101.
- 161. Martadinata, H. and A.T.n. Phan, *Structure of human telomeric RNA (TERRA): stacking of two G-quadruplex blocks in K+ solution*. Biochemistry, **2013**. 52(13): p. 2176-2183.
- 162. Xu, Y., Y. Suzuki, and M. Komiyama, Click Chemistry for the Identification of G-Quadruplex Structures: Discovery of a DNA-RNA G-Quadruplex. Angew. Chem. Int. Ed., 2009. 48(18): p. 3281-3284.
- Martadinata, H. and A.T. Phan, Structure of propeller-type parallel-stranded RNA Gquadruplexes, formed by human telomeric RNA sequences in K+ solution. J. Am. Chem. Soc., 2009. 131(7): p. 2570-8.
- 164. Joachimi, A., A. Benz, and J.S. Hartig, *A comparison of DNA and RNA quadruplex structures and stabilities.* Biorg. Med. Chem., **2009**. 17(19): p. 6811-6815.
- 165. Collie, G.W., S.M. Haider, S. Neidle, and G.N. Parkinson, *A crystallographic and modelling study of a human telomeric RNA (TERRA) quadruplex*. Nucleic Acids Res., **2010**. 38(16): p. 5569-5580.
- 166. Zahler, A.M. and J.R. Williamson, *Inhibition of telomerase by G-quartet DNA structure*. Nature, **1991**. 350(6320): p. 718.
- 167. Sun, D., B. Thompson, B.E. Cathers, M. Salazar, S.M. Kerwin, J.O. Trent, T.C. Jenkins, S. Neidle, and L.H. Hurley, *Inhibition of human telomerase by a G-quadruplex-interactive compound*. J. Med. Chem., **1997**. 40(14): p. 2113-2116.
- 168. Paeschke, K., K.R. McDonald, and V.A. Zakian, *Telomeres: structures in need of unwinding.* FEBS Lett., **2010**. 584(17): p. 3760-3772.
- 169. Sharma, S., K.M. Doherty, and R.M. Brosh, *Mechanisms of RecQ helicases in pathways of DNA metabolism and maintenance of genomic stability.* Biochem. J., **2006**. 398(3): p. 319-337.
- 170. Paeschke, K., S. Juranek, T. Simonsson, A. Hempel, D. Rhodes, and H.J. Lipps, *Telomerase recruitment by the telomere end binding protein-β facilitates G-quadruplex DNA unfolding in ciliates.* Nat. Struct. Mol. Biol., **2008**. 15(6): p. 598-604.

- 171. Chavez, A., A.M. Tsou, and F.B. Johnson, *Telomeres do the (un) twist: helicase actions at chromosome termini*. Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease, 2009. 1792(4): p. 329-340.
- 172. Postberg, J., M. Tsytlonok, D. Sparvoli, D. Rhodes, and H.J. Lipps, *A telomerase-associated RecQ protein-like helicase resolves telomeric G-quadruplex structures during replication.* Gene, **2012**. 497(2): p. 147-154.
- 173. A Juranek, S. and K. Paeschke, *Cell cycle regulation of G-quadruplex DNA structures at telomeres.* Curr. Pharm. Des., **2012**. 18(14): p. 1867-1872.
- 174. Brosh Jr, R.M., *DNA helicases involved in DNA repair and their roles in cancer*. Nature Reviews Cancer, **2013**. 13(8): p. 542-558.
- 175. Wickramasinghe, C.M., H. Arzouk, A. Frey, A. Maiter, and J.E. Sale, *Contributions of the specialised DNA polymerases to replication of structured DNA*. DNA Repair, **2015**. 29: p. 83-90.
- 176. Haeusler, A.R., C.J. Donnelly, G. Periz, E.A. Simko, P.G. Shaw, M.-S. Kim, N.J. Maragakis, J.C. Troncoso, A. Pandey, and R. Sattler, *C9orf72 nucleotide repeat structures initiate molecular cascades of disease.* Nature, **2014**. 507(7491): p. 195-200.
- 177. Mergny, J.-L. and C. Hélène, *G-quadruplex DNA: a target for drug design.* Nat. Med., **1998**.
   4(12): p. 1366-1367.
- 178. Kerwin, S.M., *G*-quadruplex DNA as a target for drug design. Curr. Pharm. Des., **2000**. 6(4): p. 441-471.
- 179. Ou, T.m., Y.j. Lu, J.h. Tan, Z.s. Huang, K.Y. Wong, and L.q. Gu, *G-Quadruplexes: Targets in Anticancer Drug Design*. ChemMedChem, **2008**. 3(5): p. 690-713.
- 180. Wheelhouse, R.T., D. Sun, H. Han, F.X. Han, and L.H. Hurley, *Cationic porphyrins as telomerase inhibitors: the interaction of tetra-(N-methyl-4-pyridyl) porphine with quadruplex DNA.* J. Am. Chem. Soc., **1998**. 120(13): p. 3261-3262.
- Haider, S.M., G.N. Parkinson, and S. Neidle, *Structure of a G-quadruplex-ligand complex*. J. Mol. Biol., **2003**. 326(1): p. 117-125.
- 182. Gavathiotis, E., R.A. Heald, M.F. Stevens, and M.S. Searle, Drug recognition and stabilisation of the parallel-stranded DNA quadruplex d (TTAGGGT)<sub>4</sub> containing the human telomeric repeat. J. Mol. Biol., 2003. 334(1): p. 25-36.
- Schouten, J.A., S. Ladame, S.J. Mason, M.A. Cooper, and S. Balasubramanian, *G-quadruplex-specific peptide- hemicyanine ligands by partial combinatorial selection.* J. Am. Chem. Soc., 2003. 125(19): p. 5594-5595.
- 184. Neidle, S. and G. Parkinson, *Telomere maintenance as a target for anticancer drug discovery*.
   Nature Reviews Drug Discovery, **2002**. 1(5): p. 383-393.
- 185. Riou, J., L. Guittat, P. Mailliet, A. Laoui, E. Renou, O. Petitgenet, F. Megnin-Chanet, C. Helene, and J. Mergny, *Cell senescence and telomere shortening induced by a new series of specific Gquadruplex DNA ligands.* Proceedings of the National Academy of Sciences, **2002**. 99(5): p. 2672-2677.

- 186. Cuesta, J., M.A. Read, and S. Neidle, *The design of G-quadruplex ligands as telomerase inhibitors.* Mini Rev. Med. Chem., **2003**. 3(1): p. 11-21.
- 187. Tauchi, T., K. Shin-Ya, G. Sashida, M. Sumi, A. Nakajima, T. Shimamoto, J.H. Ohyashiki, and K. Ohyashiki, *Activity of a novel G-quadruplex-interactive telomerase inhibitor, telomestatin (SOT-095), against human leukemia cells: involvement of ATM-dependent DNA damage response pathways.* Oncogene, **2003**. 22(34): p. 5338-5347.
- 188. Gomez, D., R. Paterski, T. Lemarteleur, K. Shin-ya, J.-L. Mergny, and J.-F. Riou, Interaction of telomestatin with the telomeric single-strand overhang. J. Biol. Chem., 2004. 279(40): p. 41487-41494.
- 189. Guittat, L., P. Alberti, D. Gomez, A. De Cian, G. Pennarun, T. Lemarteleur, C. Belmokhtar, R. Paterski, H. Morjani, and C. Trentesaux, *Targeting human telomerase for cancer therapeutics.* Cytotechnology, **2004**. 45(1-2): p. 75-90.
- 190. Pennarun, G., C. Granotier, L.R. Gauthier, D. Gomez, F. Hoffschir, E. Mandine, J.-F. Riou, J.-L. Mergny, P. Mailliet, and F.D. Boussin, *Apoptosis related to telomere instability and cell cycle alterations in human glioma cells treated by new highly selective G-quadruplex ligands.* Oncogene, **2005**. 24(18): p. 2917-2928.
- 191. Pendino, F., I. Tarkanyi, C. Dudognon, J. Hillion, M. Lanotte, J. Aradi, and E. Segal-Bendirdjian, *Telomeres and telomerase: Pharmacological targets for new anticancer strategies?* Curr. Cancer Drug Targets, **2006**. 6(2): p. 147-180.
- 192. Zhou, J., X. Zhu, Y. Lu, R. Deng, Z. Huang, Y. Mei, Y. Wang, W. Huang, Z. Liu, and L. Gu, Senescence and telomere shortening induced by novel potent G-quadruplex interactive agents, quindoline derivatives, in human cancer cell lines. Oncogene, **2006**. 25(4): p. 503-511.
- 193. Marsh, T.C. and E. Henderson, *G-wires: self-assembly of a telomeric oligonucleotide, d* (*GGGGTTGGGG*), *into large superstructures.* Biochemistry, **1994**. 33(35): p. 10718-10724.
- 194. Marsh, T.C., J. Vesenka, and E. Henderson, *A new DNA nanostructure, the G-wire, imaged by scanning probe microscopy.* Nucleic Acids Res., **1995**. 23(4): p. 696-700.
- 195. Calzolari, A., R. Di Felice, E. Molinari, and A. Garbesi, *G*-quartet biomolecular nanowires. Appl. Phys. Lett., **2002**. 80(18): p. 3331-3333.
- 196. Sondermann, A., C. Holste, R. Möller, W. Fritzsche, and W. Fritzsche. *Assembly Of G-Quartet Based DNA Superstructures (G-Wires)*. in *AIP Conf. Proc.* 2002. AIP.
- 197. Davis, J.T., *G*-quartets 40 years later: from 5'-GMP to molecular biology and supramolecular chemistry. Angew. Chem. Int. Ed., **2004**. 43(6): p. 668-698.
- 198. Kotlyar, A.B., N. Borovok, T. Molotsky, H. Cohen, E. Shapir, and D. Porath, *Long, Monomolecular Guanine-Based Nanowires.* Adv. Mater., **2005**. 17(15): p. 1901-1905.
- 199. Davis, J.T. and G.P. Spada, *Supramolecular architectures generated by self-assembly of guanosine derivatives.* Chem. Soc. Rev., **2007**. 36(2): p. 296-313.

- 200. Liu, S.P., S.H. Weisbrod, Z. Tang, A. Marx, E. Scheer, and A. Erbe, *Direct Measurement of Electrical Transport Through G-Quadruplex DNA with Mechanically Controllable Break Junction Electrodes.* Angew. Chem. Int. Ed., **2010**. 49(19): p. 3313-3316.
- 201. Kotia, R.B., L. Li, and L.B. McGown, *Separation of nontarget compounds by DNA aptamers.* Anal. Chem., **2000**. 72(4): p. 827-831.
- 202. Rehder, M.A. and L.B. McGown, Open-tubular capillary electrochromatography of bovine β-lactoglobulin variants A and B using an aptamer stationary phase. Electrophoresis, 2001. 22(17): p. 3759-3764.
- 203. Charles, J.A. and L.B. McGown, Separation of Trp-Arg and Arg-Trp using G-quartet-forming DNA oligonucleotides in open-tubular capillary electrochromatography. Electrophoresis, 2002. 23(11): p. 1599-1604.
- 204. Roe, S., D.J. Ritson, T. Garner, M. Searle, and J.E. Moses, *Tuneable DNA-based asymmetric catalysis using a G-quadruplex supramolecular assembly.* Chem. Commun., **2010**. 46(24): p. 4309-4311.
- 205. Wang, C., G. Jia, J. Zhou, Y. Li, Y. Liu, S. Lu, and C. Li, *Enantioselective Diels–Alder Reactions* with *G-Quadruplex DNA-Based Catalysts.* Angew. Chem. Int. Ed., **2012**. 51(37): p. 9352-9355.
- 206. German, I., D.D. Buchanan, and R.T. Kennedy, *Aptamers as ligands in affinity probe capillary electrophoresis.* Anal. Chem., **1998**. 70(21): p. 4540-4545.
- 207. Potyrailo, R.A., R.C. Conrad, A.D. Ellington, and G.M. Hieftje, *Adapting selected nucleic acid ligands (aptamers) to biosensors.* Anal. Chem., **1998**. 70(16): p. 3419-3425.
- 208. Hamaguchi, N., A. Ellington, and M. Stanton, *Aptamer beacons for the direct detection of proteins*. Anal. Biochem., **2001**. 294(2): p. 126-131.
- 209. Li, J.J., X. Fang, and W. Tan, *Molecular aptamer beacons for real-time protein recognition*. Biochem. Biophys. Res. Commun., **2002**. 292(1): p. 31-40.
- 210. Ueyama, H., M. Takagi, and S. Takenaka, A novel potassium sensing in aqueous media with a synthetic oligonucleotide derivative. Fluorescence resonance energy transfer associated with guanine quartet- potassium ion complex formation. J. Am. Chem. Soc., **2002**. 124(48): p. 14286-14287.
- 211. Song, S., L. Wang, J. Li, C. Fan, and J. Zhao, *Aptamer-based biosensors*. TrAC, Trends Anal. Chem., **2008**. 27(2): p. 108-117.
- 212. Guo, L., D. Nie, C. Qiu, Q. Zheng, H. Wu, P. Ye, Y. Hao, F. Fu, and G. Chen, A G-quadruplex based label-free fluorescent biosensor for lead ion. Biosensors Bioelectron., 2012. 35(1): p. 123-127.
- 213. Yue, Q., T. Shen, C. Wang, L. Wang, H. Li, S. Xu, H. Wang, and J. Liu, *Construction of a controllable förster resonance energy transfer system based on G-quadruplex for DNA sensing.* Biosensors Bioelectron., **2013**. 40(1): p. 75-81.
- 214. Bustamante, C., J.C. Macosko, and G.J. Wuite, *Grabbing the cat by the tail: manipulating molecules one by one.* Nature Reviews Molecular Cell Biology, **2000**. 1(2): p. 130-136.

- 215. Rief, M., M. Gautel, F. Oesterhelt, J.M. Fernandez, and H.E. Gaub, *Reversible unfolding of individual titin immunoglobulin domains by AFM*. Science, **1997**. 276(5315): p. 1109-1112.
- 216. Carrion-Vazquez, M., P.E. Marszalek, A.F. Oberhauser, and J.M. Fernandez, *Atomic force microscopy captures length phenotypes in single proteins.* Proceedings of the National Academy of Sciences, **1999**. 96(20): p. 11288-11292.
- 217. Carrion-Vazquez, M., A.F. Oberhauser, S.B. Fowler, P.E. Marszalek, S.E. Broedel, J. Clarke, and J.M. Fernandez, *Mechanical and chemical unfolding of a single protein: a comparison.* Proceedings of the National Academy of Sciences, **1999**. 96(7): p. 3694-3699.
- Marszalek, P.E., H. Lu, H. Li, M. Carrion-Vazquez, A.F. Oberhauser, K. Schulten, and J.M. Fernandez, *Mechanical unfolding intermediates in titin modules*. Nature, **1999**. 402(6757): p. 100-103.
- 219. Socci, N.D., J.N. Onuchic, and P.G. Wolynes, *Stretching lattice models of protein folding.* Proceedings of the National Academy of Sciences, **1999**. 96(5): p. 2031-2035.
- Yang, G., C. Cecconi, W.A. Baase, I.R. Vetter, W.A. Breyer, J.A. Haack, B.W. Matthews, F.W. Dahlquist, and C. Bustamante, *Solid-state synthesis and mechanical unfolding of polymers of T4 lysozyme.* Proceedings of the National Academy of Sciences, **2000**. 97(1): p. 139-144.
- 221. Fernandez, J.M. and H. Li, *Force-clamp spectroscopy monitors the folding trajectory of a single protein.* Science, **2004**. 303(5664): p. 1674-1678.
- 222. Chung, H.S., K. McHale, J.M. Louis, and W.A. Eaton, *Single-molecule fluorescence experiments determine protein folding transition path times.* Science, **2012**. 335(6071): p. 981-984.
- 223. Erie, D.A., G. Yang, H.C. Schultz, and C. Bustamante, DNA bending by Cro protein in specific and nonspecific complexes: implications for protein site recognition and specificity. Science, 1994. 266(5190): p. 1562.
- 224. Wang, H., Y. Yang, M.J. Schofield, C. Du, Y. Fridman, S.D. Lee, E.D. Larson, J.T. Drummond, E. Alani, and P. Hsieh, *DNA bending and unbending by MutS govern mismatch recognition and specificity.* Proceedings of the National Academy of Sciences, **2003**. 100(25): p. 14822-14827.
- 225. Radmacher, M., M. Fritz, H.G. Hansma, and P.K. Hansma, *Direct observation of enzyme activity with the atomic force microscope.* SCIENCE-NEW YORK THEN WASHINGTON-, **1994**: p. 1577-1577.
- 226. Wuite, G.J., S.B. Smith, M. Young, D. Keller, and C. Bustamante, *Single-molecule studies of the effect of template tension on T7 DNA polymerase activity.* Nature, **2000**. 404(6773): p. 103-106.
- 227. Kishino, A. and T. Yanagida, *Force measurements by micromanipulation of a single actin filament by glass needles.* **1988**.
- 228. Svoboda, K., C.F. Schmidt, B.J. Schnapp, and S.M. Block, *Direct observation of kinesin stepping by optical trapping interferometry*. Nature, **1993**. 365(6448): p. 721.
- 229. Junge, W., *ATP synthase and other motor proteins.* Proceedings of the National Academy of Sciences, **1999**. 96(9): p. 4735-4737.

- 230. Abbondanzieri, E.A., W.J. Greenleaf, J.W. Shaevitz, R. Landick, and S.M. Block, *Direct observation of base-pair stepping by RNA polymerase.* Nature, **2005**. 438(7067): p. 460-465.
- Morin, J.A., F.J. Cao, J.M. Lázaro, J.R. Arias-Gonzalez, J.M. Valpuesta, J.L. Carrascosa, M. Salas, and B. Ibarra, *Active DNA unwinding dynamics during processive DNA replication*. Proceedings of the National Academy of Sciences, **2012**. 109(21): p. 8115-8120.
- 232. Morin, J.A., F.J. Cao, J.M. Lázaro, J.R. Arias-Gonzalez, J.M. Valpuesta, J.L. Carrascosa, M. Salas, and B. Ibarra, *Mechano-chemical kinetics of DNA replication: identification of the translocation step of a replicative DNA polymerase.* Nucleic Acids Res., **2015**. 43(7): p. 3643-3652.
- 233. Morin, J.A., F. Cerrón, J. Jarillo, E. Beltran-Heredia, G.L. Ciesielski, J.R. Arias-Gonzalez, L.S. Kaguni, F.J. Cao, and B. Ibarra, *DNA synthesis determines the binding mode of the human mitochondrial single-stranded DNA-binding protein.* Nucleic Acids Res., **2017**.
- 234. Smith, S.B., L. Finzi, and C. Bustamante, *Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads.* Science, **1992**. 258(5085): p. 1122-1126.
- 235. Bustamante, C., J.F. Marko, E.D. Siggia, and S. Smith, *Entropic elasticity of lambda-phage DNA*. SCIENCE-NEW YORK THEN WASHINGTON-, **1994**: p. 1599-1599.
- 236. Marko, J.F. and E.D. Siggia, *Stretching dna*. Macromolecules, **1995**. 28(26): p. 8759-8770.
- 237. Wang, M.D., H. Yin, R. Landick, J. Gelles, and S.M. Block, *Stretching DNA with optical tweezers*. Biophys. J., **1997**. 72(3): p. 1335-1346.
- 238. Bustamante, C., S.B. Smith, J. Liphardt, and D. Smith, *Single-molecule studies of DNA mechanics*. Curr. Opin. Struct. Biol., **2000**. 10(3): p. 279-285.
- 239. Mehta, A.D., M. Rief, J.A. Spudich, D.A. Smith, and R.M. Simmons, *Single-molecule biomechanics with optical methods.* Science, **1999**. 283(5408): p. 1689-1695.
- 240. Clausen-Schaumann, H., M. Seitz, R. Krautbauer, and H.E. Gaub, *Force spectroscopy with single bio-molecules.* Curr. Opin. Chem. Biol., **2000**. 4(5): p. 524-530.
- Ishii, Y. and T. Yanagida, *Single molecule detection in life sciences*. Single Molecules, **2000**.1(1): p. 5-16.
- 242. Bockelmann, U., Single-molecule manipulation of nucleic acids. Curr. Opin. Struct. Biol., 2004. 14(3): p. 368-373.
- 243. Neuman, K.C. and A. Nagy, *Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy.* Nat. Methods, **2008**. 5(6): p. 491.
- 244. Neuman, K.C., E.H. Chadd, G.F. Liou, K. Bergman, and S.M. Block, *Characterization of photodamage to Escherichia coli in optical traps.* Biophys. J., **1999**. 77(5): p. 2856-2863.
- 245. Bennink, M.L., S.H. Leuba, G.H. Leno, J. Zlatanova, B.G. de Grooth, and J. Greve, Unfolding individual nucleosomes by stretching single chromatin fibers with optical tweezers. Nat. Struct. Mol. Biol., 2001. 8(7): p. 606-610.
- 246. Cherney, D.P., T.E. Bridges, and J.M. Harris, *Optical trapping of unilamellar phospholipid vesicles: investigation of the effect of optical forces on the lipid membrane shape by confocal-Raman microscopy.* Anal. Chem., **2004**. 76(17): p. 4920-4928.

- 247. Sacconi, L., I.M. Tolić-Nørrelykke, C. Stringari, R. Antolini, and F.S. Pavone, *Optical micromanipulations inside yeast cells*. Appl. Opt., **2005**. 44(11): p. 2001-2007.
- 248. Ashkin, A., J.M. Dziedzic, J. Bjorkholm, and S. Chu, *Observation of a single-beam gradient force optical trap for dielectric particles.* Opt. Lett., **1986**. 11(5): p. 288-290.
- 249. Arias-González, J. and M. Nieto-Vesperinas, *Optical forces on small particles: attractive and repulsive nature and plasmon-resonance conditions.* JOSA A, **2003**. 20(7): p. 1201-1209.
- 250. Ashkin, A. and J.M. Dziedzic, *Optical trapping and manipulation of viruses and bacteria.* Science, **1987**. 235: p. 1517-1521.
- 251. Ashkin, A., J.M. Dziedzic, and T. Yamane, *Optical trapping and manipulation of single cells using infrared laser beams.* Nature, **1987**. 330(6150): p. 769-771.
- 252. Ashkin, A. and J. Dziedzic, *Internal cell manipulation using infrared laser traps.* Proceedings of the National Academy of Sciences, **1989**. 86(20): p. 7914-7918.
- 253. Hormeño, S., B. Ibarra, F. Chichón, K. Habermann, B. Lange, J. Valpuesta, J. Carrascosa, and J. Arias-Gonzalez, *Single centrosome manipulation reveals its electric charge and associated dynamic structure.* Biophys. J., **2009**. 97(4): p. 1022-1030.
- 254. Arias-Gonzalez, J.R., *Optical tweezers to study viruses*, in *Structure and Physics of Viruses*. 2013, Springer. p. 273-304.
- 255. Hormeno, S. and J.R. Arias-Gonzalez, *Exploring mechanochemical processes in the cell with optical tweezers*. Biol. Cell, **2006**. 98(12): p. 679-695.
- 256. Smith, S.B., Y. Cui, and C. Bustamante, *Optical-trap force transducer that operates by direct measurement of light momentum.* Methods Enzymol., **2003**. 361: p. 134-162.
- 257. Fällman, E. and O. Axner, Influence of a glass-water interface on the on-axis trapping of micrometer-sized spherical objects by optical tweezers. Appl. Opt., 2003. 42(19): p. 3915-3926.
- 258. Neuman, K.C., E.A. Abbondanzieri, and S.M. Block, *Measurement of the effective focal shift in an optical trap.* Opt. Lett., **2005**. 30(11): p. 1318-1320.
- 259. Bustamante, C.J. and S.B. Smith, *Light-force sensor and method for measuring axial opticaltrap forces from changes in light momentum along an optic axis.* 2006, U.S. Patent No. 7133132 B2
- 260. de Lorenzo, S., M. Ribezzi-Crivellari, J.R. Arias-Gonzalez, S.B. Smith, and F. Ritort, A temperature-jump optical trap for single-molecule manipulation. Biophys. J., 2015. 108(12): p. 2854-2864.
- Selvam, S., D. Koirala, Z. Yu, and H. Mao, *Quantification of topological coupling between DNA superhelicity and G-quadruplex formation*. J. Am. Chem. Soc., **2014**. 136(40): p. 13967-13970.
- 262. Capitanio, M., D. Maggi, F. Vanzi, and F. Pavone, *FIONA in the trap: the advantages of combining optical tweezers and fluorescence.* Journal of Optics A: Pure and Applied Optics, 2007. 9(8): p. S157.

- 263. Cluzel, P., A. Lebrun, C. Heller, and R. Lavery, *DNA: an extensible molecule*. Science, 1996.
  271(5250): p. 792.
- Smith, S.B., Y. Cui, and C. Bustamente, Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. Science, 1996. 271(5250): p. 795.
- 265. Dhakal, S., Z. Yu, R. Konik, Y. Cui, D. Koirala, and H. Mao, *G*-quadruplex and *i*-motif are mutually exclusive in ILPR double-stranded DNA. Biophys. J., **2012**. 102(11): p. 2575-2584.
- 266. Dhakal, S., Y. Cui, D. Koirala, C. Ghimire, S. Kushwaha, Z. Yu, P.M. Yangyuoru, and H. Mao, Structural and mechanical properties of individual human telomeric G-quadruplexes in molecularly crowded solutions. Nucleic Acids Res., **2013**. 41(6): p. 3915-3923.
- Garavís, M., R. Bocanegra, E. Herrero-Galán, C. González, A. Villasante, and J.R. Arias-Gonzalez, *Mechanical unfolding of long human telomeric RNA (TERRA).* Chem. Commun., 2013. 49(57): p. 6397-6399.
- 268. Yangyuoru, P.M., A.Y. Zhang, Z. Shi, D. Koirala, S. Balasubramanian, and H. Mao, Mechanochemical Properties of Individual Human Telomeric RNA (TERRA) G-Quadruplexes. ChemBioChem, 2013. 14(15): p. 1931-1935.
- 269. Abraham Punnoose, J., Y. Cui, D. Koirala, P.M. Yangyuoru, C. Ghimire, P. Shrestha, and H. Mao, *Interaction of G-quadruplexes in the full-length 3' human telomeric overhang.* J. Am. Chem. Soc., **2014**. 136(52): p. 18062-18069.
- 270. Yu, Z. and H. Mao, Non-B DNA structures show diverse conformations and complex transition kinetics comparable to RNA or proteins--a perspective from mechanical unfolding and refolding experiments. Chem. Rec., **2013**. 13(1): p. 102-16.
- Williams, M.C., J.R. Wenner, I. Rouzina, and V.A. Bloomfield, *Effect of pH on the overstretching transition of double-stranded DNA: evidence of force-induced DNA melting.* Biophys. J., **2001**. 80(2): p. 874-881.
- Wenner, J.R., M.C. Williams, I. Rouzina, and V.A. Bloomfield, Salt dependence of the elasticity and overstretching transition of single DNA molecules. Biophys. J., 2002. 82(6): p. 3160-3169.
- 273. van Mameren, J., P. Gross, G. Farge, P. Hooijman, M. Modesti, M. Falkenberg, G.J. Wuite, and E.J. Peterman, *Unraveling the structure of DNA during overstretching by using multicolor, single-molecule fluorescence imaging.* Proceedings of the National Academy of Sciences, 2009. 106(43): p. 18231-18236.
- 274. Williams, M.C., I. Rouzina, and M.J. McCauley, *Peeling back the mystery of DNA overstretching*. Proceedings of the National Academy of Sciences, **2009**. 106(43): p. 18047-18048.
- 275. Fu, H., H. Chen, J.F. Marko, and J. Yan, *Two distinct overstretched DNA states*. Nucleic Acids Res., **2010**: p. gkq309.

- 276. Fu, H., H. Chen, X. Zhang, Y. Qu, J.F. Marko, and J. Yan, *Transition dynamics and selection of the distinct S-DNA and strand unpeeling modes of double helix overstretching.* Nucleic Acids Res., **2010**: p. gkq1278.
- 277. Paik, D.H. and T.T. Perkins, *Overstretching DNA at 65 pN does not require peeling from free ends or nicks.* J. Am. Chem. Soc., **2011**. 133(10): p. 3219-3221.
- 278. Bosaeus, N., A.H. El-Sagheer, T. Brown, S.B. Smith, B. Åkerman, C. Bustamante, and B. Nordén, *Tension induces a base-paired overstretched DNA conformation*. Proceedings of the National Academy of Sciences, **2012**. 109(38): p. 15179-15184.
- 279. Zhang, X., H. Chen, H. Fu, P.S. Doyle, and J. Yan, *Two distinct overstretched DNA structures revealed by single-molecule thermodynamics measurements.* Proceedings of the National Academy of Sciences, **2012**. 109(21): p. 8103-8108.
- 280. King, G.A., P. Gross, U. Bockelmann, M. Modesti, G.J. Wuite, and E.J. Peterman, *Revealing the competition between peeled ssDNA, melting bubbles, and S-DNA during DNA overstretching using fluorescence microscopy.* Proceedings of the National Academy of Sciences, **2013**. 110(10): p. 3859-3864.
- 281. Zhang, X., H. Chen, S. Le, I. Rouzina, P.S. Doyle, and J. Yan, *Revealing the competition between peeled ssDNA, melting bubbles, and S-DNA during DNA overstretching by single-molecule calorimetry.* Proceedings of the National Academy of Sciences, **2013**. 110(10): p. 3865-3870.
- 282. Peterman, E.J., F. Gittes, and C.F. Schmidt, *Laser-induced heating in optical traps.* Biophys. J.,
  2003. 84(2): p. 1308-1316.
- 283. Seol, Y., A.E. Carpenter, and T.T. Perkins, *Gold nanoparticles: enhanced optical trapping and sensitivity coupled with significant heating.* Opt. Lett., **2006**. 31(16): p. 2429-2431.
- 284. Hormeño, S., P. Gregorio-Godoy, J. Pérez-Juste, L.M. Liz-Marzán, B.H. Juárez, and J.R. Arias-Gonzalez, *Laser Heating Tunability by Off-Resonant Irradiation of Gold Nanoparticles*. Small, 2014. 10(2): p. 376-384.
- 285. Hormeno, S. and J.R. Arias-Gonzalez, *Exploring mechanochemical processes in the cell with optical tweezers.* Biol. Cell, **2006**. 98(12): p. 679-95.
- 286. Binnig, G., N. Garcia, and H. Rohrer, *Conductivity sensitivity of inelastic scanning tunneling microscopy.* Physical Review B, **1985**. 32(2): p. 1336.
- 287. Binnig, G., C.F. Quate, and C. Gerber, *Atomic force microscope*. Phys. Rev. Lett., **1986**. 56(9): p. 930.
- 288. Shao, Z., J. Yang, and A.P. Somlyo, *Biological atomic force microscopy: from microns to nanometers and beyond.* Annu. Rev. Cell. Dev. Biol., **1995**. 11(1): p. 241-265.
- 289. Bustamante, C., C. Rivetti, and D.J. Keller, *Scanning force microscopy under aqueous solutions.* Curr. Opin. Struct. Biol., **1997**. 7(5): p. 709-716.
- 290. Hugel, T. and M. Seitz, *The study of molecular interactions by AFM force spectroscopy.* Macromol. Rapid Commun., **2001**. 22(13): p. 989-1016.
- 291. Eaton, P. and P. West, *Atomic force microscopy*. 2010: Oxford University Press.

- 292. Faraldos, M. and C. Goberna, *Técnicas de análisis y caracterización de materiales*. 2002: Consejo Superior de Investigaciones Científicas.
- 293. Allen, M.J., X.F. Dong, T.E. O'Neill, P. Yau, S.C. Kowalczykowski, J. Gatewood, R. Balhorn, and E.M. Bradbury, *Atomic force microscope measurements of nucleosome cores assembled along defined DNA sequences.* Biochemistry, **1993**. 32(33): p. 8390-8396.
- 294. Bustamante, C., D. Keller, and G. Yang, *Scanning force microscopy of nucleic acids and nucleoprotein assemblies.* Curr. Opin. Struct. Biol., **1993**. 3(3): p. 363-372.
- 295. Hansma, H.G., M. Bezanilla, F. Zenhausern, M. Adrian, and R.L. Sinsheimer, *Atomic force microscopy of DNA in aqueous solutions*. Nucleic Acids Res., **1993**. 21(3): p. 505-512.
- 296. Rees, W.A., R.W. Keller, J.P. Vesenka, G. Yang, and C. Bustamante, *Evidence of DNA bending in transcription complexes imaged by scanning force microscopy.* Science, **1993**. 260(5114): p. 1646-1650.
- 297. Leuba, S.H., G. Yang, C. Robert, B. Samori, K. van Holde, J. Zlatanova, and C. Bustamante, *Three-dimensional structure of extended chromatin fibers as revealed by tapping-mode scanning force microscopy.* Proceedings of the National Academy of Sciences, **1994**. 91(24): p. 11621-11625.
- 298. Bustamante, C., D.A. Erie, and D. Keller, *Biochemical and structural applications of scanning force microscopy.* Curr. Opin. Struct. Biol., **1994**. 4(5): p. 750-760.
- 299. Hansma, H., K. Kim, D. Laney, R. Garcia, M. Argaman, M. Allen, and S. Parsons, *Properties of biomolecules measured from atomic force microscope images: a review.* J. Struct. Biol., **1997**. 119(2): p. 99-108.
- 300. Costa, L., M. Kerkmann, G. Hartmann, S. Endres, P. Bisch, W. Heckl, and S. Thalhammer, Structural studies of oligonucleotides containing G-quadruplex motifs using AFM. Biochem. Biophys. Res. Commun., 2004. 313(4): p. 1065-1072.
- 301. Moreno-Herrero, F., J. Colchero, J. Gomez-Herrero, and A. Baro, Atomic force microscopy contact, tapping, and jumping modes for imaging biological samples in liquids. Physical Review E, 2004. 69(3): p. 031915.
- 302. Xu, Y. and M. Komiyama. *The structural studies of human telomeric DNA using AFM*. in *Nucleic Acids Symp. Ser.* 2007. Oxford Univ Press.
- 303. Neaves, K.J., J.L. Huppert, R.M. Henderson, and J.M. Edwardson, *Direct visualization of G-quadruplexes in DNA using atomic force microscopy.* Nucleic Acids Res., **2009**. 37(18): p. 6269-6275.
- 304. Xu, Y., T. Ishizuka, K. Kurabayashi, and M. Komiyama, Consecutive Formation of G-Quadruplexes in Human Telomeric-Overhang DNA: A Protective Capping Structure for Telomere Ends. Angew. Chem. Int. Ed., 2009. 48(42): p. 7833-7836.
- 305. Mela, I., R. Kranaster, R.M. Henderson, S. Balasubramanian, and J.M. Edwardson, *Demonstration of ligand decoration, and ligand-induced perturbation, of G-quadruplexes in a plasmid using atomic force microscopy.* Biochemistry, **2012**. 51(2): p. 578-585.

- 306. Fuentes-Perez, M.E., M.S. Dillingham, and F. Moreno-Herrero, *AFM volumetric methods for the characterization of proteins and nucleic acids.* Methods, **2013**. 60(2): p. 113-121.
- 307. Ares, P., M.E. Fuentes-Perez, E. Herrero-Galán, J.M. Valpuesta, A. Gil, J. Gomez-Herrero, and
   F. Moreno-Herrero, *High resolution atomic force microscopy of double-stranded RNA*.
   Nanoscale, 2016. 8(23): p. 11818-11826.
- 308. Randall, A. and J.D. Griffith, *Structure of long telomeric RNA transcripts the G-rich rna forms a compact repeating structure containing G-quartets.* J. Biol. Chem., **2009**. 284(21): p. 13980-13986.
- 309. Phan, A.T., *Human telomeric G-quadruplex: structures of DNA and RNA sequences.* The FEBS journal, **2010**. 277(5): p. 1107-1117.
- 310. Petraccone, L., *Higher-order quadruplex structures*, in *Quadruplex Nucleic Acids*. 2012, Springer. p. 23-46.
- 311. Mullis, K.B., H.A. Erlich, N. Arnheim, G.T. Horn, R.K. Saiki, and S.J. Scharf, *Process for amplifying, detecting, and/or-cloning nucleic acid sequences.* 1987, Google Patents.
- 312. Horcas, I., R. Fernández, J. Gomez-Rodriguez, J. Colchero, J. Gómez-Herrero, and A. Baro, WSXM: a software for scanning probe microscopy and a tool for nanotechnology. Rev. Sci. Instrum., 2007. 78(1): p. 013705.
- 313. Saenger, W., Principles of nucleic acid structure. 2nd ed. 1984: Springer-Verlag.
- 314. Laughlan, G., A. Murchie, D. Norman, M. Moore, P. Moody, D. Lilley, and B. Luisi, *The highresolution crystal structure of a parallel-stranded guanine tetraplex.* Science, **1994**. 265(5171): p. 520-524.
- 315. Moreno-Herrero, F., J. Colchero, and A. Baro, *DNA height in scanning force microscopy*. Ultramicroscopy, **2003**. 96(2): p. 167-174.
- 316. Bustamante, C., Y.R. Chemla, N.R. Forde, and D. Izhaky, *Mechanical processes in biochemistry*. Annu. Rev. Biochem., **2004**. 73(1): p. 705-748.
- 317. Tinoco, I., G. Chen, and X. Qu, *RNA reactions one molecule at a time.* Cold Spring Harb. Perspect. Biol., **2010**. 2(11): p. a003624.
- 318. Gray, R.D., J.O. Trent, and J.B. Chaires, *Folding and unfolding pathways of the human telomeric G-quadruplex*. J. Mol. Biol., **2014**. 426(8): p. 1629-50.
- 319. Mashimo, T., H. Yagi, Y. Sannohe, A. Rajendran, and H. Sugiyama, *Folding pathways of human telomeric type-1 and type-2 G-quadruplex structures.* J. Am. Chem. Soc., **2010**. 132(42): p. 14910-8.
- Rajendran, A., M. Endo, K. Hidaka, and H. Sugiyama, Direct and single-molecule visualization of the solution-state structures of G-hairpin and G-triplex intermediates. Angew. Chem. Int. Ed. Engl., 2014. 53(16): p. 4107-12.
- 321. Bustamante, C., *Unfolding single RNA molecules: bridging the gap between equilibrium and non-equilibrium statistical thermodynamics.* Q. Rev. Biophys., **2005**. 38(04): p. 291-301.

- 322. Yu, Z., J.D. Schonhoft, S. Dhakal, R. Bajracharya, R. Hegde, S. Basu, and H. Mao, *ILPR G-quadruplexes formed in seconds demonstrate high mechanical stabilities.* J. Am. Chem. Soc., **2009**. 131(5): p. 1876-82.
- 323. de Messieres, M., J.C. Chang, B. Brawn-Cinani, and A. La Porta, *Single-molecule study of G-quadruplex disruption using dynamic force spectroscopy.* Phys. Rev. Lett., **2012**. 109(5): p. 058101.
- 324. Lorenz, R., S.H. Bernhart, J. Qin, C. Honer zu Siederdissen, A. Tanzer, F. Amman, I.L. Hofacker, and P.F. Stadler, *2D meets 4G: G-quadruplexes in RNA secondary structure prediction.* IEEE/ACM Trans. Comput. Biol. Bioinform., **2013**. 10(4): p. 832-44.
- 325. Bergues-Pupo, A.E., J.R. Arias-Gonzalez, M.C. Morón, A. Fiasconaro, and F. Falo, *Role of the central cations in the mechanical unfolding of DNA and RNA G-quadruplexes.* Nucleic Acids Res., **2015**. 43(15): p. 7638-7647.
- 326. Bugaut, A., P. Murat, and S. Balasubramanian, *An RNA hairpin to G-quadruplex conformational transition.* J. Am. Chem. Soc., **2012**. 134(49): p. 19953-6.
- 327. Hardin, C.C., T. Watson, M. Corregan, and C. Bailey, *Cation-dependent transition between the quadruplex and Watson-Crick hairpin forms of d(CGCG3GCG).* Biochemistry, **1992**. 31(3): p. 833-41.
- 328. Kuo, M.H., Z.F. Wang, T.Y. Tseng, M.H. Li, S.T. Hsu, J.J. Lin, and T.C. Chang, *Conformational transition of a hairpin structure to G-quadruplex within the WNT1 gene promoter.* J. Am. Chem. Soc., **2015**. 137(1): p. 210-8.
- Bustamante, C., *In singulo biochemistry: when less is more.* Annu. Rev. Biochem., **2008**. 77: p. 45-50.
- 330. Herrero-Galan, E., M.E. Fuentes-Perez, C. Carrasco, J.M. Valpuesta, J.L. Carrascosa, F. Moreno-Herrero, and J.R. Arias-Gonzalez, *Mechanical identities of RNA and DNA double helices unveiled at the single-molecule level.* J. Am. Chem. Soc., **2013**. 135(1): p. 122-31.
- 331. Yangyuoru, P.M., A.Y. Zhang, Z. Shi, D. Koirala, S. Balasubramanian, and H. Mao, Mechanochemical properties of individual human telomeric RNA (TERRA) G-quadruplexes. ChemBioChem, 2013. 14(15): p. 1931-5.
- 332. Alemany, A. and F. Ritort, *Determination of the elastic properties of short ssDNA molecules by mechanically folding and unfolding DNA hairpins.* Biopolymers, **2014**. 101(12): p. 1193-9.
- 333. Bizarro, C.V., A. Alemany, and F. Ritort, *Non-specific binding of Na+ and Mg2+ to RNA determined by force spectroscopy methods*. Nucleic Acids Res., **2012**. 40(14): p. 6922-35.
- 334. Bosco, A., J. Camunas-Soler, and F. Ritort, *Elastic properties and secondary structure formation of single-stranded DNA at monovalent and divalent salt conditions.* Nucleic Acids Res., **2014**. 42(3): p. 2064-74.
- 335. Dessinges, M.-N., B. Maier, Y. Zhang, M. Peliti, D. Bensimon, and V. Croquette, *Stretching single stranded DNA, a model polyelectrolyte.* Phys. Rev. Lett., **2002**. 89(24): p. 248102.
- 336. Zhang, Z., J. Dai, E. Veliath, R.A. Jones, and D. Yang, *Structure of a two-G-tetrad intramolecular G-quadruplex formed by a variant human telomeric sequence in K+ solution:*

*insights into the interconversion of human telomeric G-quadruplex structures.* Nucleic Acids Res., **2010**. 38(3): p. 1009-21.

- 337. Zhang, A.Y. and S. Balasubramanian, *The kinetics and folding pathways of intramolecular Gquadruplex nucleic acids.* J. Am. Chem. Soc., **2012**. 134(46): p. 19297-308.
- 338. Li, W., X.M. Hou, P.Y. Wang, X.G. Xi, and M. Li, *Direct measurement of sequential folding pathway and energy landscape of human telomeric G-quadruplex structures.* J. Am. Chem. Soc., **2013**. 135(17): p. 6423-6.
- You, H., X. Zeng, Y. Xu, C.J. Lim, A.K. Efremov, A.T. Phan, and J. Yan, *Dynamics and stability of polymorphic human telomeric G-quadruplex under tension*. Nucleic Acids Res., 2014. 42(13): p. 8789-95.
- 340. Koirala, D., T. Mashimo, Y. Sannohe, Z. Yu, H. Mao, and H. Sugiyama, *Intramolecular folding in three tandem guanine repeats of human telomeric DNA*. Chem. Commun. (Camb.), **2012**. 48(14): p. 2006-8.
- Long, X., J.W. Parks, C.R. Bagshaw, and M.D. Stone, *Mechanical unfolding of human telomere G-quadruplex DNA probed by integrated fluorescence and magnetic tweezers spectroscopy*. Nucleic Acids Res., **2013**. 41(4): p. 2746-2755.
- 342. Koirala, D., C. Ghimire, C. Bohrer, Y. Sannohe, H. Sugiyama, and H. Mao, *Long-loop G-quadruplexes are misfolded population minorities with fast transition kinetics in human telomeric sequences.* J. Am. Chem. Soc., **2013**. 135(6): p. 2235-41.
- 343. Hsin, J., J. Strumpfer, E.H. Lee, and K. Schulten, *Molecular origin of the hierarchical elasticity of titin: simulation, experiment, and theory.* Annu Rev Biophys, **2011**. 40: p. 187-203.
- 344. Jacobson, D.R., D.B. McIntosh, and O.A. Saleh, *The snakelike chain character of unstructured RNA*. Biophys. J., **2013**. 105(11): p. 2569-76.
- 345. Zuker, M., *Mfold web server for nucleic acid folding and hybridization prediction*. Nucleic Acids Res., **2003**. 31(13): p. 3406-15.
- 346. Collie, G.W. and G.N. Parkinson, *The application of DNA and RNA G-quadruplexes to therapeutic medicines.* Chem. Soc. Rev., **2011**. 40(12): p. 5867-5892.
- 347. Boncina, M., J. Lah, I. Prislan, and G. Vesnaver, *Energetic basis of human telomeric DNA folding into G-quadruplex structures*. J. Am. Chem. Soc., **2012**. 134(23): p. 9657-63.
- Jarzynski, C., Nonequilibrium Equality for Free Energy Differences. Phys. Rev. Lett., 1997.
   78(14): p. 2690-2693.
- 349. Mergny, J.-L., A.-T. Phan, and L. Lacroix, *Following G-quartet formation by UV-spectroscopy*. FEBS Lett., **1998**. 435(1): p. 74-78.
- 350. Yurenko, Y.P., J. Novotný, V. Sklenář, and R. Marek, *Exploring non-covalent interactions in guanine-and xanthine-based model DNA quadruplex structures: a comprehensive quantum chemical approach.* PCCP, **2014**. 16(5): p. 2072-2084.
- 351. Poudel, L., N.F. Steinmetz, R.H. French, V.A. Parsegian, R. Podgornik, and W.-Y. Ching, Implication of the solvent effect, metal ions and topology in the electronic structure and

hydrogen bonding of human telomeric G-quadruplex DNA. PCCP, **2016**. 18(31): p. 21573-21585.

- Li, H., E.-h. Cao, and T. Gisler, Force-induced unfolding of human telomeric G-quadruplex: a steered molecular dynamics simulation study. Biochem. Biophys. Res. Commun., 2009. 379(1): p. 70-75.
- 353. Li, M.-H., Q. Luo, X.-G. Xue, and Z.-S. Li, *Toward a full structural characterization of G-quadruplex DNA in aqueous solution: Molecular dynamics simulations of four G-quadruplex molecules.* Journal of Molecular Structure: THEOCHEM, **2010**. 952(1): p. 96-102.
- 354. Yang, C., S. Jang, and Y. Pak, *Multiple stepwise pattern for potential of mean force in unfolding the thrombin binding aptamer in complex with Sr2+.* The Journal of chemical physics, **2011**. 135(22): p. 12B610.
- 355. Islam, B., M. Sgobba, C. Laughton, M. Orozco, J. Sponer, S. Neidle, and S. Haider, *Conformational dynamics of the human propeller telomeric DNA quadruplex on a microsecond time scale.* Nucleic Acids Res., **2013**. 41(4): p. 2723-2735.
- 356. Stadlbauer, P., M. Krepl, T.E. Cheatham, J. Koča, and J. Šponer, *Structural dynamics of possible late-stage intermediates in folding of quadruplex DNA studied by molecular simulations.* Nucleic Acids Res., **2013**: p. gkt412.
- 357. Linak, M.C., R. Tourdot, and K.D. Dorfman, *Moving beyond Watson–Crick models of coarse grained DNA dynamics.* The Journal of chemical physics, **2011**. 135(20): p. 11B613.
- 358. Rebic, M., F. Mocci, A. Laaksonen, and J. Ulicny, *Multiscale simulations of human telomeric G-quadruplex DNA*. The Journal of Physical Chemistry B, **2014**. 119(1): p. 105-113.
- 359. Bergues-Pupo, A.E., I. Gutierrez, J.R. Arias-Gonzalez, F. Falo, and A. Fiasconaro, *Mesoscopic model for DNA G-quadruplex unfolding.* arXiv preprint arXiv:1705.10778, **2017**.
- 360. Dudko, O.K., G. Hummer, and A. Szabo, *Intrinsic rates and activation free energies from single-molecule pulling experiments.* Phys. Rev. Lett., **2006**. 96(10): p. 108101.
- Friddle, R.W., A. Noy, and J.J. De Yoreo, *Interpreting the widespread nonlinear force spectra of intermolecular bonds.* Proceedings of the National Academy of Sciences, **2012**. 109(34): p. 13573-13578.
- 362. Hurley, L.H., *DNA and its associated processes as targets for cancer therapy.* Nature reviews. Cancer, **2002**. 2(3): p. 188.
- 363. Shay, J.W. and W.E. Wright, *Telomerase: a target for cancer therapeutics.* Cancer Cell, 2002.
  2(4): p. 257-265.
- 364. Moorhouse, A.D., A.M. Santos, M. Gunaratnam, M. Moore, S. Neidle, and J.E. Moses, *Stabilization of G-quadruplex DNA by highly selective ligands via click chemistry.* J. Am. Chem. Soc., 2006. 128(50): p. 15972-15973.
- 365. Parkinson, G.N., R. Ghosh, and S. Neidle, *Structural basis for binding of porphyrin to human telomeres*. Biochemistry, **2007**. 46(9): p. 2390-2397.
- 366. De Cian, A., L. Lacroix, C. Douarre, N. Temime-Smaali, C. Trentesaux, J.-F. Riou, and J.-L. Mergny, *Targeting telomeres and telomerase*. Biochimie, **2008**. 90(1): p. 131-155.

- 367. Neidle, S. and G.N. Parkinson, *Quadruplex DNA crystal structures and drug design*. Biochimie, **2008**. 90(8): p. 1184-1196.
- 368. Parkinson, G.N., F. Cuenca, and S. Neidle, *Topology conservation and loop flexibility in quadruplex-drug recognition: crystal structures of inter-and intramolecular telomeric DNA quadruplex-drug complexes.* J. Mol. Biol., **2008**. 381(5): p. 1145-1156.
- 369. Han, F.X., R.T. Wheelhouse, and L.H. Hurley, *Interactions of TMPyP4 and TMPyP2 with quadruplex DNA. Structural basis for the differential effects on telomerase inhibition.* J. Am. Chem. Soc., **1999**. 121(15): p. 3561-3570.
- 370. Koeppel, F., J.-F. Riou, A. Laoui, P. Mailliet, P.B. Arimondo, D. Labit, O. Petitgenet, C. Hélène, and J.-L. Mergny, *Ethidium derivatives bind to G-quartets, inhibit telomerase and act as fluorescent probes for quadruplexes.* Nucleic Acids Res., **2001**. 29(5): p. 1087-1096.
- 371. Riou, J.F., L. Guittat, P. Mailliet, A. Laoui, E. Renou, O. Petitgenet, F. Mégnin-Chanet, C. Hélène, and J.L. Mergny, *Cell senescence and telomere shortening induced by a new series of specific G-quadruplex DNA ligands.* Proceedings of the National Academy of Sciences, 2002. 99(5): p. 2672-2677.
- 372. Burger, A.M., F. Dai, C.M. Schultes, A.P. Reszka, M.J. Moore, J.A. Double, and S. Neidle, *The G-quadruplex-interactive molecule BRACO-19 inhibits tumor growth, consistent with telomere targeting and interference with telomerase function.* Cancer Res., **2005**. 65(4): p. 1489-1496.
- 373. Franceschin, M., G-Quadruplex DNA Structures and Organic Chemistry: More Than One Connection. Eur. J. Org. Chem., 2009. 2009(14): p. 2225-2238.
- 374. Bare, G.A., B. Liu, and J.C. Sherman, *Synthesis of a single G-quartet platform in water.* J. Am. Chem. Soc., **2013**. 135(32): p. 11985-11989.
- 375. Chung, W.J., B. Heddi, M. Tera, K. Iida, K. Nagasawa, and A.T.n. Phan, Solution structure of an intramolecular (3+ 1) human telomeric G-quadruplex bound to a telomestatin derivative. J. Am. Chem. Soc., 2013. 135(36): p. 13495-13501.
- 376. Ranjan, N., E. Davis, L. Xue, and D.P. Arya, *Dual recognition of the human telomeric G-quadruplex by a neomycin–anthraquinone conjugate.* Chem. Commun., **2013**. 49(51): p. 5796-5798.
- 377. Wang, X., C.-X. Zhou, J.-W. Yan, J.-Q. Hou, S.-B. Chen, T.-M. Ou, L.-Q. Gu, Z.-S. Huang, and J.-H. Tan, Synthesis and evaluation of quinazolone derivatives as a new class of c-KIT G-quadruplex binding ligands. ACS Med. Chem. Lett., 2013. 4(10): p. 909-914.
- 378. Reed, J.E., A.A. Arnal, S. Neidle, and R. Vilar, Stabilization of G-quadruplex DNA and inhibition of telomerase activity by square-planar nickel (II) complexes. J. Am. Chem. Soc., 2006. 128(18): p. 5992-5993.
- 379. Ma, D.-L., C.-M. Che, and S.-C. Yan, *Platinum (II) complexes with dipyridophenazine ligands as human telomerase inhibitors and luminescent probes for G-quadruplex DNA.* J. Am. Chem. Soc., **2008**. 131(5): p. 1835-1846.

- 380. Georgiades, S.N., N.H. Abd Karim, K. Suntharalingam, and R. Vilar, *Interaction of metal complexes with G-quadruplex DNA*. Angew. Chem. Int. Ed., **2010**. 49(24): p. 4020-4034.
- 381. Castor, K.J., J. Mancini, J. Fakhoury, N. Weill, R. Kieltyka, P. Englebienne, N. Avakyan, A. Mittermaier, C. Autexier, and N. Moitessier, *Platinum (II) Phenanthroimidazoles for Targeting Telomeric G-Quadruplexes.* ChemMedChem, **2012**. 7(1): p. 85-94.
- 382. Sun, D., Y. Liu, D. Liu, R. Zhang, X. Yang, and J. Liu, *Stabilization of G-quadruplex DNA*, *inhibition of telomerase activity and live cell imaging studies of chiral ruthenium (II) complexes.* Chemistry-a European Journal, **2012**. 18(14): p. 4285-4295.
- 383. Suntharalingam, K., A. Łęczkowska, M.A. Furrer, Y. Wu, M.K. Kuimova, B. Therrien, A.J. White, and R. Vilar, A cyclometallated platinum complex as a selective optical switch for quadruplex DNA. Chemistry-A European Journal, 2012. 18(51): p. 16277-16282.
- 384. Chung, C.Y.-S. and V.W.-W. Yam, Selective label-free detection of G-quadruplex structure of human telomere by emission spectral changes in visible-and-NIR region under physiological condition through the FRET of a two-component PPE-SO 3–-Pt (ii) complex ensemble with Pt<sup>I</sup> Pt, electrostatic and  $\pi$ - $\pi$  interactions. Chemical Science, **2013**. 4(1): p. 377-387.
- 385. Wang, J., K. Lu, S. Xuan, Z. Toh, D. Zhang, and F. Shao, *A Pt (II)–Dip complex stabilizes parallel c-myc G-quadruplex.* Chem. Commun., **2013**. 49(42): p. 4758-4760.
- 386. von Grebe, P., K. Suntharalingam, R. Vilar, S. Miguel, J. Pablo, S. Herres-Pawlis, and B. Lippert, A Conformationally Flexible Dinuclear PtII Complex with Differential Behavior of its Two States toward Quadruplex DNA. Chemistry-A European Journal, 2013. 19(34): p. 11429-11438.
- 387. Tan, J.-H., L.-Q. Gu, and J.-Y. Wu, *Design of selective G-quadruplex ligands as potential anticancer agents.* Mini Rev. Med. Chem., **2008**. 8(11): p. 1163-1178.
- 388. Li, T., E. Wang, and S. Dong, Parallel G-quadruplex-specific fluorescent probe for monitoring DNA structural changes and label-free detection of potassium ion. Anal. Chem., 2010. 82(18): p. 7576-7580.
- 389. Hamon, F., E. Largy, A. Guédin-Beaurepaire, M. Rouchon-Dagois, A. Sidibe, D. Monchaud, J.L. Mergny, J.F. Riou, C.H. Nguyen, and M.P. Teulade-Fichou, *An Acyclic Oligoheteroaryle That Discriminates Strongly between Diverse G-Quadruplex Topologies.* Angew. Chem. Int. Ed., 2011. 50(37): p. 8745-8749.
- 390. Wilson, T., M.P. Williamson, and J.A. Thomas, *Differentiating quadruplexes: binding preferences of a luminescent dinuclear ruthenium (II) complex with four-stranded DNA structures.* Org. Biomol. Chem., **2010**. 8(11): p. 2617-2621.
- Xu, L., D. Zhang, J. Huang, M. Deng, M. Zhang, and X. Zhou, *High fluorescence selectivity and visual detection of G-quadruplex structures by a novel dinuclear ruthenium complex.* Chem. Commun., 2010. 46(5): p. 743-745.
- 392. Liao, G.-L., X. Chen, L.-N. Ji, and H. Chao, *Visual specific luminescent probing of hybrid G-quadruplex DNA by a ruthenium polypyridyl complex.* Chem. Commun., **2012**. 48(87): p. 10781-10783.

- 393. Shi, S., J. Zhao, X. Gao, C. Lv, L. Yang, J. Hao, H. Huang, J. Yao, W. Sun, and T. Yao, Molecular "light switch" for G-quadruplex DNA: cycling the switch on and off. Dalton Transactions, 2012. 41(19): p. 5789-5793.
- Yao, J.-L., X. Gao, W. Sun, X.-Z. Fan, S. Shi, and T.-M. Yao, A naked-eye on-off-on molecular "light switch" based on a reversible "conformational switch" of G-quadruplex DNA. Inorg. Chem., 2012. 51(23): p. 12591-12593.
- 395. Wachter, E., D. Moyá, S. Parkin, and E.C. Glazer, Ruthenium Complex "Light Switches" that are Selective for Different G-Quadruplex Structures. Chemistry-A European Journal, 2016. 22(2): p. 550-559.
- Wang, Y., Y. Hu, T. Wu, H. Liu, L. Zhang, X. Zhou, and Y. Shao, Specific G-quadruplex structure recognition of human telomeric RNA over DNA by a fluorescently activated hyperporphyrin. Analyst, 2015. 140(15): p. 5169-5175.
- 397. Rajput, C., R. Rutkaite, L. Swanson, I. Haq, and J.A. Thomas, *Dinuclear monointercalating Rull complexes that display high affinity binding to duplex and quadruplex DNA.* Chemistry-A European Journal, **2006**. 12(17): p. 4611-4619.
- Shi, S., J. Liu, T. Yao, X. Geng, L. Jiang, Q. Yang, L. Cheng, and L. Ji, *Promoting the formation and stabilization of G-quadruplex by dinuclear Rull complex Ru2 (obip) L4.* Inorg. Chem., 2008. 47(8): p. 2910-2912.
- 399. Chen, T., Y. Liu, W.-J. Zheng, J. Liu, and Y.-S. Wong, *Ruthenium polypyridyl complexes that induce mitochondria-mediated apoptosis in cancer cells.* Inorg. Chem., **2010**. 49(14): p. 6366-6368.
- 400. Shi, S., X. Geng, J. Zhao, T. Yao, C. Wang, D. Yang, L. Zheng, and L. Ji, *Interaction of [Ru (bpy) 2 (dppz)] 2+ with human telomeric DNA: Preferential binding to G-quadruplexes over i-motif.*Biochimie, **2010**. 92(4): p. 370-377.
- 401. Xu, L., G.-L. Liao, X. Chen, C.-Y. Zhao, H. Chao, and L.-N. Ji, *Trinuclear Ru (II) polypyridyl* complexes as human telomeric quadruplex DNA stabilizers. Inorg. Chem. Commun., **2010**. 13(9): p. 1050-1053.
- 402. Sun, D., R. Zhang, F. Yuan, D. Liu, Y. Zhou, and J. Liu, *Studies on characterization, telomerase inhibitory properties and G-quadruplex binding of η 6-arene ruthenium complexes with 1, 10-phenanthroline-derived ligands.* Dalton Transactions, **2012**. 41(6): p. 1734-1741.
- 403. Chen, X., J.-H. Wu, Y.-W. Lai, R. Zhao, H. Chao, and L.-N. Ji, *Targeting telomeric G-quadruplexes with the ruthenium (II) complexes [Ru (bpy) 2 (ptpn)] 2+ and [Ru (phen) 2 (ptpn)] 2+.* Dalton Transactions, **2013**. 42(13): p. 4386-4397.
- 404. Yao, J.-L., X. Gao, W. Sun, S. Shi, and T.-M. Yao, [Ru (bpy) 2 dppz-idzo] 2+: a colorimetric molecular "light switch" and powerful stabilizer for G-quadruplex DNA. Dalton Transactions, 2013. 42(16): p. 5661-5672.
- Yu, Q., Y. Liu, J. Zhang, F. Yang, D. Sun, D. Liu, Y. Zhou, and J. Liu, *Ruthenium (II) polypyridyl complexes as G-quadruplex inducing and stabilizing ligands in telomeric DNA*. Metallomics, 2013. 5(3): p. 222-231.

- 406. Liao, G., X. Chen, J. Wu, C. Qian, H. Wang, L. Ji, and H. Chao, *Novel ruthenium (II) polypyridyl complexes as G-quadruplex stabilisers and telomerase inhibitors.* Dalton Transactions, **2014**. 43(21): p. 7811-7819.
- Xu, L., X. Chen, J. Wu, J. Wang, L. Ji, and H. Chao, *Dinuclear Ruthenium (II) Complexes That Induce and Stabilise G-Quadruplex DNA*. Chemistry-A European Journal, 2015. 21(10): p. 4008-4020.
- 408. He, L., X. Chen, Z. Meng, J. Wang, K. Tian, T. Li, and F. Shao, *Octahedral ruthenium complexes* selectively stabilize *G*-quadruplexes. Chem. Commun., **2016**. 52(52): p. 8095-8098.
- Wang, X., L. Pei, X. Fan, and S. Shi, [Ru (L) 2 (3-tppp)] 2+(L= bpy, phen) stabilizes two different forms of the human telomeric G-quadruplex DNA. Inorg. Chem. Commun., 2016. 72: p. 7-12.
- 410. Krautbauer, R., S. Fischerländer, S. Allen, and H.E. Gaub, *Mechanical fingerprints of DNA drug complexes*. Single Molecules, **2002**. 3(2-3): p. 97-103.
- 411. Tessmer, I., C. Baumann, G. Skinner, J. Molloy, J. Hoggett, S. Tendler, and S. Allen, *Mode of drug binding to DNA determined by optical tweezers force spectroscopy.* Journal of Modern Optics, 2003. 50(10): p. 1627-1636.
- 412. Sischka, A., K. Toensing, R. Eckel, S.D. Wilking, N. Sewald, R. Ros, and D. Anselmetti, *Molecular mechanisms and kinetics between DNA and DNA binding ligands.* Biophys. J., 2005. 88(1): p. 404-411.
- Vladescu, I.D., M.J. McCauley, I. Rouzina, and M.C. Williams, *Mapping the phase diagram of single DNA molecule force-induced melting in the presence of ethidium.* Phys. Rev. Lett., 2005. 95(15): p. 158102.
- 414. McCauley, M.J. and M.C. Williams, *Mechanisms of DNA binding determined in optical tweezers experiments.* Biopolymers, **2007**. 85(2): p. 154-168.
- Almaqwashi, A.A., T. Paramanathan, I. Rouzina, and M.C. Williams, *Mechanisms of small molecule–DNA interactions probed by single-molecule force spectroscopy*. Nucleic Acids Res., 2016. 44(9): p. 3971-3988.
- 416. Mihailovic, A., I. Vladescu, M. McCauley, E. Ly, M.C. Williams, E.M. Spain, and M.E. Nuñez, *Exploring the interaction of ruthenium (II) polypyridyl complexes with DNA using single-molecule techniques.* Langmuir, **2006**. 22(10): p. 4699-4709.
- 417. Almaqwashi, A.A., T. Paramanathan, P. Lincoln, I. Rouzina, F. Westerlund, and M.C. Williams, *Strong DNA deformation required for extremely slow DNA threading intercalation by a binuclear ruthenium complex*. Nucleic Acids Res., **2014**. 42(18): p. 11634-11641.
- 418. Koirala, D., S. Dhakal, B. Ashbridge, Y. Sannohe, R. Rodriguez, H. Sugiyama, S. Balasubramanian, and H. Mao, *A single-molecule platform for investigation of interactions between G-quadruplexes and small-molecule ligands.* Nat. Chem., **2011**. 3(10): p. 782-787.
- 419. Bolger, J., A. Gourdon, E. Ishow, and J.-P. Launay, *Mononuclear and Binuclear Tetrapyrido [3, 2-a: 2 ', 3 '-c: 3 '', 2 ''-h: 2 ''', 3 '''-j] phenazine (tpphz) Ruthenium and Osmium Complexes.* Inorg. Chem., **1996**. 35(10): p. 2937-2944.

- 420. Martínez-Peña, F. and A.M. Pizarro, *Unpublished results*. **2017**.
- 421. Ibarra, B., Y.R. Chemla, S. Plyasunov, S.B. Smith, J.M. Lazaro, M. Salas, and C. Bustamante, *Proofreading dynamics of a processive DNA polymerase.* The EMBO journal, **2009**. 28(18): p. 2794-2802.
- 422. Bhattacharyya, D., G.M. Arachchilage, and S. Basu, *Metal cations in G-quadruplex folding and stability*. Frontiers in chemistry, **2016**. 4.
- 423. Evans, E. and K. Ritchie, *Dynamic strength of molecular adhesion bonds.* Biophys. J., **1997**. 72(4): p. 1541-1555.