Università degli Studi di Padova

Facoltà di Scienze MM. FF. NN.

Laurea Magistrale in Biologia Sanitaria

A G-rich sequence within the 3'UTR of KiSS1 forms G-quadruplex

Relatore : *Natascia Tiso*
Dipartimento di Biologia, Padova.

Correlatore *: Nicolas de Roux*
Hôpital Robert Debré, Paris

Laureanda: GIULIA BONAFÈ

ANNO ACCADEMICO 2009 / 2010

Acknowledgements

D’abord, je tiens à remercier le Docteur Nicolas de Roux pour m’avoir accueilli dans son équipe, pour avoir insisté pour que j’apprenne le français et pour m’avoir toujours patiemment aidé surtout au début quand j’ai rencontré de nombreuses difficultés en raison de la langue.
Je remercie également toute *l’équipe Avenir*: Lucie Chevrier, Camille Vassord Dang, Laure Villoing, Carine Villanueva Bouanani, Lukas Huijbregts, Sandrine Jacquier and Fabien Guimiot qui m’ont toujours très gentiment aidé chaque fois que j’en ai eu besoin.
Merci aussi à tous les stagiaires présents au laboratoire: Eva Hernandez, Rebecca Serfati, Kevin Guillemeau et Blaise Weber.

Vorrei ringraziare di cuore anche la Dottoressa Patrizia Alberti che mi ha ospitata nel suo laboratorio (Régulation et Dynamique des Génomes, 43 rue Cuvier Paris à côté du Jardin des Plantes) per aiutarmi nella detezione degli spettri CD e la Dottoressa Natascia Tiso per il supporto e l’incoraggiamento ricevuti da Padova.

E soprattutto vorrei ringraziare la Dottoressa Silvia Baggio (la mia colloc e copine preferita) per essersi divertita un sacco insieme a me e avermi sempre aiutata in tutta quest’avventura bellissima, per avermi dato sempre la forza e il coraggio di non fermarmi mai, per aver reso quest’anno a Parigi uno dei più belli di tutta la mia vita.
E i miei genitori e la mia famiglia...per aver permesso che tutto questo accadesse.

Summary :

Riassunto

Il meccanismi di inizio della pubertà sono molto complessi e ancora poco conosciuti. La pubertà ha inizio con una secrezione pulsante di GH, rilasciato dall'ipotalamo, che porta alla secrezione di LH e FSH dalle cellule gonadotropiche dell'ipofisi, che saranno poi responsabili della sintesi e della secrezione degli ormoni sessuali da parte delle gonadi. Questa cascata ormonale che porta allo sviluppo dei caratteri sessuali secondari e che coinvolge l'ipotalamo, l'ipofisi e le gonadi costituisce l'asse gonadotropico che è attivato per la prima volta alla sedicesima settimana della vita fetale, viene inattivato dopo la nascita e ritorna attivo al momento della pubertà per svolgere le funzioni della riproduzione.
Nel 2003, le Kisspeptine e il loro recettore GPR54 sono stati individuati come elementi chiave nella regolazione dell'asse gonadotropico e nella sua riattivazione al momento della pubertà.

Studi genetici antecedenti hanno trovato tre polimorfismi nella regione, ricca in guanine, del 3'UTR del gene KiSS1, nei pazienti affetti da Pubertà Precoce Centrale (CPP). L'analisi di questi polimorfismi ha rilevato l'esistenza di differenze funzionali nette. Analisi *in silico* hanno inoltre dimostrato che la sequenza ricca in guanine del 3'UTR del gene KiSS1, che codifica per le Kisspeptine, puo assumere conformazioni particolari chiamate G-quadruplex. Queste strutture secondarie sono state recentemente descritte come regolatrici dell'espressione genica attraverso diversi meccanismi molecolari.

L'obiettivo del mio tirocinio è stato quello di dimostrare che la regione ricca in guanine del 3'UTR del gene KiSS1 forma effettivamente strutture secondarie G-quadruplex e di capire se i polimorfismi trovati nei pazienti affetti da pubertà precoce interferiscono con la formazione di queste strutture e quindi con la regolazione del gene. È stato anche iniziato uno studio preliminare sulle potenziali interazioni tra proteine e i G-quadruplex di KiSS1.

Per fare questo è stata innanzitutto svolta un'analisi tramite FRET con l'oligonucleotide F22T, che rappresenta la regione ricca in guanine del gene in esame, marcato alle due estremità con il cromoforo FAM e il quencher TAMRA, per capire se avviene effettivamente la formazione di strutture secondarie particolari. Questo è stato dimostrato in quanto si è potuto assistere a un progressivo aumento della fluorescenza di FAM con le temperatura, segno della presenza iniziale di una strutture secondaria: se l'oligonucleotide viene denaturato, le sue due estremità si allontanano e l'effetto silenziatore di TAMRA su FAM diminuisce progressivamente permettendo a FAM di emettere la propria fluorescenza.

In secondo luogo, per capire se queste strutture secondarie osservate siano veramente delle strutture G-quadruplex, è stato fatto il Circular Dichroism, che fornisce grafici

specifici per ogni tipo di struttura che l'oligonucleotide analizzato assume. Siamo cosi riusciti a dimostrare che la regione d'interesse assume effettivamente una conformazione G-quadruplex perchè il grafico ottenuto è un grafico tipico di strutture G-quadruplex. Sono stati analizzati anche gli oligonucleotidi contenenti i polimorfismi (H2, H3, H4) e la cosa interessante è stata osservare la formazione di strutture G-quadruplex anche in questi casi, ma aventi conformazioni diverse rispetto all'aplotipo presente nella popolazione normale (H1). A seconda della direzione dei filamenti del DNA che formano la struttura G-quadruplex, questa può essere descritta come parallela, antiparallela o mista. H1 assume una conformazione antiparallela mentre H2, H3 e H4 assumono una conformazione mista. È stato dimostrato per la prima volta che la regione del 3'UTR di KiSS1 forma strutture G-quadruplex e che i polimorfismi trovati nei pazienti con pubertà precoce interferiscono con la formazione di queste strutture particolari; questo potrebbe contribuire al mal funzionamento del gene nei soggetti malati.

Infine, è stata dimostrata l'interazione di queste strutture del gene KiSS1 con proteine tramite FRET. È stata fatta l'analisi con l'oligonucleotide F22T in presenza e in assenza dell'estratto proteico della linea cellulare MDA MB 231, che esprime in maniera endogena KiSS1, ed è stata confrontata la cinetica delle due reazioni (attraverso lo studio della curva della prima derivata della fluorescenza di FAM). Sono risultate evidenti molte differenze dovute alla presenza di proteine, che sembrano avere un effetto di destabilizzazione sulle strutture G-quadruplex di KiSS1.

Abbiamo quindi dimostrato che la regione ricca in guanine del 3'UTR di KiSS1 si struttura in G-quadruplex e che i polimorfismi trovati nei pazienti che sviluppano pubertà precoce interferiscono con la formazione di queste strutture secondarie particolari, facendo loro assumere una conformazione differente. Infine, abbiamo dimostrato che i G-quadruplex di KiSS1 interagiscono in modo specifico con le proteine nucleari.
Questo studio si è focalizzato sulla sequenza di DNA. La capacità del DNA di formare G-quadruplex sembra essere coinvolta in diversi meccanismi molecolari come la regolazione genica. Molti G-quadruplex sono infatti stati trovati nelle regioni promotrici dei geni mentre il ruolo di queste strutture in altre parti del gene, come sequenze codificanti o regioni UTR, è ancora poco chiaro. Possiamo ipotizzare che i G-quadruplex nelle sequenze esoniche o introniche interferiscano con l'attività dell'RNA polimerasi.
Il G-quadruplex trovato in KiSS1 può anche agire come regolatore post-trascrizionale dell'espressione del gene attraverso diversi meccanismi molecolari e sarà quindi necessario eseguire le stesse analisi anche sull'RNA. Infine, questa analisi strutturale dovrebbe essere affiancata anche da un'analisi funzionale per confermare che questo G-quadruplex può effettivamente diventare un nuovo bersaglio farmacologico per regolare l'espressione del gene KiSS1, ad esempio nei pazienti affetti da pubertà precoce.

Abstract

The mechanisms of pubertal initiation are very complex and not still fully understood. It depends on the increase in pulsatile secretion of GnHR by the hypothalamus leading to the release of LH (Luteinizing Hormone) and FSH (Follicle Stimulating Hormone) by the gonadotropic cells of pituitary and they are responsible for the synthesis of sex steroids by the gonads. This endocrine axis defines the gonadotropic axis.

In 2003, the Kisspeptins/GPR54 couple was identified as a key element in the gonadotropic axis regulation and in its reactivation at puberty. In fact, mutations in KiSS1R gene, which encodes Kisspeptins Receptor, cause the Isolated Hypogonadotropic Hypogonadism (IHH), a genetic disease which leads to an absence of the gonadotropic axis activation at puberty.

Previous genetic studies, performed in the *Avenir Team*, have found three polymorphisms in a guanine-rich (G-rich) region of the 3'UTR of KiSS1, in patients with a Central Precocious Puberty (CPP). Functional analysis of these polymorphisms have reveled net functional differences using a reporter gene.

In silico analysis revealed that the 3'UTR G-rich sequence, in which polymorphisms were found, may be fold in G-quadruplex structure. This secondary structure was recently described as potentially involved in gene expression regulation by different molecular mechanisms.

The task of this project was to demonstrate that this G-rich region effectively forms G-quadruplex structures and to study if these polymorphisms change the G-quadruplex conformation. We have also tried to show that this G-rich sequence binds proteins.

Our results showed that the 3'UTR region of KiSS1 indeed folds into antiparallele and mixed G-quadruplexes. We also observed that polymorphisms of this sequence change G-quadruplex conformations. Our results also indicate that G-quadruplex of KiSS1 is able to specifically bind proteins.

This study, although preliminary, opens up new possibilities in order to better understand the regulation of gonadotropic axis by the Kisspeptins/GPR54 system and the role of KiSS1 in the onset of puberty.

Introduction

The *Avenir Team* works on the genetics and physiology of the initiation of puberty. Puberty is a complex mechanism resulting in pulsatile secretion of GnRH by the hypothalamus and subsequent synthesis of LH and FSH. In 2003 the *Avenir Team* discovered the essential role of Kisspeptins and their receptor GPR54 in the pubertal initiation.

Gonadotrope axis and puberty

The gonadotrope axis is composed of three major structures: the hypothalamus, the pituitary and the gonads.

The hypothalamus is the site of the synthesis of GnRH (Gonadotropin Releasing Hormone) which represents an important step in the reproductive cascade and which induces and controls the synthesis of LH (Luteinizing Hormone) and FSH (Follicular Stimulating Hormone) by pituitary. LH and FSH, through their respective receptors, LH-R and FSH-R, led to the maturation of the gonads and stimulate the secretion of sex steroids (estrogen and testosterone). Sex steroids exert negative feedback effect at both the hypothalamus and pituitary with the exception of estrogen that undergoes positive feedback at the time of the mid-cycle ovulatory surge.

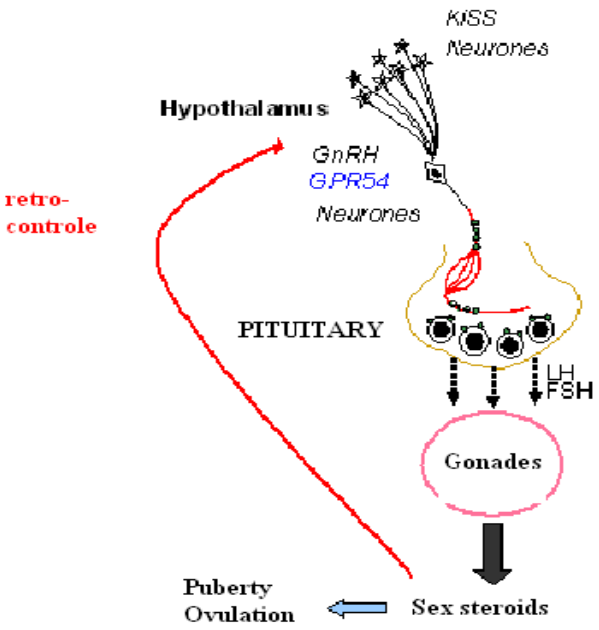

Fig.1: hypothalamic-pituitary gonadal axis

The gonadotrope axis is activated for the first time at the 16th weeks of fetal life, it becomes dormant after birth and it is reactivated at the time of puberty. It will remain active during the entire adult life to ensure reproductive function.

Puberty is a complex biological process involving the development of secondary sexual characteristics, accelerated linear growth and adrenal maturation. The activation of pulsatile hypothalamic secretion of GnRH is a key event in the onset of puberty. The study of genetic diseases of puberty at onset, such as Hypogonadotropic Hypogonadism, allowed us to discover that the Kisspeptins/GPR54 system plays an important role in pubertal initiation.

Central Precocious Puberty (CPP)

Precocious Puberty is the medical term for puberty occurring at an unusually early age. The onset of puberty, governed by the complex process of hypothalamic-pituitary-gonadal axis, involves: the release of GnRH by hypothalamus, the pituitary stimulation by GnRH to release LH and FSH, the production of sex hormones (estrogen and testosterone) by ovaries and testicles stimulated by LH and FSH and the physical changes of puberty caused by sex hormones (15). It is a complex genetic disease with familial aggregation in which the entire gonadotrope axis starts too soon, but although all the steps begin earlier than they should, their pattern and timing are otherwise normal in the process. If the cause can be traced to the hypothalamus or pituitary, it is considered central.

Kisspeptins

GPR54 is a G protein coupled receptor isolated in 1999, highly expressed in the brain (particularly in the hypothalamus), in the placenta, pituitary, spinal cord, pancreas, heart, muscle, testis and other organs in which its role has not yet been clarified.

GPR54 ligands are called Kisspeptins and they are the products of KiSS1 gene which was originally identified as a human metastasis suppressor gene in 1996 by subtractive cloning between melanoma cells having a strong metastatic ability and those same cells having lost this metastatic ability after transfection of human chromosome 6 (6). KiSS1 gene maps to chromosome 1, suggesting that the element causing inhibition of metastasis on chromosome 6 may be an important regulator of Kisspeptins (32).

The KiSS1 gene encodes a 145 amino acid residue peptide that is further processed to a final peptide, metastin. It has been identified as a putative metastasis suppressor gene in human melanoma and breast cancer cell lines, in fact there is an important link between expression levels of KiSS1 and patients prognosis. Its levels are higher in tumor compared to the normal tissues and they significantly increase in invasive tumors. Over expression of Kiss1 is particularly evident in patients with aggressive breast cancer and with mortality (34).

The normal timing of puberty is genetically determined but there are also many environmental factors involved. KiSS1 expression increases at the end of childhood and to understand why this happen it was decided to look for mutations in patients with a precocious puberty (CPP) and to extend this analysis to the 3'UTR region of KiSS1 gene, which is involved in the post-transcriptional regulation of the gene. Several polymorphisms are been found (***Fig.2***) and some of these are more frequent in patients with CPP rather than in the normal population. They are localized in the G-rich region of the 3'UTR of the gene, which could form particular secondary structures called G-quadruplex involved in post-transcriptional regulation of gene expression. The functional analysis of these polymorphisms showed net differences using a reporter gene (*unpublished data, Nicolas De Roux et al*.). These polymorphisms were therefore suspected to interfere with the capacity of this region to form G-quadruplex.

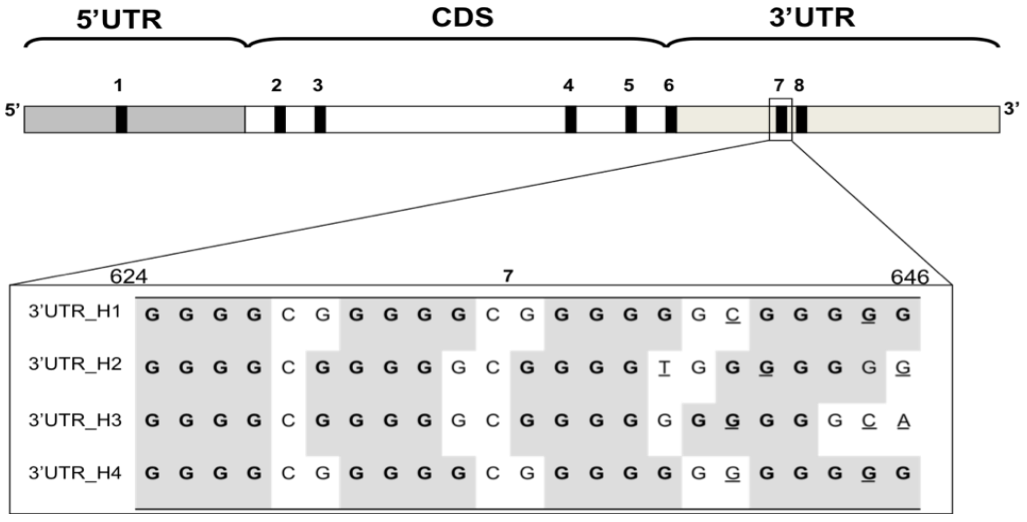

Fig.2: localization of G-quadruplex in KiSS1 mRNA. They are indicated by closed boxes and numbered 1 to 8. G-repeats involved in the planer organization of G-quadruplexes are in bold and highlighted in gray. (unpublished data, Nicolas De Roux et al.)

G-Quadruplex

Elsewhere in the genome there are guanine-rich single-stranded DNA sequences which have the ability to fold into novel four-stranded structures, called G-Quadruplex (1). Enrichment for sequences with intra molecular G-Quadruplex forming potential (QFP) have been found in transcriptional promoters of humans (3), in the first intron of human genes (23), in human telomeres (2), in the Fragil X Syndrome repeat $d(CGG)_n$ (1), in oncogenes and tumor suppressor genes, in ribosomal DNA, minisatellites, immunoglobulin (3) and in some genes involved in metabolism. The formation of these structures would require the canonical Watson-Crick base-pairing to be broken and they consist of a square arrangement of G (a tetrad) stabilized by Hoogsteen hydrogen bonding (3) and of a formation of a multi-storey. The opportunity of this arises during DNA replication, transcription and recombination, when regions of double-stranded DNA become transiently single-stranded (2).

G-quadruplexes are generally folded into one of three principal topologies: parallel, antiparallel or mixed (***Fig.3b***). In the parallel topology, all guanine nucleotides adopt the *anti* conformation in the glycosidic bond, while on the other hand, in the antiparallel and mixed topologies, half of the guanine nucleotides possess the *anti* conformation and the other half possesses the *syn* conformation (36).

The formation of G-quadruplex is strongly dependent on monovalent cations concentrations such as K^+, Na^+, NH_4^+ (2); monovalent ion sits in a channel at the centre of the G-quadruplex (in the centre of the tetrads) (***Fig.3a***).

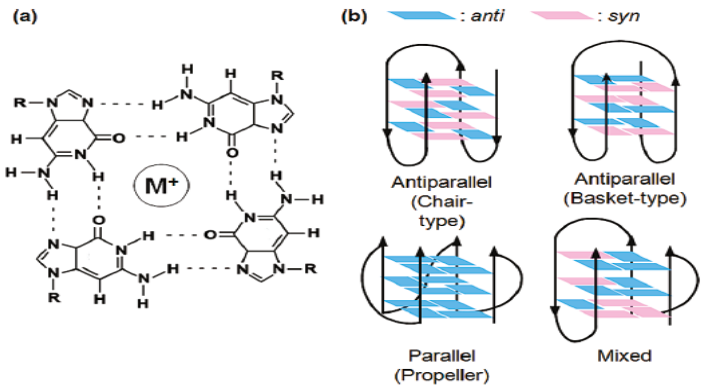

Fig.3: (a) chemical structure of a G-quartet observed in a chair-type antiparallel G-quadruplex. A monovalent cation (M^+) occupies the central position. (b) Schematic representation of different structures of the intramolecular chair-type and basket-type antiparallel, parallel and mixed G-quadriplexes (reference 38).

G-quadruplex formation might influence regulation of genes expression (***Fig.4***) and have an important role in transcriptional regulation (3); for example, it has been demonstrated that the formation of G-quadruplex structures in genes promoters can inhibit transcription whereas in telomeres they suppress telomerase activity and so telomeres extension. This is an attractive idea for the design of new anti-cancer agents (for example ligands which bind and stabilize G-quadruplex) because telomerase is an enzyme which is up-regulated in ~85% of human cancer.

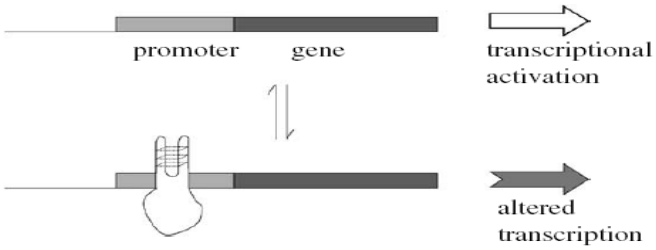

Fig.4: model for the effect of a promoter G-quadruplex (reference 27)

G-quadruplex can be formed by DNA or RNA sequences. Although many studies have dealt with the formation and consequences of DNA quadruplexes, little is known about four-stranded RNA quadruplexes *in vivo.* RNA folds into equally stable quadruplex structures, preferring the all-parallel orientation of the participating strands (26). Contrary to duplex DNA, cellular RNA is generated as single stand and it favors the formation of non-duplex secondary structures, as the G-quadruplexes are, because of the absence of the competition with duplex formation. The formation of these structures can interfere with translation, binding of specific proteins, splicing (***Fig.5***).

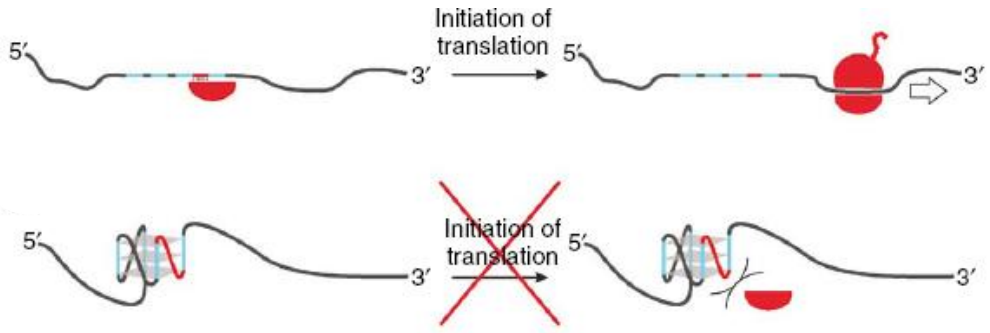

Fig.5: in order to detect quadruplex-formation in mRNA sequences, a potential quadruplex forming sequence was positioned at a crucial site for translational initiation (in red). If this site is incorporated into a loop of the four-stranded structure, it becomes inaccessible to ribosomal subunits and translation is inhibited (reference 26).

The aim of the work

The task of my work was to demonstrate that the 3'UTR region of KiSS1 folds into G-quadruplex structure and to show that the polymorphisms found in CPP patients behave differently from the haplotype presents in the most part of the population and, to do this, we have done FRET analysis and calculated CD spectra. We have shown for the first time that this region folds into G-quadruplex structure and that G-quadruplex conformation changes between the polymorphisms. We have also tried to demonstrate that KiSS1 G-quadruplex is bound by proteins. We have used the MDA MB 231 cells line protein extract, which endogenously expresses KiSS1 and done FRET analysis and we have observed that some proteins interact with G-quadruplex of KiSS1 to destabilize it.

Materials and methods

Oligonucleotides

The fluorescent oligonucleotides

We have ordered two different G-rich oligonucleotides from Eurogentec (Belgium) to perform FRET analysis:

- A typical human telomeric sequence: F21T= 5'-FAM-GGGTTAGGGTTAGGGTTAGGG-TAMRA-3'
- The KiSS1 3'UTR region: F22T= 5'-FAM-GGGGCGGGGGCGGGGGGCGGGG-TAMRA-3'

Oligonucleotides with polymorphisms of Kiss1

Four different oligo G-rich synthesized and purified by Eurogentec (Belgium) were ordered, H1 is the haplotype of 3'UTR region of Kiss1 which is the most frequent in the normal population, H2, H3 and H4 are the 3'UTR regions with the most frequent polymorphisms in patients with a central precocious puberty (CPP):

H1 = 5'- GGGG CG GGGG CG GGGG GC GGGG A -3'

H2 = 5'- GGGG C GGGG GC GGGG AG GGGGG G -3'

H3 = 5'- GGGG C GGGG GC GGGG G GGGG GCA -3'

H4 = 5'- GGGG CG GGGG CG GGGG G GGGGG A -3'

FRET analysis

FRET (Fluorescence Resonance Energy Transfer) is a methods describing energy transfer between two chromophores and it can be used to determine the proximity between two molecules (for example the receptor and its ligand, protein-protein interactions, protein-DNA interactions) and it is a useful tool to quantify molecular dynamics in biophysics and biochemistry and to follow protein conformational changes. It also can be used for monitoring the complex formation between two molecules. A donor chromophore, initially in its excited state, may transfer energy to an acceptor chromophore (in proximity, typically less than 10 nm) through non-radiative dipole-dipole coupling. When they are dissociated, the donor emission is detected upon the donor excitation whereas when they are in proximity due to interactions of the two

molecules, the acceptor emission is predominantly observed because of the intermolecular FRET between the donor and the acceptor.

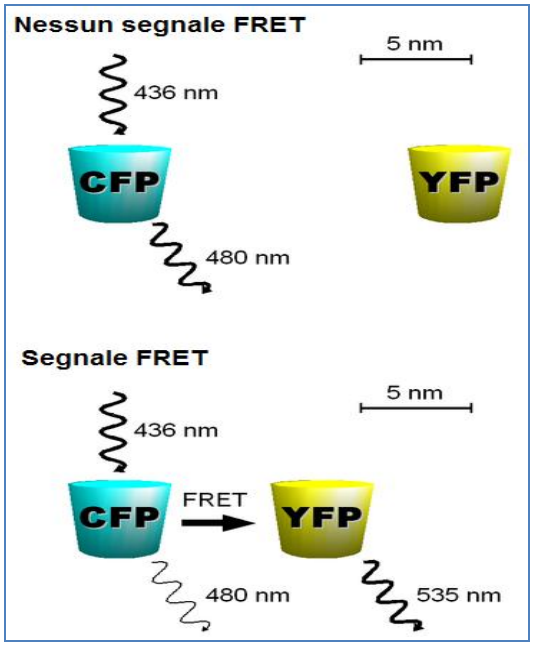

Fig.6: example of FRET between CFP and YFP (Wavelength vs. Absorption): a fusion protein containing CFP and YFP excited at 440 nm. The fluorescent emission peak of CFP overlaps the excitation peak of YFP. Because the two proteins are adjacent to each other, the energy transfer is significant and a large proportion of the energy from CFP is transferred to YFP and creates a much larger YFP emission peak.

FRET analysis can be also used to discover if an oligo can assume a particular secondary structure (***Fig.7***):

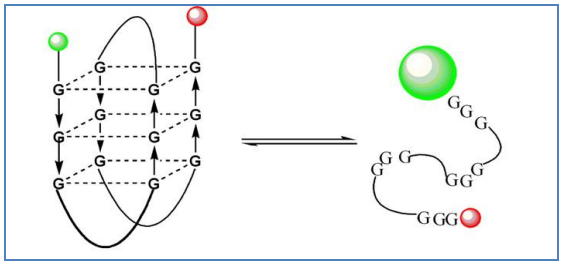

Fig.7: schematic representation of the melting of a quadruplex-forming oligonucleotide that has been labeled with a fluorophore (green) and a quencher (red) (reference 28).

The basis of FRET assays is *quenching*, that refers to any process which decreases the fluorescence intensity of a given substance. A *quencher* is a substance that absorbs excitation energy from a fluorophore and re-emits much of this energy as light. FRET is a dynamic quenching mechanism because energy transfer occurs while the donor is in the excited state. FRET is based on classical dipole-dipole interactions between the transition dipoles of the donor and acceptor and is extremely dependent on the donor – acceptor distance. FRET also depends on the donor-acceptor spectral overlap and the relative orientation of the donor and acceptor transition dipole moments. One of the most common quencher is TAMRA (Abs 565 nm, Em 580 nm). This modification is

used to quench the emission of a reporter dye and it is recommended as a quencher for FAM.

Buffer: choice of the monocation

Most real-time PCR machines work in a temperature range between 25 and 95°C so one needs a F21T or F22T quadruplex with a melting temperature between 45 and 55°C. My experiments were performed in a total monocation concentration of 100 mM and in a sodium cacodylate concentration of 10 mM, pH 7.2 buffer. Taking into account that the folding is intramolecular (meaning that the T_m in concentration-independent) and that the stability of the quadruplex depends on the nature and concentration of the cation, I selected 3 different buffer conditions:

i. 10-100 mM KCl/NaCl/NH_4Cl
ii. 90-0 mM LiCl
iii. 10 mM NaCacodylate
iv. 0,2 µM F21T or F22T

We have done the same experiment at 10 mM of KCl/NaCl/NH_4Cl, changing the concentration of DNA:

i. 10 mM KCl/NaCl/NH_4Cl
ii. 90 mM LiCl
iii. 10 mM NaCacodylate
iv. 0,1/0,4/0,5 µM F22T

At last we have studied the formation of G4 in the presence of unlabelled DNA (the same sample of F22T): H1= 5'-GGGGCGGGGGCGGGGGGCGGGGA-3' (KiSS1 3'UTR region):

i. 10 mM KCl/NaCl/NH_4Cl
ii. 90 mM LiCl
iii. 10 mM NaCacodylate
iv. 0,1 µM F22T
v. 0/0,1/0,5/1/2/5 µM n1

Experimental setting

All experiments are performed onto 96 wells; each condition is tested at least in triplicate in a volume of 20 µl for each sample. The FAM tag is excited at 492 nm and the emission is collected at 516 nm. A PCR machine (7300 Real Time PCR System) was used only to exploit its temperature cycles, not for amplification. In order to study G-quadruplex formation a short protocol was used:

- Denaturation of the oligonucleotide: 85°C for 10 min
- Overnight incubation at RT to refold G-quadruplex structures
- FRET analysis:
    - Equilibration at low temperature (1 min at 25°C + 1 min at 26°C)
    - Stepwise increase of 1°C every min for 70 cycles to reach 96°C
    - The recording is performed every minute between 30°C to 96°C

Determination of T_m

Manual baseline determination, required for true T_m determination, is actually impractical because Real-time PCR machines generate a lot of data and only two feasible automatizable alternatives are possible:

i. *$T_{1/2}$ determination.* The emission of fluorescein is normalized between 0 and 1, and the $T_{1/2}$ is defined as the temperature for which the normalized emission is 0.5.

ii. *First derivative analysis.* The T_{max} is defined as the maximum of the first derivative of the fluorescence *vs.* temperature curve.

Circular Dichroism (CD)

Circular dichroism (CD) spectroscopy measures differences in the absorption of left-handed polarized light versus right-handed polarized light that arises due to structural asymmetry. The absence of regular structure results in zero CD intensity, while an ordered structure results in a spectrum that can contain both positive and negative signals. CD has a wide range of applications in many different fields such as the investigation of secondary or tertiary structure of proteins and the structural family to which they belong, the comparison of the structures of a protein obtained from different source or from different mutants of the same protein, the study of the conformational stability of a protein under stress or how this stability is altered by buffer composition.

UV-melting experiments

The thermal stability of different oligonucleotides (H1, H2, H3 and H4) was characterized in heating/cooling experiments by recording the UV absorbance at 295 nm as a function of temperature using a Kontron-Uvikon 940 UV/Vis spectrophotometer using 1-cm path length quartz cuvettes. UV-melting experiments were conducted in 100 mM NaCl and KCl buffer. DNA concentration was 3 μM.

Thermal Difference Spectra (TDS)

The thermal difference spectra (TDS) of a nucleic acid were obtained by recording the UV absorbance spectra over a range of temperatures (2°-90°C) and subsequently taking the difference between each spectrum and that at 90°C. TDS provide specific signatures of different DNA structural conformations. Spectra were recorded between 220 and 320 nm on a Kontron-Uvikon 940 UV/Vis spectrophotometer using quartz cuvettes with an optical path length of 1 cm. DNA concentration of H1, H2, H3 and H4 was 3 μM.

Protocols for: Thermal difference spectra, UV-melting curves and circular dichroism

UV-absorbance measurements were acquired on a Kontron Uvikon 940 spectrophotometer. For melting experiments, samples were heated at 92°C for a few minutes, cooled from 92°C to 2°, kept at 2°C for 30 minutes and heated from 2°C to 92°C at a rate of 0.2°C min^{-1}. The temperature was varied with a circulating water bath; evaporation at high temperature and condensation at low temperature were prevented by a layer of mineral oil and by a dry air flow in the sample compartment, respectively. The absorbance was monitored at 245, 260, 273, 295 and 405 nm; the absorbance at 405 nm was subtracted from each wavelength. For each melting curve, the melting temperature (T_m) was graphically determined as the intercept between the melting curve and the median line between low-temperature and high-temperature absorbance linear baselines. For each sample, the thermal difference spectrum (TDS) was obtained by subtracting the absorbance spectrum at 2°C from the one at 92°C. The spectrum at 92°C was recorded after heating the sample at 92°C for a few minutes, whereas the spectrum at 2°C was recorded after annealing from 92°C to 2°C at a rate of 0.2°C min^{-1}. Circular

dichroism (CD) spectra were recorded on a Jasco J-810 spectropolarimeter at 4, 60, 80 and 100°C, after annealing from 92°C at a rate of 0.2°C min^{-1}. Each spectrum was obtained by averaging three scans, at a speed of 500 nm min^{-1}, and by subtracting the contributions of the solvent and the quartz cell. All experiments (melting, TDS and CD) were carried out at 3 oligonucleotide strand concentration, in a cacodylate buffer (10 mM), pH 7.2 (adjusted with LiOH), containing KCl or NaCl (100 mM), using 1 cm (for 3 and 10 μM strand concentration) or 0.2 cm (for 30 μM strand concentration) path-length quartz cells.

Gel electrophoresis

Sample preparation

A solution of oligo (H1, H2, H3 and H4) 30 μM in a 10 mM Cacodylate and 100 mM KCl/NaCl buffer has been prepared:

- Denaturing of the oligo: 90°C for 2 min
- Rapid centrifugation of the sample
- Incubation at RT for 2h
- After annealing, sucrose to final concentration of 11-12% (10 μl of oligo solution + 3 μl of sucrose 50%) has been added

Gel preparation

- Acrylamide/Bisacrylamide gel (19:1) 12% with 20 mM KCl/NaCl
- Plates: 16,5 cm x 14,5 cm
- Spacer: 1 mm
- Comb: 16 or 20 wells

Running buffer

- TBE 1X
- KCl/NaCl 20 mM

Gel run

- In cold room
- Premigration of the gel at 3W for 30 min

- Depositing of 13 µl of the oligo/sucrose solution
- Migration at 3W for 2 h

Detection

- UV-shadow: DNA was detected by UV-shadow at 254 nm with a G:BOX (Syngene).
- The human telomeric sequence GGGTTAGGGTTAGGGTTAGGG, known to fold into intramolecular G-quadruplexes both in sodium and in potassium, was loaded into the gels to provide an internal migration marker

Cell culture

The MDA MB 231 human breast carcinoma cell line was obtained from a patient in 1937 at M. D. Anderson Cancer Centre (Houston, TX). It's an invasive cell line isolated from the pleural effusions of patients with breast cancer (35) and it has been showed that these cells express very high levels of KiSS1 (34): KiSS1 expression is not reduced in human breast cancer tissues and conversely it's associated with poor prognosis in breast cancer patients and it behaves as a facilitator of invasive and motile behavior in human breast cancer cells.

The cells were maintained in monolayer culture in Dulbecco's Modified Eagle Medium 1X 4,5g/L of glucose (DMEM+GlutaMAX^{TM}-1) supplemented with 10% of fetal calf serum (FCS), 1% of L-glutamine, 1% of penicillin-streptomycin and 25 mM Hepes; they were placed in the incubator at 37°C with 5% of CO_2 and they were sub-cultured when they reached 100% of confluence.

Preparation of nuclear extracts

This protocol is intended for nuclear proteins extraction of approximately 10^7 cells, in a box of 100 mm of diameter, when the cells are confluent. The expected concentration is 2-3 µg/µl of proteins in a final volume of 100 µl.

After having washed the cells with PBS 1X, they are centrifuged for 5 min at 4°C, 120g (1000 rpm) and the pellet is resuspened in 400 µl of buffer A. after a second centrifugation (20 sec), the pellet is resuspended in 100 µl of buffer B and incubated for

20 min at 4°C and at least quickly centrifuged. The supernatant is aliquoted, the proteins concentration is determined and the tubes are stored at –80°C until use.

Buffer A

Products	**Final concentration**
Hepes KOH pH 7,9	25 mM
$MgCl_2$	25 mM
KCl	10 mM
DTT	0,5 mM
PMSF	0,2 mM

Buffer B (lysis buffer)

Products	**Final concentration**
Hepes KOH pH 7,9	25 mM
Glycerol	25%
$MgCl_2$	1,5 mM
NaCl	420 mM
DTT	0,5 mM
PMSF	0,2 mM

Results

1) Evidences that the G-rich sequence of KiSS1 folds into G-quadruplex structure

FRET analysis

At the beginning we have tested FRET analysis with the oligonucleotide F21T, a sequence of a typical human telomere, because it is known that the telomeres are rich in G and are usual to form G-quadruplex structures *in vitro* (10). We used different concentrations (from 10 mM to 100 mM) of different cations (KCl, NaCl, NH_4Cl) because it is known that cations affect the formation of G-quadruplex structures (7). To test the functionality of FRET we have also done the experiment with different F21T concentrations (0,2/0,4/0,8 µM) and different volumes (15 and 20 µl) in order to choose the best experimental conditions (***Fig.8***).

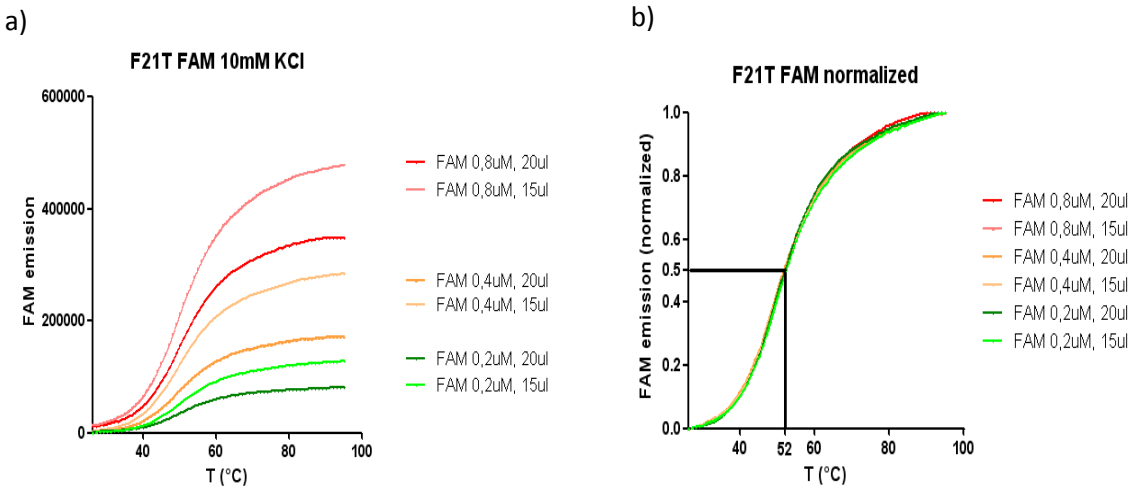

Fig.8: FRET analysis at different F21T concentrations (0,2/0,4/0,8 µM) and different volumes (15 and 20 µl). a) Raw data. b) Normalized melting curves.

F21T is an oligonucleotide marked with a chromophore, FAM, and an acceptor chromophore, TAMRA. We observed FRET when, passing from low temperature to high temperature, the FAM fluorescence increases whereas TAMRA fluorescence decreases (***Fig.9.***). This is due to DNA denaturation; when the oligonucleotide loses its secondary structure, its 5' and 3' ends are moved away from each other, TAMRA can no longer quench FAM and so it starts emitting fluorescence (FAM).

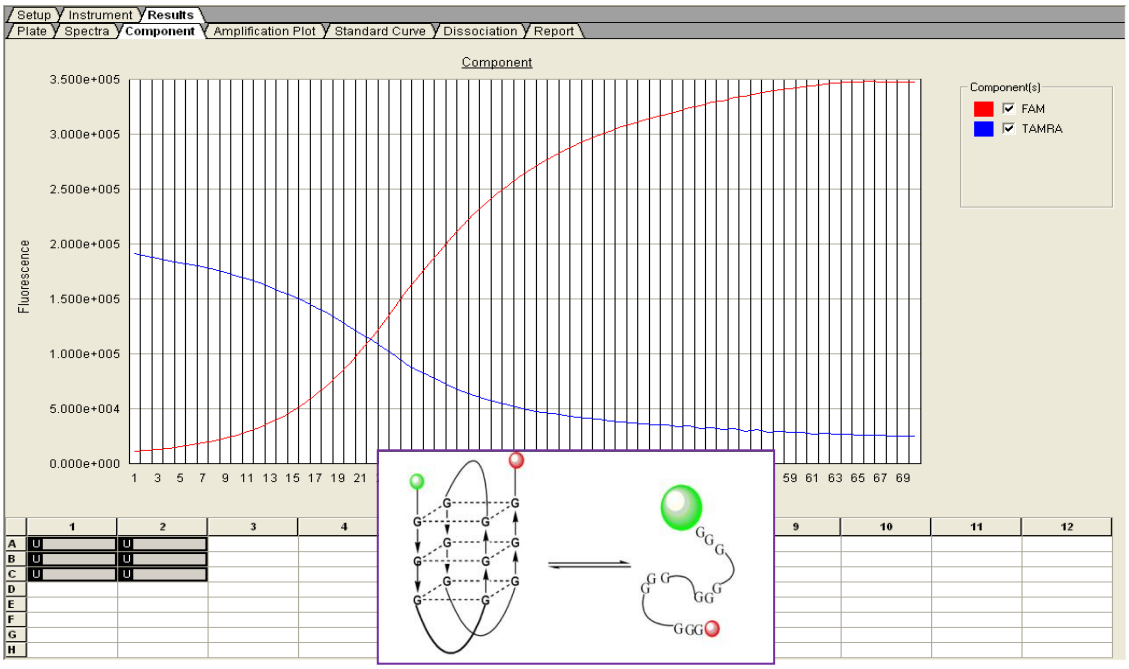

Fig.9: a typical out put of FRET analysis by 7300 RealTime PCR System

Using temperature cycles of PCR machine, we have observed FRET, and so the formation of a particular secondary structures, in the presence of all three cations but we have noticed that the structures formed are different because the curves detected are different from each other. On the contrary, changing the oligonucleotide concentration and the final volume, results do not change. The same experiments have been made with F22T, which is the 3'UTR G-rich region of KiSS1 gene, and it is marked with FAM and TAMRA too. Similar results have been obtained: there is a different behavior in the presence of different cations (***Fig.10***).

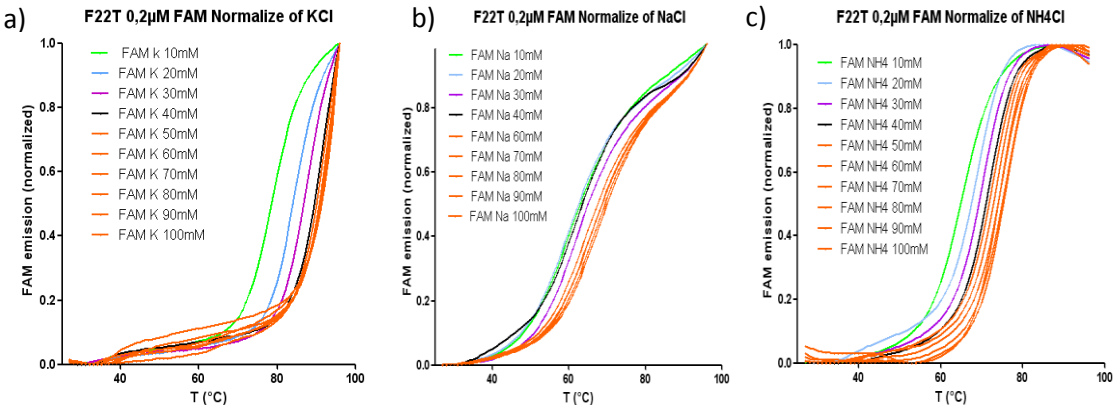

Fig.10: FAM normalized curves at different cations concentrations (10-100 mM), in the presence of K^+ (a), Na^+ (b) and NH_4^+ (c).

We have also calculated the melting temperature (T_m) in the presence of K^+, Na^+ and NH_4^+ and we have seen that the T_m is directly proportional to the cations concentration (***Fig.11***) and this means that the secondary structures found are stabilized by an increase in K^+, Na^+ or NH_4^+ concentration: the oligonucleotide denaturation becomes more and more difficult with the increasing of cations concentration and it also causes a rise in the T_m. This could be due to the tetraplexes formation and these seem to be particularly stable in the present of K^+.

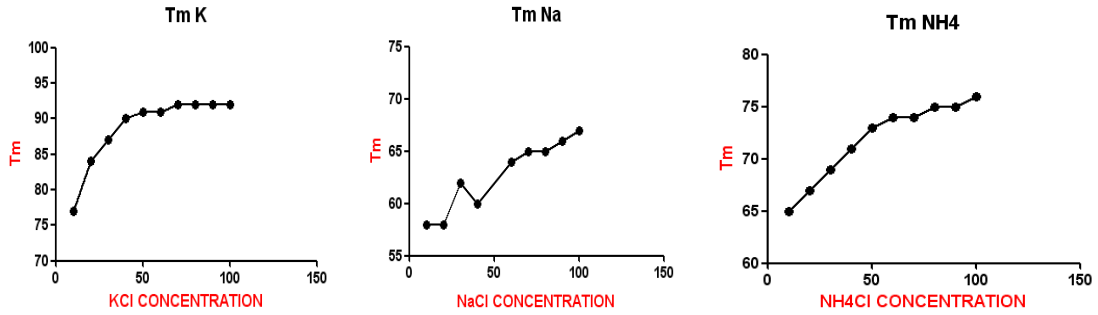

Fig.11: curves of T_m as a function of cations concentration

We have also tested the formation of these particular secondary structures at different DNA concentrations maintaining a constant cation concentration of 10 mM: FRET was observed for each DNA concentration. T_m did not significantly change, but it undergoes small changes and it means that DNA concentration slightly affects the formation of G-quadruplex and probably a minority of molecules was as intermolecular structures.

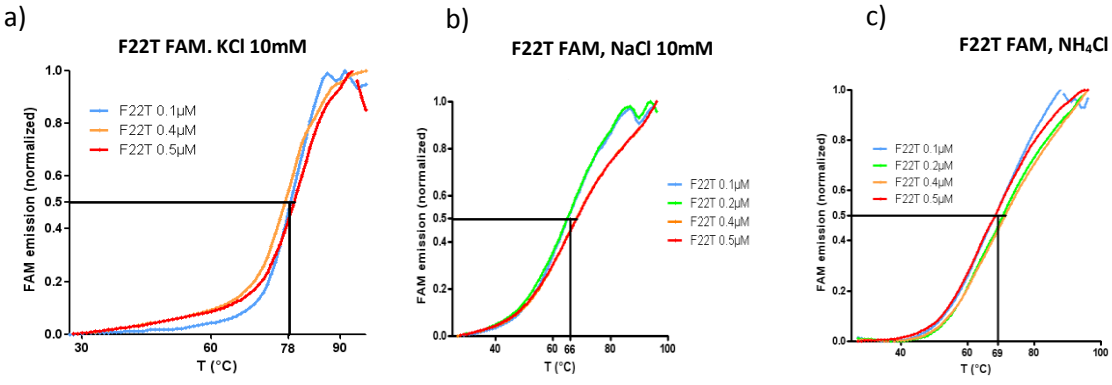

Fig.12: FAM normalized curves in the presence of different concentration of F22T. a) In the presence of K^+. b) In the presence Na^+. c) In the presence of NH_4^+.

To confirm this hypothesis we have repeated the experiment by adding different concentrations of unlabeled DNA of the same sequence of F22T: we have prepared the following mixture and perform FRET analysis:

- 10 mM of KCl or NaCl or NH_4Cl
- 90 mM of LiCl
- 10 mM of Cacodylate
- 0,1 µM of F22T
- 0/0,1/0,5/1/2/5 µM of H1 (which has the same sequence of F22T but it is not marked)

This experiment should confirm or not the formation of intramolecular structures: if we detect FRET, it means that the structures observed before are intramolecular, if instead we dose not detect FRET, it means that we observed intermolecular structures and that there is a competition between the marked oligonucleotide and the unlabeled one to form these secondary structures. We have observed FRET in the presence of K^+ and Na^+ but in the presence of NH_4^+ it is not so clear. The T_m slightly varies by increasing H1 concentration and it means that a minority of intermolecular structures is folded.

a) b) c)

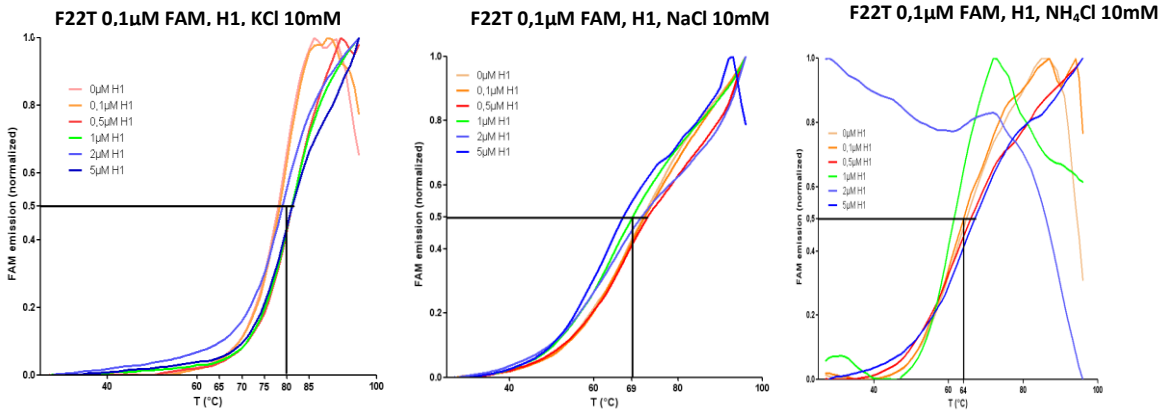

Fig13: normalized melting curves of F22T in the presence of 10 mM of K^+ (a), Na^+ (b) and NH_4^+ (c) using different H1 concentration (0,1/0,5/1/2/5 µM).

Gel electrophoresis

To have a further confirmation that the structures observed with the previous experiment, in the presence of KCl and NaCl, was mostly intramolecular, a gel electrophoresis was done, in a KCl buffer and in NaCl buffer too.

Migration of intermolecular structures is slower than migration of intramolecular structures. We have obtained different results in the presence of NaCl and KCl (***Fig.14***):

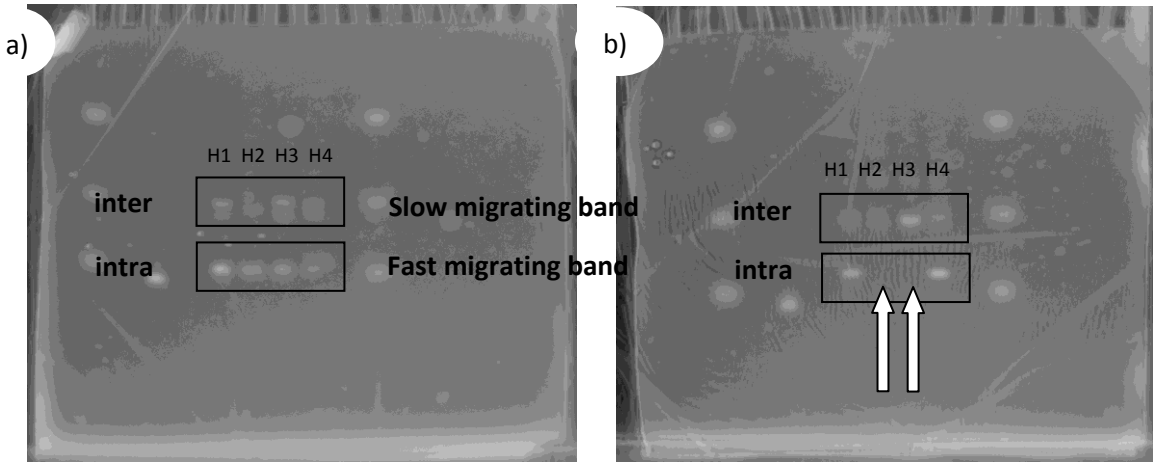

Fig.14: acrylamide/bisacrylamide (19:1) gel electrophoresis. a) In the presence of NaCl. b) In the presence of KCl

The fast migrating band has mobility comparable to that of human telomeric sequence that is known to fold into intramolecular G-quadruplex both in Na^+ and in K^+, while the slow migration band, composed by two very close bands, has a lower mobility than that of the human telomeric sequence. This slower band probably corresponds to intermolecular structures migration, but we cannot be totally sure of this: it could also be another intramolecular structure.

We have observed the presence of two bands with different mobility that indicates the formation of at least two very different structures in KCl (***Fig.14a***) and in NaCl (***Fig.14b***).

Circular Dichroism analysis

To prove that the secondary intramolecular structures observed in the presence of K^+ and Na^+ are really G-quadruplex structures as we have assumed, we have analyzed the 3'UTR region of KiSS1 by CD.

Depending on the direction of the strands or parts of a strand that form the tetrads, structures may be described as parallel, antiparallel or mixed (***Fig.15***).

In the CD spectra, an antiparallel G-quadruplex shows positive and negative peaks around 290 and 265 nm respectively, while a parallel G-quadruplex shows positive and

negative peaks around 260 and 240 nm respectively. A mixed G-quadruplex shows positive peaks around 260 and 295 nm (38).

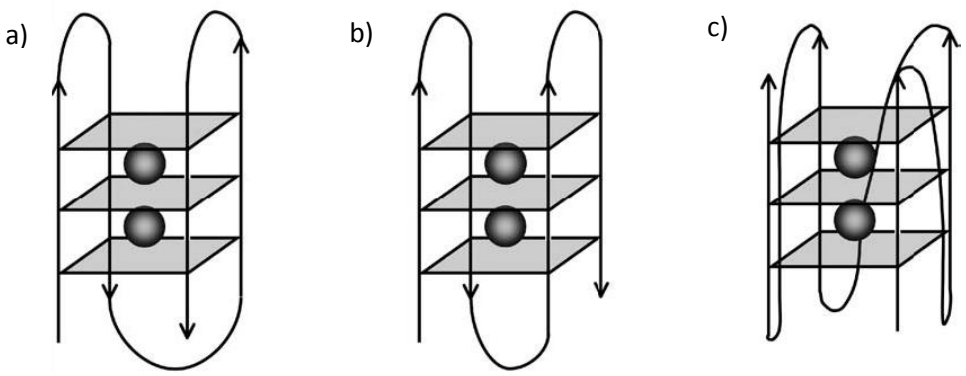

Fig.15: sample schematic quadruplex folds. a) Unimolecular antiparallel. b) Unimolecular antiparallel. c) Unimolecular parallel. Many sequences can also form mixed structures (reference 27).

CD analysis was performed in presence of K^+ or Na^+ and we have observed for the first time the formation of G-quadruplex for both conditions tested (***Fig.16***): depending on the buffer used, the conformation of G-quadruplex changes: in the presence of KCl the oligo assumes an antiparallel conformation (***Fig.16a***) whereas in the presence of NaCl buffer a mixed conformation (parallel/antiparallel) (***Fig.16b***):

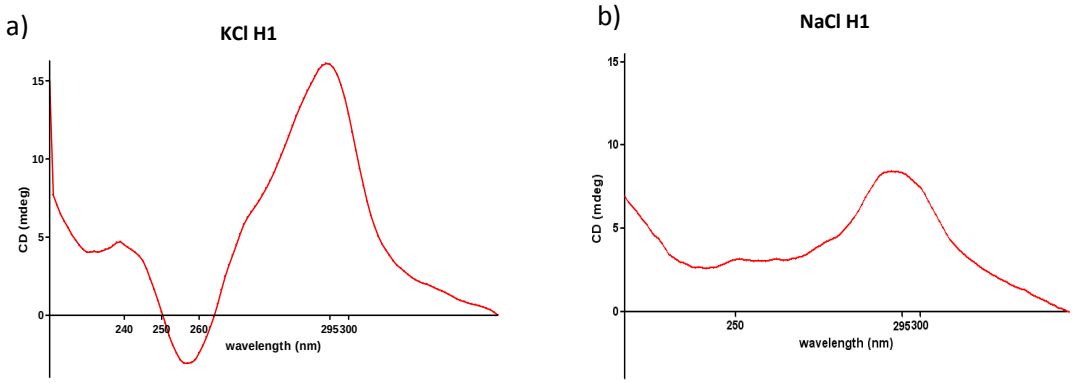

Fig.16: out-put of CD analysis by GraphPad Prism. a) G-quadruplex structure of H1 in the presence of K^+, a typical curve of an antiparallel G4 conformation. b) G-quadruplex structure of H2 in the presence of Na^+, a typical curve of a mixed G4 conformation.

Then CD spectra were also calculated at different temperatures (4°C, 60°C, 80°C and 100°C) to analyze how these secondary structures are stable. The same curve was obtained at all the temperatures tested, it means that the structure is very stable and therefore it keeps the same conformation without denaturing despite changes in

temperature. Changes in the curves mean that the secondary structure of the oligonucleotide is less stable and is easily denatured by increasing the temperature or it also means that the oligo do not form only one structure (***Fig.17***).

We have observed that, in the presence of K^+ (***Fig.17a***), until 80°C, G-quadruplex structures of H1 are more stable than in the presence of Na^+ (***Fig.17b***), indeed all the curves (at 4°C, 60°C and 80°C) are perfectly overlapped if we calculate the CD spectra in the presence of K^+, while this does not happen in the presence of Na^+.

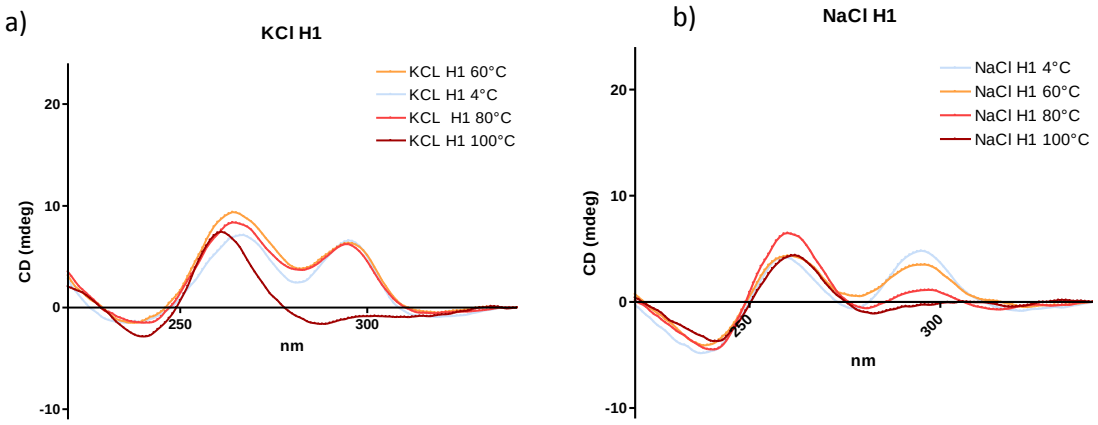

Fig.17: CD spectra of H1 at different temperatures (4°C, 60°C, 80°C and 100°C). a) In the presence of KCl. b) In the presence of NaCl.

With this experiment we have been able to demonstrate that the G-rich 3'UTR region of KiSS1 effectively fold into a particular secondary structure and that this structure is really a G-quadruplex structure as we had assumed. Moreover, we have shown that G-quadruplexes are stabilized by high cations concentrations, particularly by K^+ and Na^+.

2) Modifications of G-quadruplex conformation as a function of polymorphisms

Circular Dichroism analysis

To understand if the polymorphisms into the 3'UTR region of KiSS1 interfere with the conformation of the G-quadruplex, we have submitted H2, H3 and H4 to CD analysis. H1 is the most common haplotype in the normal population and the others three different haplotypes of KiSS (H2, H3, H4) contain polymorphisms that are often found in patients with CPP (Central Precocious Puberty). Depending on the type of sequence, we have obtained different G-quadruplex structures: only H1, in the presence of K^+, can

form an antiparallel G-quadruplex conformation (***Fig.16a and Fig.18a***) while in all the others cases we got mixed conformation: parallel/antiparallel (***Fig18.***):

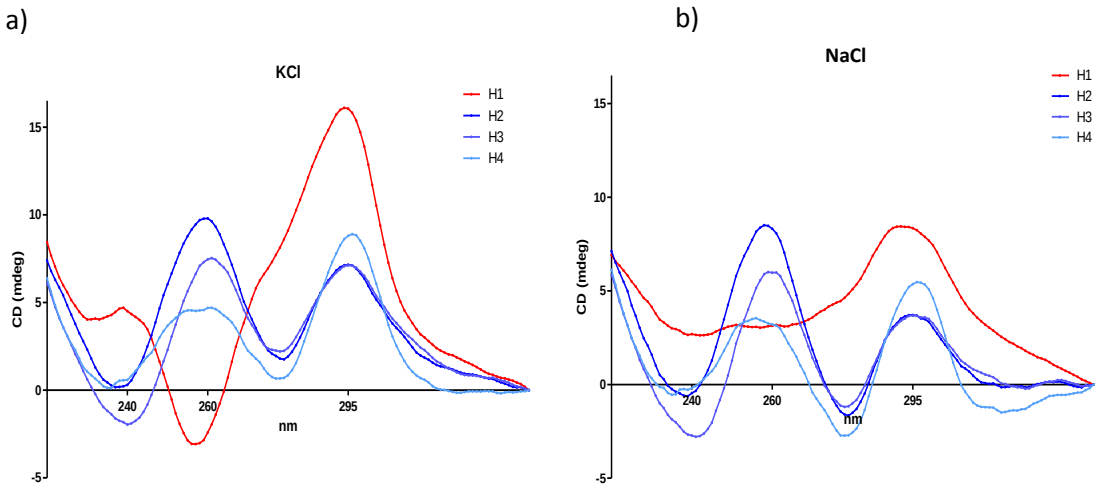

Fig.18: CD spectra of H1, H2, H3 and H4. a) In the presence of K^+. b) In the presence of Na^+.

We have also done CD analysis at different temperatures for each polymorphism. We have observed that, in the presence of K^+ (***Fig.19a, c, e***), until 80°C, structures of all oligonucleotides analyzed are more stable than in the presence of Na^+ (***Fig.19b, d, f***), indeed all the curves (at 4°C, 60°C and 80°C) are perfectly overlapped if we calculate the CD spectra in the presence of K^+, while this does not happen in the presence of Na^+.

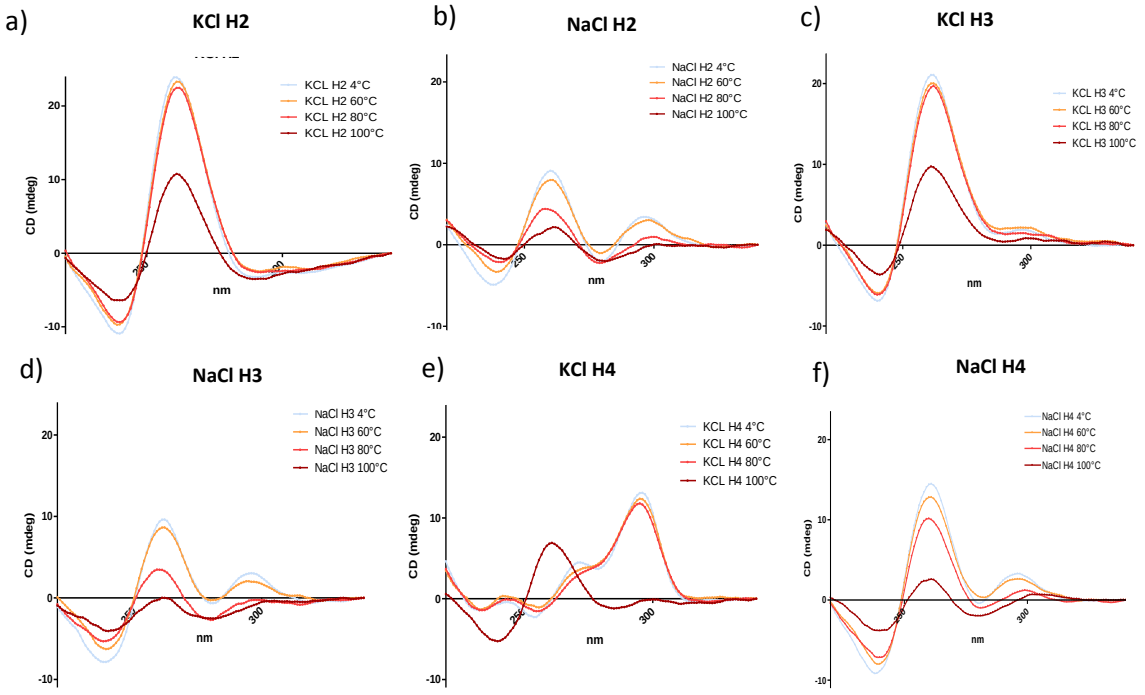

Fig.19: CD spectra of H2 (panels a and b), H3 (panels c and d), H4 (panels e and f) and H4 (panels g and h) at different temperatures (4, 60, 80 and 100°C) in the presence of K^+ (panels a, c and e) and in the presence of Na^+ (panels b, d and f).

We have observed that H2, H3 and H4 behave differently from H1 in the presence of K^+:

oligo	4°C K^+	60°C K^+	80°C K^+	100°C K^+
H1	Mixed	Mixed	Mixed	Parallel
H2	Parallel	Parallel	Parallel	Parallel
H3	Parallel	Parallel	Parallel	Parallel
H4	Antiparallel	Antiparallel	Antiparallel	Parallel

We can observe that the polymorphisms behave differently, in particular H2 and H3 in the presence of KCl (***Fig.19a and c***) fold into parallel G-quadruplexes that are typical of intermolecular structures.

Gel electrophoresis

The gels done in the presence of KCl and NaCl for all the polymorphisms (***Fig.14***) confirm the results obtained by CD. In the presence of NaCl (***Fig.14a***) we can observe for each polymorphism both slow and fast migrating band. In the presence of KCl (***Fig.14b***) instead, we still observe the formation of intermolecular G-quadruplexes for each polymorphism but intramolecular structures are present only with H1 and H4.
We have thus confirmed that polymorphisms behave differently from each other.
Moreover, the absence of intramolecular structures with H2 and H3 in the presence of K^+ is also confirmed by the CD spectra at different temperatures (***Fig.19a, c***) that have detected G-quadruplex structures with a parallel conformation. It is known that intermolecular G-quadruplexes are more likely to take a parallel conformation. The spectra detected for H1 and H4 instead, till 80°C, show a mixed and an antiparallel conformation respectively, while at 100°C they are also inclined to assume a parallel conformation. It is probably due to the formation of intermolecular structures favored by high temperature and DNA denaturing.

UV-melting experiment

We have also calculated the T_m of the four oligonucleotides at different wavelengths that is 245, 260, 273 and 295 nm which are the wavelengths at which the peaks of CD quadruplex spectra are recorded.

In the presence of Na^+ (***Fig.20a, b***), the thermal denaturation curves at 295 nm indicate that the G-quadruplexes formed by all the 4 sequences analyzed have approximately the same T_m (about 75°C). In the presence of K^+ (***Fig.20c***) instead the curves are flat and it is probably due to the fact that the G-quadruplexes formed are so stable that do not denature even at high temperatures.

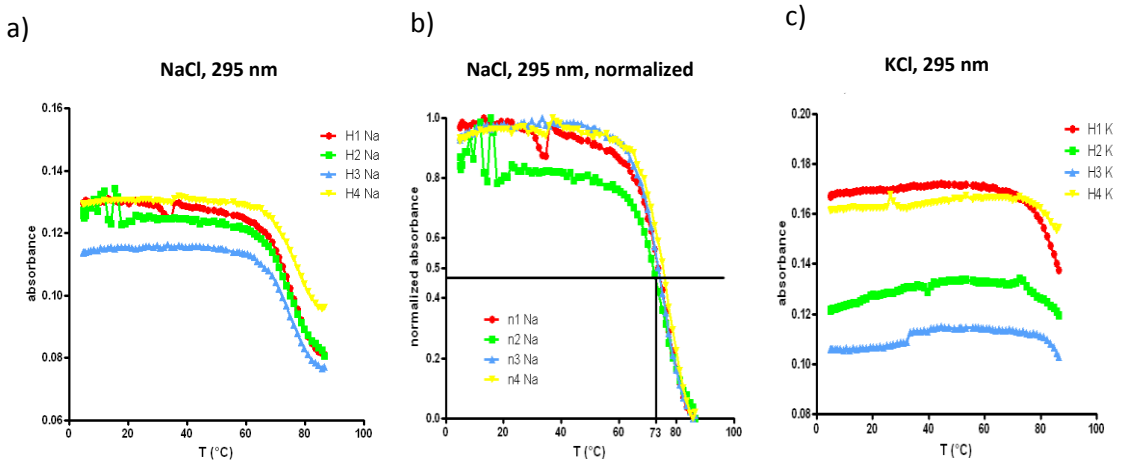

Fig.20: Thermal denaturation Curves at 295nm. a) Raw data in the presence of Na^+. b) Normalized curves in the presence of Na^+. c) Raw data in the presence of K^+.

Thermal Difference Spectra

At least, for a further confirmation of our results, we have determined the Thermal Difference Spectra (TDS) of the oligonucleotides of interest in the presence of K^+ (***Fig.21a***) and also in the presence of Na^+ (***Fig.21b***). A TDS is obtained for a nucleic acid by simply recording the ultraviolet absorbance spectra of the unfolded and folded states at temperatures above (92°C) and below (2°C) its melting temperature. The difference between these two spectra is the TDS. The TDS has a specific shape that is unique for each type of nucleic acid structure and it reflects the subtleties of base stacking interactions that occur uniquely within each type of nucleic acid structure.

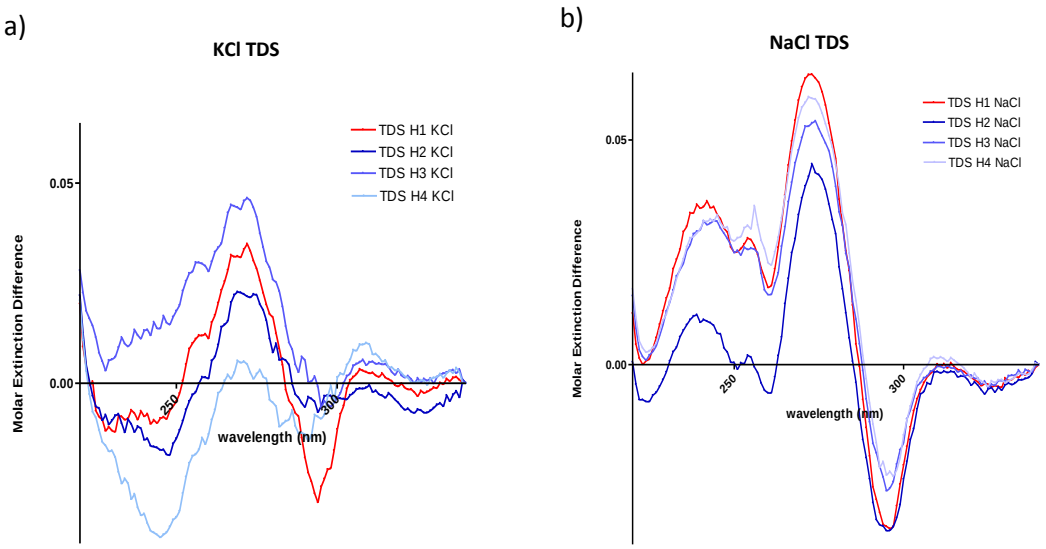

Fig.21: Thermal Difference Spectra of H1, H2, H3 and H4. a) TDS in the presence of K^+. b) TDS in the presence of Na^+.

We have obtained a typical spectrum of G-quadruplex structures with two well distinguishable peaks at 273 and 295 nm and another peak a little less marked at 245°C. In the presence of K^+ curves are less clear and precise because the oligonucleotides are more stable and they are still structured and folded at high temperature.

Whit this experiments we have demonstrated that the polymorphisms, found into CPP patients, affect the conformation of the secondary structure of the 3'UTR of KiSS1, without abrogating it. We have shown that they fold into different G-quadruplex conformation (parallel, antiparallel, mixed) and that they are stabilized especially in the presence of K^+. Their different behavior from H1 could partly explain the ill regulation of the gene in CPP patients.

3) Preliminary study of interactions between G-quadruplex of KiSS1 and proteins

To study proteins-DNA interactions we have performed FRET analysis after incubated DNA (F22T) with proteins one hour to allow the attachment of the proteins on the DNA.

The results we have obtained (***Fig.22***) show that there are some differences between the denaturation of F22T in the presence of proteins (***Fig.22*** *blue line*) end its denaturation

in the absence of them (***Fig.22*** *red line*) and it suggests that the oligonucleotide analyzed (KiSS1) interacts with proteins:

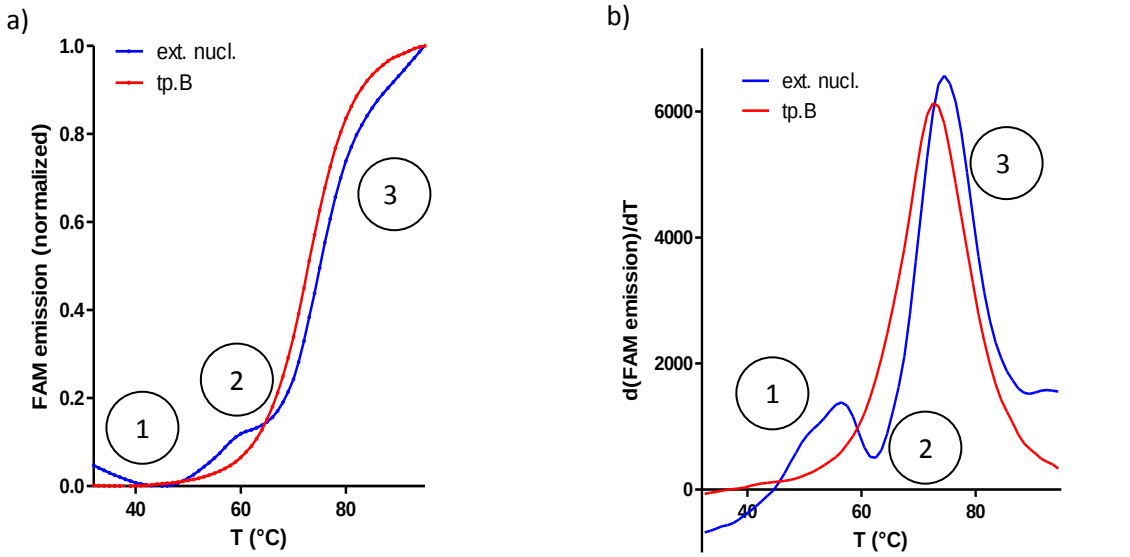

Fig.22: FRET analysis in the presence (blue line) and in the absence (red line) of proteins and stabilization of F22T in the presence of them. a) Normalized melting curves. b) First derivative of the melting curves.

To note:

in the normalized curves (***Fig.22a***):

1) higher initial value of fluorescence in the presence of proteins
2) the hump of the curve in the presence of proteins at about 60°C
3) the T_m is higher in the presence of proteins

in the first derivative curves (***Fig.22b***):

1) a velocity increase between 50 and 55°C
2) a rapid decrease in velocity between 55 and 60°C
3) a very quickly resumed velocity above 60°C

To show that these effects are characteristic of Kiss1 we have performed the same experiment with F21T and we have obtained a normalized curve without the hump and the initial higher value:

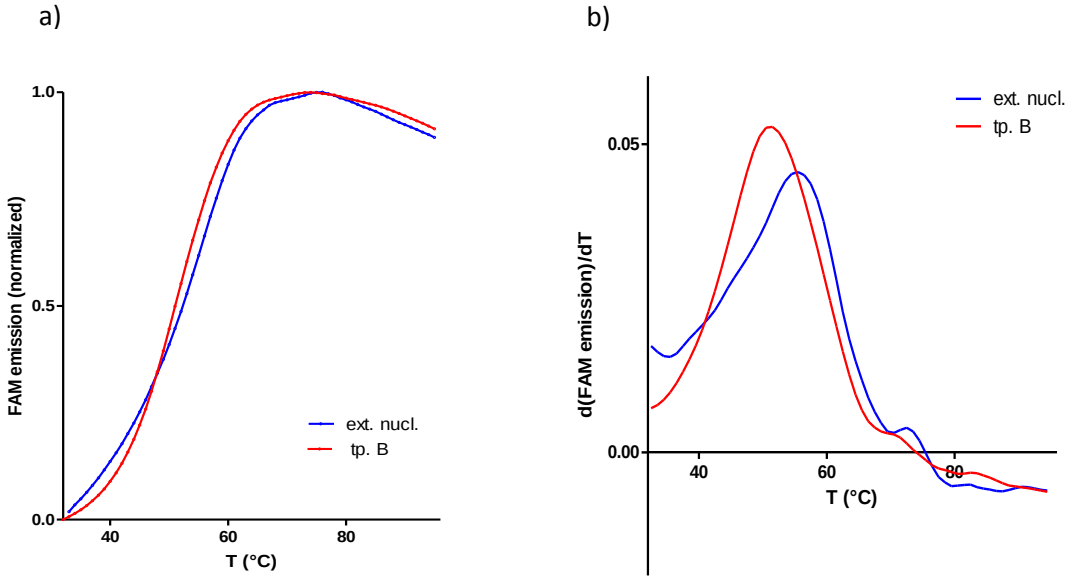

Fig.23: FRET analysis of F21T in the presence (blue line) and in the absence (red line) of proteins. a) Normalized melting curves. b) First derivative of the melting curves.

Competition experiments

To confirm the interaction between the proteins and the G-quadruplex of KiSS1, a competition experiment was performed. We have mixed in the same solution the nuclear extract (29 µg/µl), F22T (at a final concentration of 0.2 µM) and H1 at different concentrations (1/2/10/20 µM). As a control we have prepared a solution without the nuclear extract and H1.

We have seen that if the concentration of the competitor increases the T_m and the stability of F22T (KiSS1) decreases:

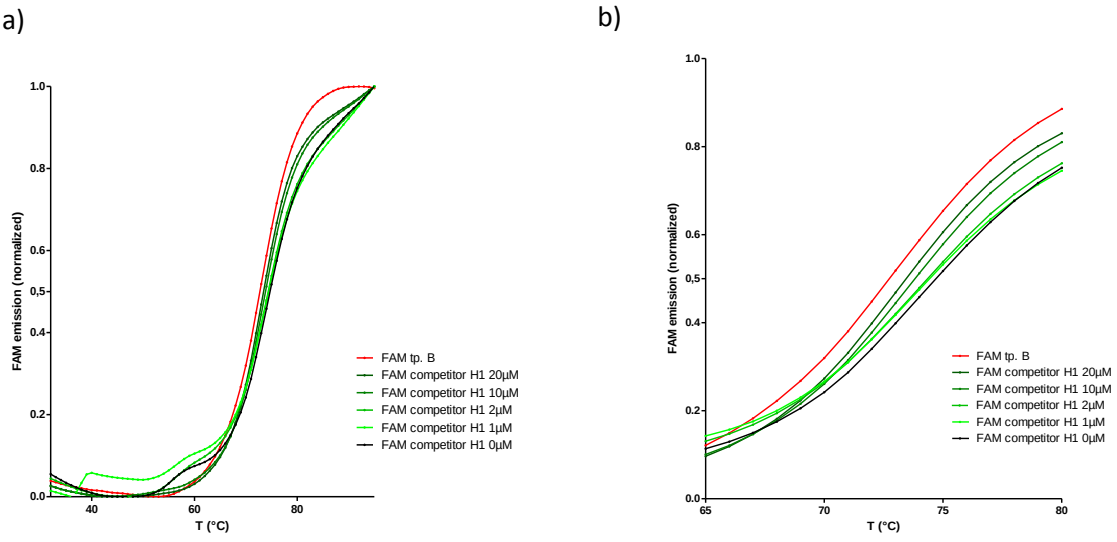

Fig.24: FRET analysis of F22T in the presence of protein (green and black curves) and in the absence of them (red curves). The thermal denaturation profile of F22T is

recorded alone (black line) and in the presence of different H1 concentrations (green lines, 0/1/2/10/20 µM). a) Normalized melting curves. b) Enlargement of panel a.

We have also noticed that in the absence (***Fig.24***, *black line*) and at low concentrations (1 and 2 µM) of competitor H1 the curves show the characteristic hump seen in the presence of proteins. However, as the unlabelled DNA increases, the hump disappears. This result highly suggests that the protein interaction with DNA is specific to KiSS1 sequence.

This last experiment shows that the G-rich 3'UTR of KiSS1 interacts specifically with proteins and that these proteins seem to destabilize the G-quadruplex structure.
The importance of finding proteins that bind G-quadruplexes of KiSS1 to alter their stability, is crucial and could have implications on a pharmacological point of view. In fact, if KiSS*1*-stabilizing proteins will be identified, it is possible to act on them in order to decrease KiSS1 expression levels and so delay puberty in CPP patients.

Discussion

In this study, we provided evidence that a G-rich sequence located in the 3'UTR of KiSS1 fold into G-quadruplex structure. This G-quadruplex structure is highly stable and it folds in parallel, antiparallel and mixed conformation. We also showed that single nucleotide polymorphisms of this sequence change this folding. Finally, we obtained preliminary results indicating that this sequence specifically binds nuclear proteins.

We initially decided to study DNA sequence rather than RNA sequence for several reasons. G-quadruplex located in promoter regions was described as potentially involved in the regulation of gene expression (38). This sequence may interfere with the capacity of the DNA polymerase to replicate DNA and one may suspect a possible modulation by G-quadruplex of the RNA polymerase activity as well. Finally, biophysical analysis of DNA sequence is easier to interpret than RNA sequence.

Biophysical studies of this DNA sequence using FRET, CD and melting temperature analysis showed that this G-rich sequence is extremely stable and could potentially forms G-quadruplex structures under physiological conditions. With K^+ concentration as low as 10 mM, F22T adopt G-quadruplex conformation. We observed an increase of the T_m with the concentration of K^+. In K^+ concentration as low as 50 mM, the complete denaturation of the peptide was impossible. This very high stability for a DNA G-quadruplex is uncommon. RNA G-quadruplexes are usually more stable than DNA G-quadruplexes (24). It is therefore highly probable that the RNA molecule of this G-rich sequence will be more stable than the DNA sequence analyzed in this work.

To understand if the structures that are formed and stabilized in the presence of monocations are intramolecular or intermolacular, FRET analysis was made with different conditions.

1. Changing F22T concentration, we still observe FRET and the normalized curves drawn for each concentration are perfectly overlapped. It means that the folding and the T_m do not depend on the oligonucleotide concentration suggesting that G-quadruplexes are mostly intramolecular.

2. Addition of unlabeled DNA did not change the FRET analysis. There is no competition between the marked DNA and the unlabeled one to form G-quadruplex which reinforces the conclusion that the quenching of the FAM fluorescence by TAMRA occurred by intramolecular rather than intermolecular quenching.
3. In the electrophoresis gel indeed, the presence of two bands indicates possible intramolecular and intermolecular structure. It is important to point out that the native gel was performed with very high concentration of DNA (30 μM) that favors the formation of intermolecular structures. FRET and CD were performed with DNA concentrations at 0,2 and 3 μM respectively.

CD analysis revealed that F22T folds in parallel and antiparallel conformations. In the first experiment, we only analyzed H1 sequence. We clearly observed different conformation in K^+ and Na^+ buffer. In K^+ buffer, we observed an antiparallel conformation although a mixed conformation was observed in Na^+ buffer. Such difference of conformation between K^+ and Na^+ has been already described for other G-quadruplexes (Guedin et al, 2008). This difference seems to be more frequent for G-quadruplex with long loops between G-repeats (Guedin et al, 2008). F22T has loops with one or two nucleotides which indicate that the capacity of G-rich sequences to fold into parallel and antiparallel G-quadruplex conformations is not only determined by the size of the loops.

The TDS, performed in the presence of the two different cations, has confirmed the formation of G-quadruplex secondary structures because the peaks detected are typical of G-quadruplex structures: the peaks of the curves were observed at 240, 273 and 295 nm. Moreover, TDS analysis has also confirmed the greater stability of G-quadruplex in the presence of K^+ rather than Na^+. Our results confirm that folding schemes in sodium appear more diverse than in potassium.

The second interest of this study was to analyze whether single nucleotide polymorphisms or small insertion of two nucleotides, described as polymorphisms, may interfere with the capacity of F22T to adopt a G-quadruplex structure.

To answer to this question, we analyzed H1, H2, H3 and H4 sequences by CD, TDS and native gel electrophoresis. *In silico* analysis has revealed that these polymorphisms

indeed change the size of the loops. They also change the guanines directly involved in the formation of the G-quartets without changing the number of G-quartets. As shown in figure 2, analysis with the QGRS mapper software revealed that H1 folds into a G-quadruplex with four G-tetrads connected by two nucleotides loops. The difference between H1 and the three others sequences H2, H3 and H4 resides in the size of the loops and the guanines involved in the G-quartets. None of these polymorphisms change the number of G-tetrads. G-rich sequences are more prone to into G-quadruplexes if loop lengths are short and equal in size and if the number of G-tetrads is high (10).

Our results confirm the *in silico* analysis. These polymorphisms alter, but do not abrogate, G-quadruplex structures formation. Indeed, all the spectra, both in the presence of K^+ and in the presence of Na^+, show typical G-quadruplexes curves. In K^+ buffer H1 has an antiparallel conformation whereas H2, H3, H4 have a mixed conformation; in Na^+ buffer we have detected the formation of mixed conformations for all oligonucleotides.

To quantify the stability of the oligonucleotides and to see if the polymorphisms cause a change in stability or if the temperature affects the conformation of G-quadruplex, we have calculated the CD spectra at different temperature for each oligonucleotide in the presence of K^+ and in the presence of Na^+. Indeed we have noticed that, in the presence of K^+, haplotype H1 shows a mixed conformation until 80°C but at 100°C it assumes a parallel conformation; H2 and H3 show a parallel conformation at each temperature tested; at least, H4 shows an antiparallel conformation until 80°C and it assumes a parallel conformation at 100°C. In the presence of Na^+ they all have the same mixed conformation until 80°C and they all assume a parallel conformation at 100°C. High temperatures seem to favor a parallel conformation and, if we look at the curves, we can observe a greater stability in K^+ buffer because the curves until 80°C are perfectly overlapped, while in Na^+ they are not.

In addition, results of the CD spectra as a function of temperature suggest that the slow migration bands of H2 and H3 in the presence of K^+ are probably due to the formation of tetraplex structures (or intermolecular structures) that consist of the association of 4

different strands. This is because H2 and H3 form typical spectra of parallel G-quadruplexse structures, very infrequent in DNA G-rich sequences, but typical of intermolecular structures.

All together, these results demonstrate that the polymorphisms of the 3' UTR, found in our cohort of controls and patients with central precocious puberty, disturb the formation of G-quadruplex. They clearly show a net difference between H1 folding and the folding of three other sequences H2, H3 and H4. The functional analysis of these polymorphisms was initially performed using a reporter gene. For that purpose, the 3' UTR sequence was inserted at the 3' end of the luciferase cDNA. Then breast cancer cell line was transfected with each construct. 48 hours after transfection, luciferase activity was measured in cell extracts. The highest luciferase activity was observed in H1-transfected cells. The differences of the luciferase activity due to H1-transfection and the three others sequences were significant between H1- and H3- or H4-transfected cells but not between H1- and H2-transfected cells. It is thus tempting to propose that the capacity of H1 to fold into a different conformation than H2, H3 and H4 explains the difference of luciferase activity. Additional functional analyses are needed to definitively confirm this hypothesis.
The relationship between the functional consequences of these polymorphisms and the precocious onset of puberty is unclear. In very small cohort of patients, we observed an association between H3 and the idiopathic central precocious puberty. H3 has the lowest frequency in the control population, less than 5%. The frequency of H3 was 21% in our cohort of patients with idiopathic central precocious patients. We therefore suggested H3 interferes with the expression level of KiSS1. Our results do not support this hypothesis as we did not observe any difference between H2, H3 and H4 G-quadruplex conformation. However, we have restricted our analysis to DNA conformation. It is now necessary to reproduce this analysis for RNA molecules.

FRET analysis with F22T (KiSS1), in the presence of nuclear extract of MDA MB 231 cells, has pointed out some differences between the F22T denaturing in the presence and in the absence of proteins. It means that some proteins probably interact with this G-quadruplex and interfere with its stability. We have distinguished three different points

in the behavior of F22T in the presence of proteins that we did not observe in the absence of proteins:

1) Higher initial value of fluorescence:

This higher value of FAM emission means that the G-quadruplex is less stable than in the absence of proteins, the two ends of the sequence are more distant and TAMRA does not perfectly quench FAM. At low temperature, nuclear proteins seem to destabilize G-quadruplex.

2) The hump of the curve at about 60°C:

We have observed this result only with high proteins concentrations (from 28µg/µl). This result indicates that nuclear proteins could destabilize the F22T conformation.

3) A higher T_m:

At temperature higher than 60°C, the curve of F22T in the presence of proteins is shifted to the right and the T_m is higher than in the absence of proteins; this result suggests that at high temperatures, F22T is more resistant to the denaturation in the presence of nuclear proteins.

Similarly we can observe some differences between the first derivative curves in the presence and in the absence of proteins:

1) between 50 and 55°C, the velocity of the denaturation increase: it means that G-quadruplexes are probably destabilized by some proteins;
2) between 55 and 60°C velocity decreases rapidly: it means that G-quadruplexes destabilization remains stable;
3) above 60°C velocity starts to increase again very quickly and this is probably due to the combined effect of proteins and G-quadruplexes denaturation induces by heat.

We observed that H1 was able to compete the effects of protein on F22T denaturation. It means that the proteins that normally stabilize the G-quadruplex of F22T (KiSS1) are now committed also with H1. Proteins cannot completely stabilize G-quadruplex of F22T because a part of them is now bound to H1. The effect of proteins on the denaturation of F22T seems to be specific to the G-rich sequence.

Conclusion

We have demonstrated in this work that a G-rich sequence located in the 3' untranslated region of KiSS1 is able to fold into G-quadruplex structure. We have also evidence that polymorphisms which change the size of the loops but not the number of G-tetrads interfere with the G-quadruplex conformation *in vitro*. Finally, this G-rich sequence specifically binds nuclear proteins. This study has been focused on DNA sequence. The capacity of DNA to fold into G-quadruplex seems to be involved in different molecular mechanisms regulating gene expression. Most of these G-quadruplexes are located in the promoter region. The role of G-quadruplex in other parts of a gene, as the coding sequence or the untranslated regions, remains poorly understood. We may speculate that G-quadruplex in exonic or intronic sequence could also interfere with the activity of the RNA polymerase. It would be interesting to indeed test this hypothesis.

This G-quadruplex located in the 3' UTR of KiSS1 may also act as a post-ranscriptional regulator of gene expression by different molecular mechanisms. It is thus necessary to perform this analysis with RNA molecules. This structural analysis must be coupled to a functional analysis to confirm that this G-quadruplex may be a new pharmacological target to modulate KiSS1 expression.

References

[1] Julian L. Huppert, Shankar Balasubramanian (2005). **Prevalence of quadruplexes in the human genome.** *Nucleic Acids Research, Vol. 33, No 9, 2908-2916.*

[2] Stave G. Hershman, Qijun Chen, Julia Y. Lee, Marina L. Kozak, Peng Yue, Li-San Wang, F. Brad Johnson (2008). **Genomic distribution and functional analyses of potential G-quadruplex-forming sequences in *Saccharomyces cerevisiae.*** *Nucleic Acids Research, Vol. 36, No 1, 144-154.*

[3] Johanna Eddy, Nancy Maizels (2008). **Conserved elements with potential to form polymorphic G-quadruplex structures in the first intron of human genes.** *Nucleic Acids Research, Vol. 36, No 4, 1321-1333.*

[4] Hans J. Lipps, Daniela Rhodes (2009). **G-quadruplex structures: *in vivo* evidence and function.** *Trends in Cell Biology, Review.*

[5] L. de Vries, B. Shtaif, M. Phillip, G. Gat-Yablonski (2009). **Kisspeptin serum levels in girl with central precocious puberty.** *Clinical Endocrinology 71, 524-528.*

[6] Masato Kotani, Michel Detheux, Ann Vandenbogaerde, David Communi, Jean-Marie Vanderwinden, Emmanuel Le Poul, Stéphane Brézillon, Richard Tyldesley, Nathalie Suarez-Huerta, Fabrice Vandeput, Cédric Blanpain, Serge N. Schiffmann, Gilbert Vassart, Marc Parmentier (2001). **The metastasis suppressor gene Kiss-1 encodes kisspeptins, the natural ligands of orphan G protein-coupled receptor GRP54.** *The Journal of Biological Chemistry, Vol. 276, No 37, 34631-34636.*

[7] Anne De Cian, Lionel Guittat, Markus Kaiser, Barbara Saccà, Samir Amarne, Anne bourdoncle, Patrizia Alberti,Marie-Paule Teulade-Fichou, Laurent Lacroix, Jean-Louis Mergny (2006). **Fluorescence-based melting assay for studying quadruplex ligands.** *Methods 42 (2007) 183-195.*

[8] Susanna Cogoi, Manikandan Paramasivan, Barbara Spolaore, Luigi E. Xodo (2008). **Structural polymorphism within a regulatory element of the human KRAS promoter: formation of G4-DNA nuclear proteins.** *Nucleic Acids Research, Vol. 36, No 11, 3765-3780.*

[9] Aurore Guédin, Anne De Cian, Julien Gros, Laurent Lacroix, Jean-Louis Mergny (2008). **Sequence effects in single-base loops for quadruplexes.** *Biochimie 90 (2008) 686-696.*

[10] Ja Yil Lee, D. S. Kim (2009). **Dramatic effect of single-base mutation on the conformational dynamics of human telomeric G-quadruplex.** *Nucleic Acids Research, Vol. 37, No 11, 3625-3634.*

[11] Yan Jin, Hongyan Li, Jinyan Bai (2009). **Homogeneous selecting of a quadruplex-binding ligand-based gold nanoparticle Fluorescence Resonance Transfer assay.** *Anal. Chem. 2009, 81, 5709-5715.*

[12] Jean-Louis Mergny, Jing Li, Laurent Lacroix, Samir Amrane, Jonathan B. Chaires (2005). **Thermal difference spectra: a specific signature for nucleic acid structures.** *Nucleic Acids Research, 2005, Vol. 33, No 16 e138.*

[13] Kah Wai Lim, Patrizia Alberti, Aurore Guédin, Laurent Lacroix, Jean-Francois Riou, Nicola J. Royle, Jean-Louis Mergny, Anh Tuan Phan (2009). **Sequence variant (CTAGGG)n in the human telomere favors a G-quadruplex structure containing a G·C·G·C tetrad.** *Nucleic Acids Research, 2009, Vol. 37, No 18, 6239-6248.*

[14] Sigve Nakken, Torbjorn Rognes, Eivind Hovig (2009). **The distruptive positions in human G-quadruplex motifs are less polymorphic and more conserved than their counterparts.** *Nucleic Acids Research, 2009, 1-8.*

[15] X. Luan, H. Yc, X. Wei, Y. Zhou, W. Wang, P. Li, X. Gan, D. Wei, J. Xiao (2007).**GPR54 polymorphisms in Chinese Girls with Central Precocious Puberty.** *Neuroendocrinology 016.*

[16] Samir Amrane, Anne De Cian, Frédéric Rosu, Markus Kaiser, Edwin De Pauw, Marie-Paule Teulade-Fichou, Jean-Louis Mergny (2008). **Identification of trinucleotide repeat ligands with a FRET melting assay.** *ChemBioChem 2008, 9, 1229-1234.*

[17] Richard A. J. Darby, Matthieu Sollogoub, Catherine McKeen, Lynda Brown, Antonina Risitano, Nicholas Brown, Christopher Barton, Tom Brown, Keith R. Fox (2002). **High throughput measurement of duplex, triplex and quadruplex melting**

curves using molecular beacons and a LightCycler. *Nucleic Acids Research, 2002, Vol. 30, No 9 e39.*

[18] Bérangère Guyen, Christoph M. Schultes, Pascale Hazel, John Mann, Stephen Neidle (2004). **Synthesis and evaluation of analogues of 10H-indolo[3,2-b]-quinoline as G-quadruplex stabilizing ligands and potential inhibitors of the enzyme telomerase.** *Org. Biomol. Chem., 2004, 2, 981-988.*

[19] Ernest Blackwell, Xing Zhang, Stephanie Ceman (2010). **Arginines of the RGG box regulate FMRP association with polyribosomes and mRNA.** *Human Molecular Genetics, 2010.*

[20] Jean-Louis Mergny, Laurent Lacroix, Marie-Paule Teulade-Fichou, Candide Hounsou, Lionel Guittat, Magali Hoarau,Paola B. Arimondo, Jean-Pierre Vigneron, Jean-Marie Lehn, Jean-Francois Riou, Thérèse Garestier, Claude Hélène (2000). **Telomerase inhibitors based on quadruplex ligands selected by fluorescence assay.** *PNAS, Vol. 98, No 6.*

[21] Aaron Mendez-Bermudez, Mark Hills, Hilda A. Pickett, Anh Tuan Phan, Jean-Louis Mergny, Jean-Francois Riou, Nicola J. Royle (2009). **Human telomeres thet contain $(CTAGGG)_n$ repeats show replication dependent instability in somatic cells and the male germline.** *Nucleic Acids Research, 2009, Vol. 37, No 18 6225-6238.*

[22] Julian Leon Huppert, Anthony Bugaut, Sunita Kumari, Shankar Balasubramanian (2008). **G-quadruplexes: the beginning and the end of UTRs.** *Nucleic Acids Research, 2008, Vol. 36, No 19 6260-6268.*

[23] Jixun Dai, Megan Carver, Danzhou Yang (2008). **Polymorphism of human telomeric quadruplex structures.** *Biochimie 90 (2008) 1172-1183.*

[24] Oleg Kikin, Zachary Zappala, Lawrence D'Antonio, Paramjeet S. Bagga (2007). **GRSDB2 and GRS_UTRdb: databases of quadruplex forming G-rich sequences in pre-mRNAs and mRNAs.** *Nucleic Acids Research, 2008, Vol. 36, Database issue D141-D148.*

[25] Jeong-Hyung Lee, Mary E. Miele, Deana J. Hicks, Karen K. Phillips, Jeffrey M. Trent, Bernard E. Weissman, Danny R. Welch (1996). **Kiss-1 a novel human**

malignant melanoma metastasis-suppressor gene. *Journal of the National Cancer Institute, Vol. 88, No 23, 1731-1737.*

[26] Markus Wieland, Jorg S. Harting (2009). **Investigation of mRNA quadruplex formation in *Escherichia coli.*** *Nature 2009, Vol. 4, No 11, 1632-1640.*

[27] Julian Leon Huppert (2007). **Four-stranded DNA: cancer, gene regulation and drug development.** *Philosophical Transactions of the Royal Society A.*

[28] Phillip A. Rachwal, Keith R. Fox (2007). **Quadruplex melting.** *Methods 43 (2007)291-301.*

[29] Simina M. Popa, Donald K. Clifton, Robert A. Steiner (2008). **The role of Kisspeptins and GPR54 in the neuroendocrine regulation of reproduction.** *Annu. Rev. Physiol. 2008. 70: 213-38.*

[30] Sophie Messager, Emmanouella E. Chatzidaki, Dan Ma, Alan G. Hendrick, Dirk Zahn, John Dixon, Rosemary R. Thresher, Isabelle Malinge, Didier Lomet, Mark B. L. Carlton, William H. Colledge, Alan Caraty, Samuel A. J. R. Aparicio (2004). **Kisspeptin directly stimulates gonadotropin-releasing hormone release via G protein-coupled receptor 54.** *PNAS, 2005, Vol. 102, No 5, 1761-1766.*

[31] Tetsuya Ohtaki, Yasushi Shintani, Susumu Honda, Hirokazu Matsumoto, Akira Hori, Kimiko Kanehashi, Yasuko Terao, Satoshi Kumano, Yoshihiro Takatsu, Yasushi Masuda, Yoshihiro Ishibashi, Takuya Watanabe, Mari Asada, Takao Yamada, Masato Suenaga, Chieko Kitada, Satoshi Usuki, Tsutomu Kurokawa, Haruo Onda, Osamu Nishimura, Masahiko Fujino (2001). **Metastasis suppressor gene KISS-1 encodes peptide ligand of a G-protein-coupled receptor.** *NATURE 2001, Vol. 411, 613-617.*

[32] E. J. Mead, J. J. Maguire, R. E. Kuc, A. P. Davenport (2007). **Kisspeptin: a multifunctional prptide system with a role in reproduction, cancer and the cardiovascular system.** *British Journal of Pharmacology 151, 1143-1153.*

[33] Luigi Xodo, Manikandan Paramasivam, Alexandro Membrino, Susanna Cogoi (2008). **Protein hnRNPA1 binds to a critical G-rich element of KRAS and unwinds G-quadruplex structures: implications in transcription.** *Nucleic Acids Symposium Series No 52, 159-160.*

[34] Tracey A. Martin, Gareth Watkins, Wen G. Jiang (2005). **KISS-1 expression in human breast cancer.** *Clinical & Experimental Metastasis 22: 503-511.*

[35] Jenet E. Price, Aristidis Polyzos, Ruo Dan Zhang, Lisa M. Daniels (1989). **Tumorigenicity and metastasis of human breast carcinoma cell lines in nude mice.** *Cancer Research 50, 717-721.*

[36] L. Vries, A. Kauschansky, M. Shohat, M. Phillip (2004). **Familial central precocious puberty suggests autosomal dominant inheritance.** *Clin. Endocrinol. Metab. 89: 1794-1800.*

[37] B. Karges, N. De Roux (2005). **Molecular genetics of isolated hypogonadotropic hypogonadism and Kallmann syndrome.** *Endocr. Dev 8: 67-80.*

[38] Dong-Hao Zhang, Takeshi Fujimoto, Sarika Saxena, Hai-Qing Yu, Daisuke Miyoshi, Naoki Sugimoto (2010). **Monomorphic RNA G-quadruplex and polymorphic DNA G-quadruplex structures responding to cellular environmental factots.** *Biochemistry, 2010.*