UNIVERSITÀ DEGLI STUDI DI UDINE

PhD course in Biomedical Sciences and Biotechnology
Cycle XXIV

PhD thesis

G4-DNA in gene promoters and development of anticancer strategies

PhD student: Alexandro Membrino

Tutor: Luigi E. Xodo

Co-tutor: Susanna Cogoi

Academic year
2010/2011

Table of contents

1. Abstract

Recent bioinformatic studies have shown that quadruplex-forming sequences occurs with a high frequency in gene promoter regions, upstream of the Transcription Start Sites (TSS). This important finding has raised the question whether quadruplexes represent structural elements playing a role on transcription regulation. To address this question our laboratory focused on *HRAS* and *KRAS* genes as their promoters contain G-rich elements that can fold into quadruplexes.

An insight into the structure of the *RAS* quadruplexes has been obtained by circular dichroism (CD) and DMS-footprinting assays. Additionally, we found that the *RAS* quadruplexes are bound by transcription factors such as MAZ and Sp1 and that protein hnRNP A1 is able to unfold the *KRAS* quadruplex. To investigate on the role played by quadruplex DNA, we engineered reporter plasmids in which quadruplex-forming sequences were either wild-type or mutated. Transcription data were also obtained with cells treated with small molecules, such as cationic porphyrins and guanidinio phthalocyanines, that stabilize quadruplexes. Together our results allowed us to suggest for the first time a transcription mechanism for the *HRAS* proto-oncogene, according to which quadruplex DNA behaves as a repressor element of transcription.

From this model we developed a decoy strategy to inhibit *KRAS* or *HRAS* transcription in cancer cells. This strategy is based on the use of chemically modified oligonucleotides (this task has been carried out in collaboration with Prof. Erik Pedersen from the University of Southern Denmark, Odense) that mimic the *RAS* quadruplexes which are formed in the critical promoter regions. Upon introduction in the cells, the decoys should bind to MAZ and Sp1 and deprive these promoters of their natural transcription factors. As the *RAS* gene give to cells a mitogenic signal, the decoys, developed in the course of this PhD, showed to be powerful antiproliferative agents with a potential in cancer therapy either to directly cure cancer or to sensitize cancer cells to conventional treatments. Work has been directed towards the elucidation of the mechanism of cell death caused by the decoys.

This study is the starting point of a new anticancer therapy against the tumours bearing mutant *RAS* alleles.

2. Introduction

2.1. Quadruplex DNA

In 1910, Bang reported that concentrated solutions of guanylic acid formed a gel. In fact concentrated solution of about 25 mg/ml of guanylic acids at pH 5 are extremely viscous and, if cooled, form a clear gel. From examination of the optical properties of the gel and investigation of the structure of fibres obtained from the gel drying, at least in the case of 5' isomers, the phenomenon may be explained as being due to helix formation by the guanylic acid (Gellert M et al. 1962). Even if this inclination of guanine-rich oligonucleotides to form polymers has been recognized for nearly forty years, the level of interest in the structure that these G-rich sequences can adopt, has increased markedly in the past three decades to date (Keniry MA 2001). The molecular basis for the association was subsequentially determined by fibre diffraction and biophysical studies using the concept that Hoogsteen hydrogen-bonded guanine-tetrad is the basic structural motif (Burge S et al. 2006).

These structures are variously referred to as G4-DNA, quadruplexes, G-quadruplexes and G-tetraplexes. Quadruplexes can be made from one to four strand of guanines. In the first case they are called intramolecular quadruplexes while in the others they are called intermolecular quadruplexes. Especially in intramolecular quadruplexes, four runs of two or three guanines are separated by short sequence that create loops into the structure (Figure 1).

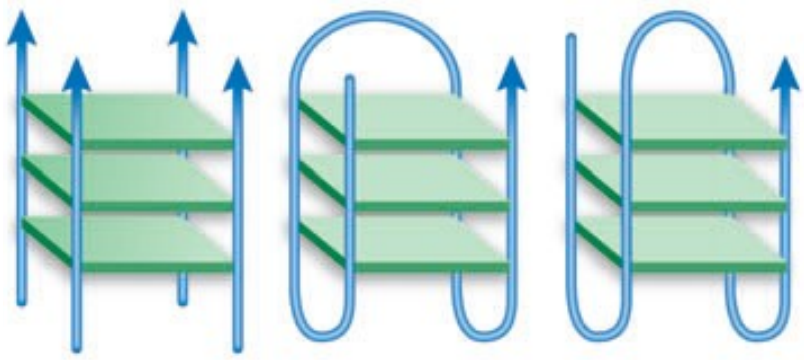


Figure 1

Maizels N **Dynamic roles for G4 DNA in the biology of eukaryotic cells.** Nature Structural & Molecular Biology 2006, 13, 1055-1059

A quadruplex is stabilized by G-quartets, planar arrays of four guanines in which each guanine pairs with two neighbours by Hoogsteen bonding. Basically the quadruplex itself consists of multiple vertically stacked guanine tetrads. Hoogsteen hydrogen bonds between the N_{12}, N_7, O_6 and N_2 guanine bases associate each of the four guanines to form a planar G-tetrad (Figure 2).

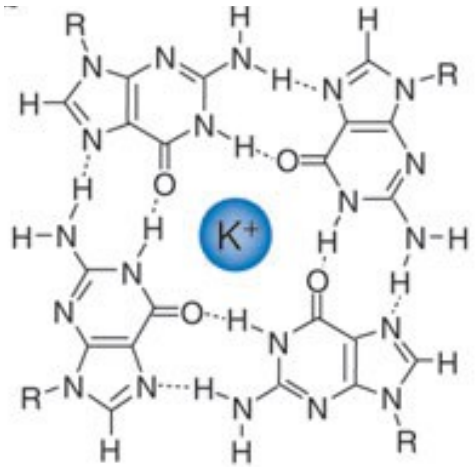


Figure 2

Maizels N **Dynamic roles for G4 DNA in the biology of eukaryotic cells.** Nature Structural & Molecular Biology 2006, 13, 1055-1059

The involvement of N_7 of guanines in the Hoogsteen base-pairing of a quadruplex protects the site from chemical modification, such as methylation by dimethyl sulfate (DMS). This unique feature makes it possible to distinguish a quadruplex structure from single-strand or B-form duplex DNA by DMS footprinting (Qin Y et al. 2008).

Each tetrad forms a central core, comparable in size to a heme moiety, which is encircled by four phosphodiester backbones (Maizels N 2006). The formation and stability of quadruplexes is monovalent cation-dependent. This has been ascribed to the strong negative electrostatic potential created by the guanines O_6 oxygen atoms, which form a central channel of the G-tetrad stacked with the cation located within this channel. The precise location of the cations between the tetrads is dependent on the nature of the ion, with Na^+ ions within the channel being observed in a range of geometries; in some structures a Na^+ ion is in plane with a G-tetrad whereas in other it is between two successive G-tetrads (Figure 3). K^+ ions are always equidistant between each tetrad plane and from the eight oxygen atoms in a symmetric tetragonal bipyramidal configuration (Burge S et al. 2006). In fact quadruplexes form spontaneously in physiological salts. They are stabilized by K^+ at concentration far below the

levels typical of mammal cells (140 mM). As might be expected, intramolecular quadruplex formation occurs in a concentration-independent manner, whereas high DNA concentrations accelerate intermolecular quadruplex formation (Maizels N 2006).

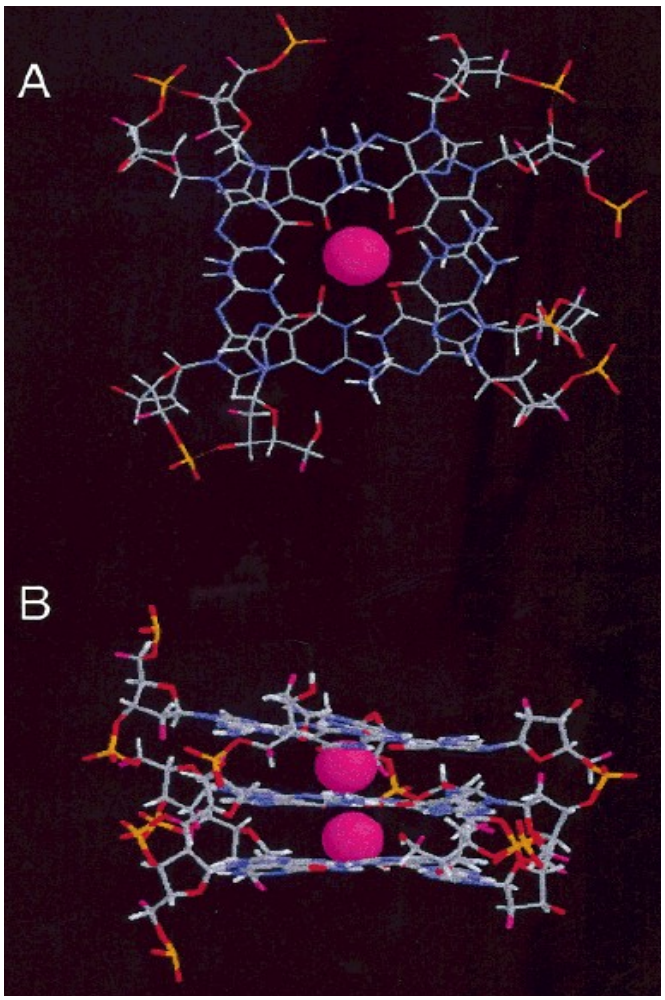


Figure 3

Keniry MA **Quadruplex structures in nucleic acids.** Biopolymers 2001, 56, 123-146

As mentioned above quadruplexes can be formed from one, two or four separate strands of DNA and can display a wide variety of topologies, which are in part a consequences of various possible combinations of strand polarities, as well as variations in loop size and sequence. When the four guanine strands are all parallel the structure is called parallel quadruplex. Strands in such fashion require a connecting loop to link the bottom G-tetrad with the top G-tetrad, leading to propeller type loops (Figure 4c). This feature has been found in crystal structures and in solution for quadruplexes formed from human telomeric DNA sequences.

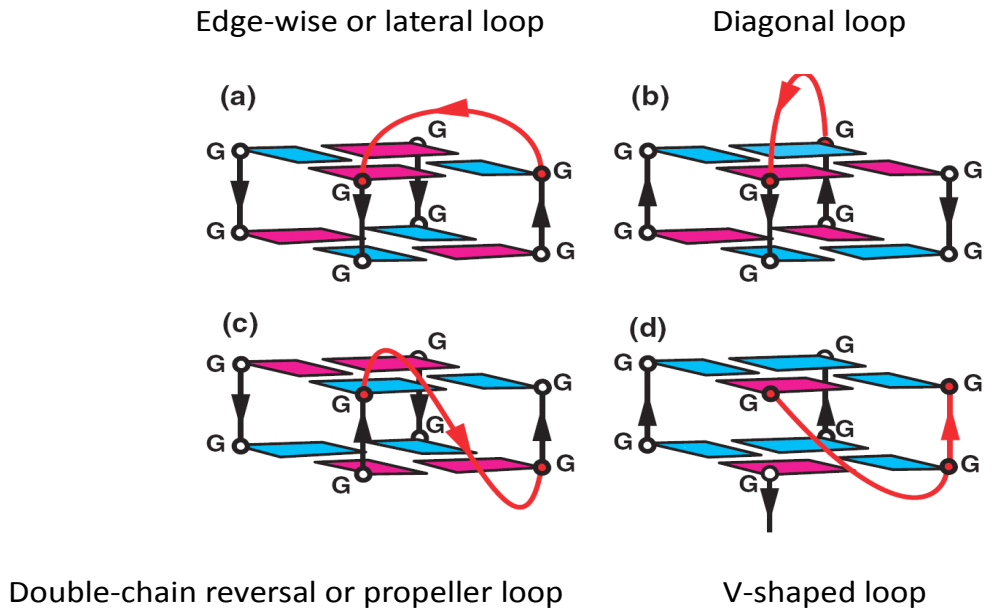


Figure 4

Patel DJ **Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics** Nucleic Acids Research, 2007, Vol. 35, No 22 7429-7455

Quadruplexes are designated as anti-parallel when at least one of the four strands is anti-parallel to the others. This type of topology is found in the majority of bimolecular and in many unimolecular quadruplex structures determined to date. In these structures two further types of loops have been observed. The first type is called lateral (sometimes termed edge-wise) loop and joins adjacent G-strands (Figure 4a). Intramolecular quadruplexes with three lateral loops or edge-wise (Figure 4a) are frequently referred to as chair structures (Keniry MA 2001). Two of these lateral loops can be located either on the same or opposite faces of a quadruplex, corresponding to head-to-head or head-to-tail respectively when in bimolecular quadruplexes. The second type of anti-parallel loop, the diagonal loop, joins opposite G-strands (Figure 4b). In this instance the directionalities of adjacent strands must alternate between parallel and anti-parallel, and are arranged around a core of four stacked G-tetrads.

All parallel quadruplexes have all guanine glycosidic angles in an *anti* conformation. Anti-parallel quadruplexes have both *syn* and *anti* guanines (Figure 5), arranged in a way that is particular for a given topology and set of strand orientations, having the four strands in differing positions relative to each other (Burge S et al. 2006).

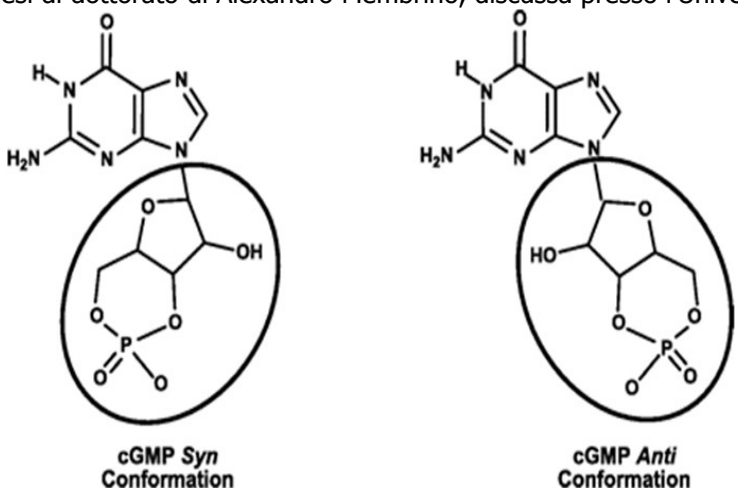


Figure 5

Francis SH, Blount MA, et al. **Mammalian Cyclic Nucleotide Phosphodiesterases: Molecular Mechanisms and Physiological Functions** Physiol. Rev. *2011,* 91, 651–690

All quadruplex structures have four grooves, defined as the cavities, bounded by the phosphodiester backbones. Grooves dimensions are variable and depend on the overall topology and the nature of the loops. Grooves in quadruplexes with only lateral and diagonal loops are structurally simple and the walls of these grooves are bounded by monotonic sugar phosphodiester groups. In contrast, grooves that incorporate propeller loops have more complex structural features that reflect the insertion of the variable sequence loop into the grooves.

Quadruplexes composed of a single dG_n strand almost invariably form right-handed parallel structures that contain the bases in the *anti* conformation. The four strands are oriented parallel to each other and consequently all the four grooves in these quadruplexes are identical. Quadruplexes formed from double dG_n repeats tend to assemble into dimeric structures by the association of a pair of hairpins (Keniry MA 2001) (Figure 6B-F).

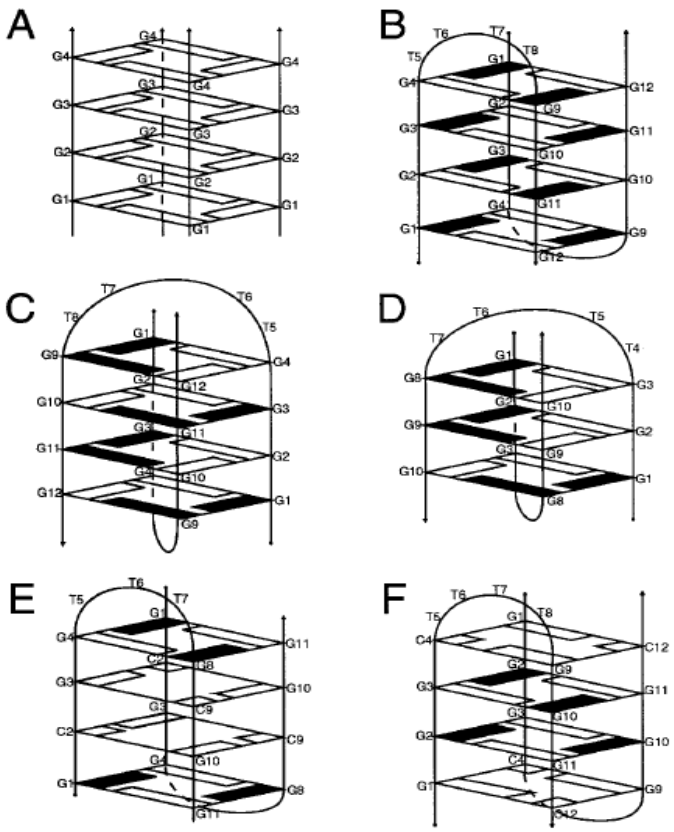


Figure 6

Keniry MA **Quadruplex structures in nucleic acids.** Biopolymers 2001, 56, 123-146

2.1.1. Quadruplex motifs in the human genome

Intramolecular quadruplexes formed by a single DNA strand have attracted much interest because they might form in telomeres, oncogene promoter sequences and other biologically relevant regions of the genome (Bates PJ et al. 2007). Quadruplex-forming sequences have been found in different genomes: from bacterial to human. Analysis of more than 61000 gene sequences across 18 prokaryotes show enrichment of quadruplex motifs in regulatory regions and indicate their predominance within promoters of genes pertaining to transcription, secondary metabolite biosynthesis and signal transduction. Quadruplex DNA may represent a regulatory signal. This is supported by the conservation of quadruplex motifs in promoters of orthologous genes across phylogenetically related organisms (Rawal P et al. 2006). Even in the yeast genome it is possible to find strong enrichment of sequences able to fold into quadruplex structures in upstream promoter regions, as well as weaker but significant enrichment in open reading frames (Hershman SG et al. 2008). In higher eukaryotes it is

possible to find quadruplex-forming sequences. Analysing and comparing the distribution of quadruplex motifs in genomic regions flanking the transcription start sites of 13 animal species suggested an extensive selection for quadruplex enrichment in transcriptional regulatory regions of warm-blooded animals (Zhao Y et al 2007).

The specific location and conservation of G-rich sequences in genomes from different procaryotes and eukaryotes indicates a strong selective pressure to retain the quadruplex-forming potential. A reason for this could be that the regulated formation of quadruplex structures provides an elegant nucleic-acid-based mechanism for modulating transcription and translation. The formation of quadruplexes during replication could well provide an additional level of replication control or a connection to a cell-cycle checkpoint (Todd AK 2007).

For these reasons lot of efforts have been done to understand the localization and abundance of potentially quadruplex-forming sequences throughout the human genome.

Quadruplex motifs are present in the human genome with an average incidence of ~1 quadruplex every 10000 bases (Lipps HJ et al. 2009). Using computational methods it is possible to map quadruplex motifs in gene promoters that may be involved in transcription regulation. Quadruplex were found to be highly prevalent in human gene promoters. The analysis revealed evidence of evolutionary selection pressure to concentrate quadruplex elements in gene promoters and proximal to the TSS of genes.

Promoter quadruplexes were also found to be strongly associated with the open form of chromatinised DNA as judged by experimental genome-wide nuclease hypersensitive data. Quadruplex loops are critical to the stability of folded structure and promoters quadruplexes show an enrichment of stabilizing loops that define a precise fold (Lipps HJ et al. 2009).

Putative Quadruplex Sequences (PQS) are enriched by a factor of 6.4 in gene promoters as compared to the average throughout the whole human genome. Observation rising from computational analysis is that 42.7% of gene promoters contained at least one PQS. Within promoters the chance to find a PQS is directly related to its proximity with the TSS. The highest concentration of PQS occurs very near the TSS, with the PQS density in the first 100 upstream bases being over 12 times higher than the genome average (Lipps HJ et al. 2009) (Figure 7).

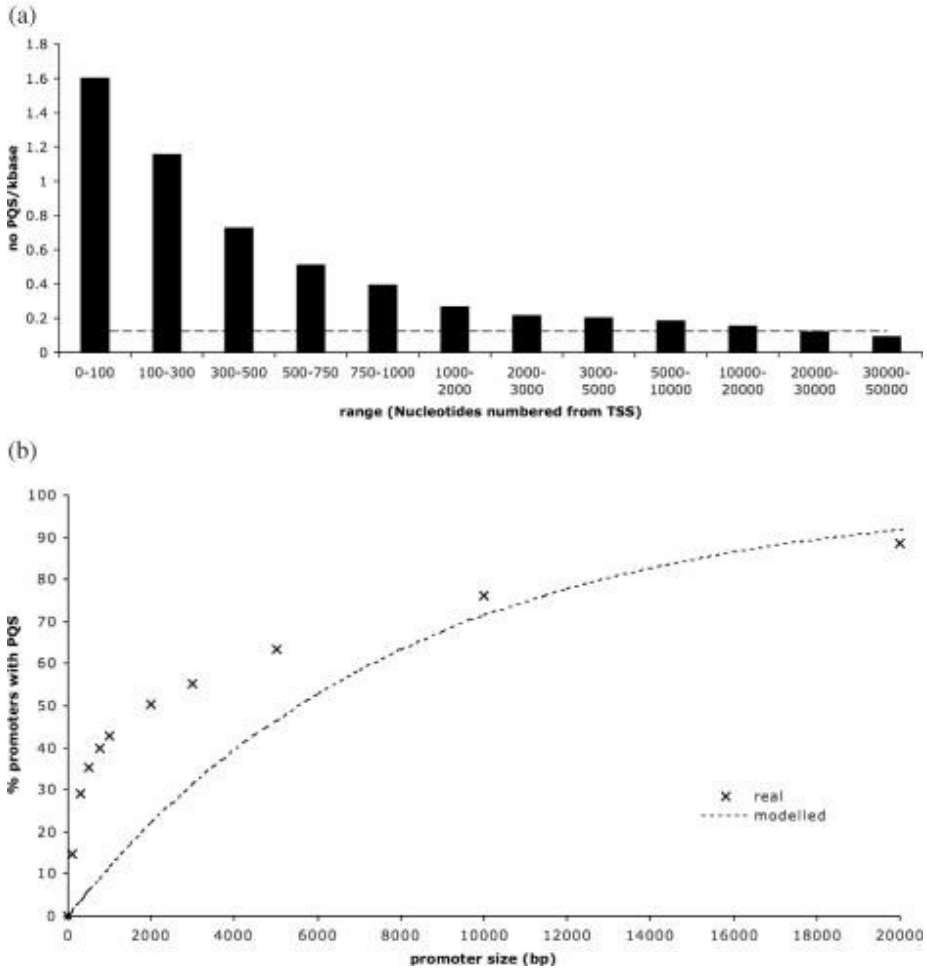


Figure 7

Huppert JL, Balasubramanian S **G-quadruplexes in promoters throughout the human genome** Nucleic Acids Research 2007, 35, 406-413

Genomic analyses can identify sequences in living cells with PQS using algorithms based on analysis of structures formed by synthetic oligonucleotides *in vitro*.

To understand how PQS are located into the human genome some software-based quadruplex search start from a simple folding rule (Huppert JL et al. 2005) describing sequences that may form quadruplexes.

Sequence searches may be used to locate potential quadruplex-forming sequences. However to determine whether they fold into a real quadruplex experimental evidences is required. The current search patterns are usually based on telomeric-like quadruplex topologies and are a useful way to determine the location of quadruplex sequences. However they may miss a number of quadruplex structures which do not follow the conventional pattern of four guanine runs (Huppert JL et al. 2007). Algorithms typically search for the presence of four run of at least three guanines within a window 40-100 nucleotides in length, sufficient for quadruplex

formation within synthetic oligonucleotides. Quadruplexes are dynamic structures and their formation depends on denaturation of the duplex, as occurs during replication, transcription or recombination (Maizels N 2006). It is important to emphasize that algorithms can establish the potential to form this conformation, because most genomic DNA is maintained as a Watson-Crick duplex, in which G-C pairing prevents formation of quadruplex.

Four aspects are usually considered in developing this kind of rule: a. strand stoichiometry; b. the number of stacked tetrads in the quadruplex core; c. the presence of mutations or deletions; and d. the length and composition of loops. This rule predicts sequences that could form a quadruplex, but does not preclude the possibility of some or all of the motifs forming an alternate structure.

For what concern the stoichiometry one should consider that since under physiological conditions the strand concentration of DNA is relatively low (in the order of nM), except in rare cases, interstrand quadruplexes will be strongly disfavoured. Single G-tetrads have only been reported in highly concentrated guanine solutions at mM concentrations, and are unlikely to be physiologically relevant. There are a few examples of double-stack quadruplexes, such as thrombin-binding aptamer and the sequences identified as responsible for the fragile X sindrome (Weisman-Shomer P et al. 2000).

Furthermore an algorithm should consider the loops' length. An intramolecular quadruplex must have three loops to link the tetrads together and they play a large role in determining both the stability and folding pattern of the quadruplex. In summary loops with lengths from 1 to 7 bases were found to form quadruplexes with stability decreasing as length increased (Huppert JL et al. 2005).

Most of these computational analysis pointed out three key repetitive functional domains characterized by high PQS: the telomeres, the ribosomal DNA (rDNA) and, in mammals, the immunoglobulin heavy-chain switch regions. High PQS also characterizes repetitive G-rich microsatellites, including some of the most unstable. In the human genome, more than 300000 distinct sites have potential to form quadruplex (Huppert JL et al. 2007).

Several kinds of evidence support the functional importance of these sites.

First, specific sequence motifs are evident at sites with high PQS, particularly within the loops separating G-runs. Second, among single-copy genes, PQS correlates with gene function as defined by gene ontology terms. Notably PQS is low in tumor-suppressor genes, which maintain genomic stability and are commonly haploinsufficient; but it is high in proto-oncogenes, which promote cell proliferation. Quadruplex formation can correlate with

genomic instability and the low PQS of tumor-suppressor genes may results in selection that protects genes in this class from instability. The high PQS of proto-oncogene is more difficult to explain, but there is an intriguing possibility that it may reflect shared regulation (Maizels N 2006).

Quadruplex can led to genomic instability and this has been taken into account by De and co-workers (2011) resulting in the finding that genomic regions, rich in quadruplexes, are on average hypermethylated in normal tissue; hypomethylation in those region is substantially associated with DNA breakpoint hotspots across a wide range of cancer types. Quadruplex structures and aberrant hypomethylation have a key role in generating genomic alteration and cancer (De A et al. 2011). At the same time one should keep in mind that whereas methylated CpGs restrict transcription, unmethylated CpGs in the vicinity of a gene allow that gene to be expressed (Saxonov S et al. 2006). Genome wide analyses have shown that PQS does correlate with GC-content but not with the number of CpG island. Both exons and introns contribute to the difference in PQS between tumor suppressor genes and proto-oncogenes. In fact tumor suppressor genes have much lower PQS than would be predicted by their genomic environment as compared to the RefSeq genes, whereas proto-oncogenes have higher PQS than would be predicted. The most straightforward interpretation of these results is that genes with specific functions have undergone selection based on quadruplex potential (Eddy J et al. 2006). Notably almost 50% of PQS contain CpG dinucleotides and, for a majority of those cases, the guanine participates in quadruplex formations. PQS show depletion of CpG methylation and nucleosome occupancy, and DNA methylation pattern have a role in the stability of other non-canonical DNA structures such as Z-DNA and H-DNA. Extensive hypomethylation and high quadruplex frequency rarely co-occur in normal tissues (De A et al. 2011).

In normal tissues, the genome is usually hypermethylated and does not show genomic instability, whereas genome-wide hypomethylation is a hallmark of many cancer types.

2.1.2. Quadruplex binding proteins

It is known that quadruplexes are recognized by different proteins and this raises the possibility that such proteins could serve not only to recognize the structure but also to promote or inhibit formation of quadruplex *in vivo*. At telomere level there are lots of proteins

able to bind the two different conformations of telomeres: the double strand and the quadruplex conformation (Shafer RH et al. 2001). For what concern the quadruplex structure proteins, human replication protein A and hnRNP A1 are able to bind to the quadruplex telomeric end and induce telomerase activation by unfolding the quadruplex (Salas TR et al. 2006, Zhang QS et al. 2006). Other than these proteins, components of telomeric complexes such as POT1 (Protection Of Telomere) and TriPeptidyl Peptidase 1 (TPP1) proteins are able to recognize the single and four-stranded DNA of telomeres and allow their elongation by telomerase (Zaug AJ et al. 2005). This happens not only in human chromosome ends but even in *Oxitricha nova* and *Stylonychia mytilis*, two kind of ciliated protozoa, in which homologous proteins of POT1 and TPP1, namely Telomere End-Binding Protein α (TEBP α) subunit and TEBP β subunit, are involved in telomere binding triggering telomerase activity (Wang F et al. 2007). Another protein of the telomere binding-proteins complex such as hnRNP D is able to bind, in this case with the Binding Domain 2 (BD2) subunit, to the quadruplex structure of the telomere and to unwind it (Enokizono Y et al. 2005).

Quadruplex binding proteins are found not only in telomeres but also in other part of the chromosomes. hnRNP A1 and its derivative Up1 for example are able to recognize and affect *KRAS* promoter sequence as we saw in the initial part of my PhD (Paramasivam M et al. 2009). Up1, the catalytic fragment of hnRNP A1 was identified in 1975 by Herrick et al. (Herrick G et al. 1976) in calf thymus as a protein that greatly destabilizes regions of DNA double helix by virtue of its selective affinity for single-stranded over double-stranded DNA. In particular proteins involved in triplet amplification diseases, such as the Bloom's syndrome and the Werner's syndrome, which are members of the RecQ family of helicases (BLM and WRN), are able to unwind DNA in the 3' to 5' direction in the presence of ATP and Mg^{2+} (Shafer RH et al. 2001). The functions of these enzymes involves unwinding quadruplex structures that may form as undesired intermediates in recombination or replication, which would explain the observed aberrant phenotype when they are defective (Shafer RH et al. 2001). Even in the fragile X syndrome expanded triplet some quadruplex-binding protein have been found to recognize this structure. Human quadruplex Telomeric DNA Binding Protein 42 (qTBP42) and unimolecular quadruplex Telomeric DNA Binding Protein 25 (uqTBP25) share close homology with the hnRNP-derived mammalian single-stranded DNA-binding proteins Up1 (Weisman-Shomer P et al. 2000) and mouse CarG-box binding factor A (hnRNP AB), both containing the RNP1 domain. They destabilize quadruplex structures

within the triplet amplification and this decreases the probability of expansion of the d(CGG)n sequence (Weisman–Shomer P et al. 2002).

The high degree of conservation of telomeric quadruplex-forming sequence and the existence of proteins that interact with quadruplexes and have telomere roles, such as RecQ family helicases, Ras proximate 1 (Rap1) and Pif1 of *Saccharomyces cerevisiae* and the corresponding mammalian TRF2, suggest that this feature may be an important component of telomere function (Smith JS et al. 2011). Pif1 protein has been studied in yeast and is able to recognize and bind quadruplex before and after replication because its regulatory function can be exerted in two steps: in the first moment it binds the quadruplex to unwind it, allowing cell to go throughout the cell cycle and then unwinds the quadruplex motifs again after replication as a sort of failsafe mechanism that makes sure that genome is free of quadruplex structures prior to chromosome condensation (Paeschke K et al. 2011).

Other than this, MyoD was one of the earliest proteins known to recognize quadruplexes. This protein is involved in the regulation of skeletal muscle development playing an important role in cellular differentiation (Shafer RH et al. 2001). In particular the intact basic part of MyoD is essential for the binding of E-box motif d(CANNTG) in promoters or enhancers of muscle-specific genes (Shklover J et al. 2007).

Another well known protein able to recognize and modulate the quadruplex structure folding or unfolding is Pur1 known also as Myc Associated Zinc-finger (MAZ) protein. This protein can bind to several promoters, such as the GAGA box in the rat insulin promoter, as well as the insulin-linked polymorphic region (ILPR) involved in genetic susceptibility to insulin-dependent diabetes mellitus (Shafer RH et al. 2001). MAZ is also able to recognize the *c-myb* promoter repressing its activity (Palumbo SL et al. 2008). Last but not least MAZ is able to bind *HRAS* promoter quadruplex-forming sequences and unwind them, as described in the final part of this thesis (Membrino et al. 2011).

Not only MyoD or MAZ where found to bind to different promoter sequences but other gene promoters require the intervention of quadruplex-binding proteins to activate or repress transcription of the downstream gene. For example in our laboratory *KRAS* promoter was found to present a quadruplex-forming sequence recognized by nuclear proteins, such as PARP-1, Ku 80, Ku 70 and hnRNP A1 (Cogoi S et al. 2008). Even Sp1 is able to bind to *KRAS* promoter region (Hoffman EK et al. 1990). NM23-H2 a metastases suppressor factor interacting with the c-*MYC* promoter is able to induce c-*MYC* expression unfolding the quadruplex structure created by the NHEIII motif (Thakur RK et al. 2009). In the promoter of

c-MYC other proteins such as hnRNP K and Sp1 are also able to regulate its transcription activity (Tomonaga T et al. 1996). Nucleophosmin, a nucleo-cytoplasmic shuttling phosphoprotein known to favour quadruplex formation, is able to recognize and bind *c-MYC* promoter quadruplex-forming region (Federici T et al. 1996). The nucleocapsid protein of the HIV-1 virus is also able to unfold quadruplex in the presence of Zn^{2+} (Kankia BI et al. 2005).
This amount of proteins that recognize quadruplexes support the evidence of an evolutionary pressure working to maintain quadruplex-forming sequences in the genome as regulatory elements of different cellular mechanisms.

2.1.3. Quadruplex DNA in promoter regions

Different studies have been done on quadruplex-forming sequences and their localization into the human genome. What came out from these analysis is that quadruplex structures can be possibly involved into regulatory mechanisms concerning replication, transcription and transduction. Most frequently quadruplex studies focused on promoter regions of different genes involved in cancers because they are hyperexpressed in different kinds of cancers. For example the first well studied promoter region has been the control element situated upstream of the P1 promoter for the *c-MYC* oncogene (Shafer RH et al. 2001). Even a polyguanine/polycytosine tract found in the proximal promoter of the Vascular Endothelial Growth Factor (VEGF) has been shown to form specific secondary structures, characterized as quadruplexes and I-motifs; the latter located in the complementary C-rich sequence (Guo K et al. 2008). The proximal 5'-flanking region of the human Platelet-Derived Growth Factor A (PDGF-A) promoter contains one Nuclease Hypersensitive Element (NHE) able to fold into quadruplex structures and modulate transcription (Qin Y et al. 2007). *c-KIT* oncogene presents a possible parallel-type quadruplex arrangement in its promoter region (Rankin S et al. 2005). Finally the *RAS* genes present quadruplexes in their promoter region such as *KRAS* (Cogoi S et al. 2004) and *HRAS* (present thesis, Membrino A et al. 2011), regulating transcription, and in the 5'UTR of the *NRAS* proto-oncogene which modulates translation (Kumari S et al. 2007).
Being involved in so many functional processes quadruplex structures could be seen as therapeutic targets. Cylene Pharmaceutical (San Diego, USA) has developed CX-3543 (Quarfloxin), which can disrupt the interaction between nucleolin and a quadruplex DNA

structure in the ribosomal DNA template, and this compound has now entered phase II clinical trials (Ou T et al. 2008). Stabilizing the quadruplex structure involved in transcription regulation can allow the blockage of transcription. For example Isoalloxazines can interact with the *c-KIT* oncogene (Ou T et al. 2008), 2,6-diamido anthraquinones (e.g. BSU-1051) (Hurley LH 2001), cationic porphyrins (e.g. TMPyP4) (Han H et al. 2000), perylenes (PIPER) (Neidle S et al. 2001), metal-free phthalocyanines (Azleer J et al. 2009, Membrino A et al. 2010), oxazole-based peptide macrocycles (Jantos K et al. 2006), TriArylPyridine ligand family (TAPs) which disrupts the quadruplex and enhance transcription of the *c-KIT* gene (Waller ZAE et al. 2009). All these different molecules can interact with quadruplexes by placing themselves on the top or the bottom of the structure, between tetrads or in the grooves (Neidle S et al. 2009). To investigate on new molecule properties of binding quadruplex DNA, a simple and high throughput technique is competition dialysis in which a nucleic acid structure is dialyzed against different kind of molecules (Ragazzon P et al. 2007). This has allowed to understand the binding affinity of some molecules used for quadruplex binding.

Given that quadruplex-forming sequences regulate gene transcription, and sometimes translation, are recognized by nuclear proteins and can be bound by artificial molecules, a strategy to modulate quadruplex-related genes is to use synthetic oligonucleotides resembling the sequence of interest. This would subtract proteins necessary to exploits the quadruplex function *in vivo*. Some G-rich libraries strongly inhibited cancer cell growth while sparing non-malignant cells (Choi EW et al. 2010) and many quadruplex-forming sequences have been created by SELEX technique and tested on cancer cell lines after an initial characterization of their features by CD (Paramasivan S et al. 2007), S1 hypersensitivity -given that Homopurin-homopyrimidine stretches adopting unusual DNA conformations are more sensible (Pestov DG et al. 1991)-, NMR and structural crystallography. Quadruplex-forming sequences, when introduced into cells, can be used as therapeutic agents. They competed for the binding between transcriptional factors and the host quadruplex structure of interest (Bates PJ et al. 1999). For example G-rich oligonucleotides which form intramolecular quadruplexes have been used to inhibit Stat3 protein avoiding its binding to DNA (Jing N et al. 2004). The most recent synthetic oligonucleotide that reached phase II clinical trial is AS1411 a quadruplex-forming oligonucleotide able to target nucleolin (Teng Y et al. 2007).

The advantages of using quadruplex oligonucleotides as therapeutic agents is that they are non-immunogenic, heat stable and they have increased resistance to serum nucleases and

enhanced cellular uptake compared to unstructured sequences (Bates PJ et al. 2009). Different techniques have been studied to enhance their delivery into cells such as modification of the oligonucleotide bases with lipophilic conjugation (Spiller DG et al. 1998) or with modified oligonucleotides as shown in this thesis (Membrino et al. 2011).

Already, quadruplexes are both targets for drug design and potential therapeutic drugs in their own right. Guanine quartets are suspected in some deoxyribozimes because of the K^+ dependence of the enzymatic activity. The discovery of fluorescent dyes that specifically target quadruplexes shows promise for locating and tracking their formation within the cell. There will be, in the future, more applications utilizing structures containing quadruplexes, particularly biologically active aptamers selected by systematic evolution of ligands by exponential enrichment (SELEX) techniques and structural studies will play a role in these applications. Not all future applications will be biomedical. Quadruplexes show promise as components for nanowires, ion channels and building blocks for directing the assembly of nanoscale components into sophisticated structures (Keniry MA 2001).

2.2. *RAS*

2.2.1. Bladder cancer

The specialized lining of the urinary tract, the urothelium, which extends from the renal pelvis to the urethra, is a source of the 5th most common cancer, responsible for approximately 3% of all cancer-related deaths in the United States. In the Western world, cigarette smoking is the most important risk factor, contributing to approximately 50% of all bladder cancers, followed by petrochemical and other industrial exposures. Bladder stone, chronic indwelling Foley catheters, schistosomiasis, and chemical irritations are well-documented risk factors (Dinney CP et al. 2004). The peak prevalence of the disease is among patients 60-70 years old (Vangeli D et al. 1996).

The human *RAS* genes represent an important prototypical family of cellular transforming genes originally identified in the T24 human urothelial cancer cell line (Reddy EP et al. 2004). Transformed cells exhibit changes in nuclear morphology and chromatin structure. Cancer is often diagnosed by an abnormal nuclear morphology. Also fibroblasts transformed

with oncogenes including *HRAS* have an abnormal nuclear morphology: the nuclei show a more rigid and spherical shape (Dunn KL et al. 2005). Two mechanisms for *RAS* gene transformation have been discovered. The most frequent is mutation of codons 12, 13, 59 or 61, which affect the enzymatic activity of the protein. In the second mechanism, internal splicing within the last intron mediates *RAS* gene expression. Mutations of the coding sequence of the *HRAS* gene, especially at codon 12, are relatively frequent, appearing in approximately 30-40% of urothelial malignancies (Dinney CP et al. 2004). *RAS* activation, via point mutation, overexpression or intensified signalling from Fibroblast Growth Factor Receptor 3 (FGFR3), occurs in 70-90% of bladder tumors in humans.

Cell growth is normally under the tight control of positive regulators (proto-oncogenes and growth factors) and negative regulators (senescence- and apoptosis-inducing molecules and tumor suppressors) of cell cycle progression. It is the intricate balance between these two opposing forces that ensures the necessary tissue renewal while maintaining homeostasis. Such a balance is not static, but rather is quite dynamic, particularly when cells are called upon to respond to various pathogenic insults. When an oncogene is activated by mutation, affected cells can turn up senescence and tumor suppressor genes to counter the oncogenic effects, thereby regaining cell-cycle control and putting a brake on cell proliferation. Failure to do so could lead to uncontrolled proliferation and tumorigenesis (Mo L et al. 2007).

2.2.2. Pancreatic cancer

Each year are diagnosed 28000 pancreatic tumours and almost all the affected people dye because of this tumour, making it one of the first causes of death in the world. 85% of the patients do not show any chance of intervention to solve the problem because of the formation of metastasis and the dramatic invasiveness of the tumor. The remaining part of the patients can barely hope in life expectation that does not exceed 5 years.

More than 90% of pancreatic cancers are related to the exocrine part of the organ. The origin of the neoplasias is due to point mutation that can be found into the genealogic trees of families with an high degree of pancreatic cancer. There are no doubts about the genetic origin of pancreatic cancers and their prognosis involves the accumulation of molecular abnormalities inside the ductal pancreatic cells.

In the tumorigenesis of pancreatic cancer are involved different suppressor genes that are altered so to be unable to control the cell cycle and so develop neoplasia. The most important of these repressor genes is p16 located in the 9p21 chromosome that is altered by means of mutations, deletions and hypermethylations of DNA. Another gene involved into the tumorigenic development of pancreatic cells is Deleted in Pancreatic adenoCarcinoma 4 (DPC4) located in chromosome 18q, inactivated only in 50% of cancers due to mutations and allelic deletions. Even p53, one of the most common genes inactivated in tumor cells, is inactivated in 50-70% of pancreatic cancers. The tumorigenic transformation of pancreatic cells can be due to genomic abnormalities of BRCA2 gene usually associated to breast and ovarian cancers.
In pancreatic cancer the most mutated gene is *KRAS* located in chromosome 12p12.1 being one of the first events in pancreatic tumor progression.

2.2.3. *RAS* family

RAS genes were originally identified in the mid-1960s as the transforming elements of Harvey and Kirsten strains of rat sarcoma viruses. Investigation of the biological properties of their protein products did not start until the early 1980s, when mutated alleles of cellular *RAS* genes were identified as dominant oncogenes in various types of tumors. Mutations within the *RAS* genes have been reported in diverse tumors including myeloid leukemia, colon, pancreatic, thyroid, and renal cell carcinomas (Burmer GC et al. 1989).
RAS genes are expressed at low level in most tissues (Jordano J et al. 1986). The *RAS* genes are expressed in a tissue-specific fashion: *HRAS* is highly expressed in skin and skeletal muscles, *KRAS* in colon and thymus and *NRAS* in male germinal tissue (Lory DR et al. 1993). *HRAS*, *KRAS* and *NRAS* have been found in a significant percentage of human tumors activated by somatic mutation of their corresponding normal proto-oncogene (Jordano J et al. 1986). The *RAS* genes have all similar structures and sequences with five exons, the first of which is not encoding, and conserved splicing sites. The introns, instead, have different lengths and sequences (Lory DR et al. 1993).

2.2.4. RAS proteins

The three mammalian *RAS* genes encode four highly related GTPases of 188 (K-RASB) or 189 (HRAS, NRAS, and K-RASA) amino acids in length. The A and B isoform of *KRAS* diverge solely in their COOH-terminal 25 amino acids (Campbell SL et al. 1998) generated by an alternative splicing of the fourth exon of this gene, and the abundance of *KRASB* transcripts is higher in comparison to that of *KRASA* transcripts (Bar-Sagi D 2001). The *RAS* family comprises other homologous proteins such as RRAS, TC21, RAP and RAL (Rebollo A et al. 1999). RAS proteins are ubiquitously expressed and located at the inner side of the plasma membrane (Bos JL 1998).

RAS proteins are membrane-bound guanosine triphosphate (GTP)/guanosine diphosphate (GDP)-binding (G) proteins that act to relay signals from the lipid-rich cellular membrane to the nucleus (Lee JT et al. 2002). RAS superfamily GTPases function as GDP/GTP-regulated molecular switches (Wennerberg K et al. 2005). RAS proteins exist in two conformations, a GTP-bound active state and a GDP-bound inactive state: the ratio of GTP to GDP bound to cellular RAS proteins is controlled by Guanine nucleotide Exchange Factors (GEFs) and GTPase-Activating Proteins (GAPs), the enzymatic activity of which responds to extracellular stimuli such as growth factors (Downward J 1998). All of the critical domains for GTPase function are present within the N-terminal 165 amino acids of RAS proteins. RAS proteins are made of three contiguous region. The first region encompasses the N-terminal 86 amino acids, which are 100% identical among the different RAS proteins. Within this regions lies the RAS effector binding domain (amino acids 32-40) which is the critical interaction site with all known downstream targets of RAS. The next 80 amino acids define the second region where mammalian RAS proteins diverge only slightly from each other, exhibiting an 85% homology between any protein pair. The remaining C-terminal sequence, known as the hypervariable region, starts at amino acid 165 and shows no sequence similarity among RAS proteins except for a conserved CAAX motif (C, cysteine; A, aliphatic amino acid; X, methionine or serine) at the very C-terminal end, which is present in all RAS proteins and directs post-translational processing (Bar-Sagi D 2001).

RAS proteins are post-translationally modified by prenylation, a process that involves the addition of a 15-carbon farnesyl isoprenoid moiety to the conserved CAAX motif by a farnesyl protein transferase (FPT). After prenylation, the C-terminal tripeptide is removed by proteolysis and the newly exposed C-terminal is methylated. In addition to post-translational

modifications, RAS proteins require binding of GTP to develop functional activity (Rebollo A et al. 1999).

RAS effectors preferentially bind to the GTP-bound state of RAS. Two regions of RAS differ in conformation between their GDP- and GTP-bound states. They are referred to as Switch I and Switch II. The Switch I region is composed of residues 30-38, which constitute part of the L2 loop and part of the β2 β strand. Within Switch I, residues 32-38 are conserved among all members of the RAS subfamily, so specificity determinants must lie outside this core of conserved residues. The Switch II region comprises residues 60-76 and is highly mobile, existing in multiple conformation (Marshall CJ 1996). The GTP-bound conformation exhibits increased affinity for downstream effectors (Mitin N et al. 2005).

RAS, discovered in two urothelial cancer cell lines (T24 and EJ), was the first named human oncogene in the early 1980. Although *HRAS* (as opposed to *KRAS* and *NRAS*) remains the main target of mutational activation in bladder tumors, the mutation frequency in different patients cohorts ranges from 0% to 84%, with no satisfactory explanation for such a wide variations. More than half of the human bladder tumors overexpress *RAS* mRNA and protein. In addition, several Receptor Tyrosine Kinases (RTKs) are believed to be constitutively active in human bladder tumors (Mo L et al. 2007). Not only do RTKs activate RAS-dependent pathways that drive proliferation, but they activate PI3K-dependent pathways which also contribute to the oncogenic mechanism. PI3K can initiate changes in gene transcription, cytoskeletal changes through β-catenin, changes in cell mobility through the tumor suppressor Adenomatous Polyposis Coli (APC) and phosphorylation of BAD, a protein involved in apoptotic and anti-apoptotic signalling (Porter AC et al. 1998).

Low-level expression of the activated HRAS could induce urothelial hyperproliferation, leading to simple urothelial hyperplasia. In a study conducted by Mo et al. after a long (>10 months) incubation with the activated form of HRAS, about 60% of transgenic mice develop low-grade, non-invasive papillary bladder tumors (Mo L et al. 2007). Such a long tumor latency and incomplete penetrance suggest that a rate-limiting step is present during persistent simple urothelial hyperplasia and that a cooperating event, such as the loss of a tumor suppressor gene must take place in order to efficiently transform the urothelium. Inactivating senescence-inducing molecules such as p16Ink4a and p19Arf, acting through Rb and p53, have been shown to be a prerequisite for cells to escape senescence and become transformed (Mo L et al. 2007).

Emerging evidence suggests that tumorigenicity of activated RAS in a specific tissue also depends on the spectrum and the extent of the signalling pathways that RAS exploits. The signal of activated RAS is propagated mainly through three structurally distinct, but functionally overlapping, signalling cascades: a. the Raf/Mek/Erk (MAPK) cascade, which promotes cell proliferation, b. the PI3K/AKT cascade, which mediates cell survival, and c. the Ral guanine nucleotide exchange factor (RalGEF)/Ral family GTPase, which was recently shown to be important for transformation and tumorigenesis (Mo L et al. 2007) (Figure 8).

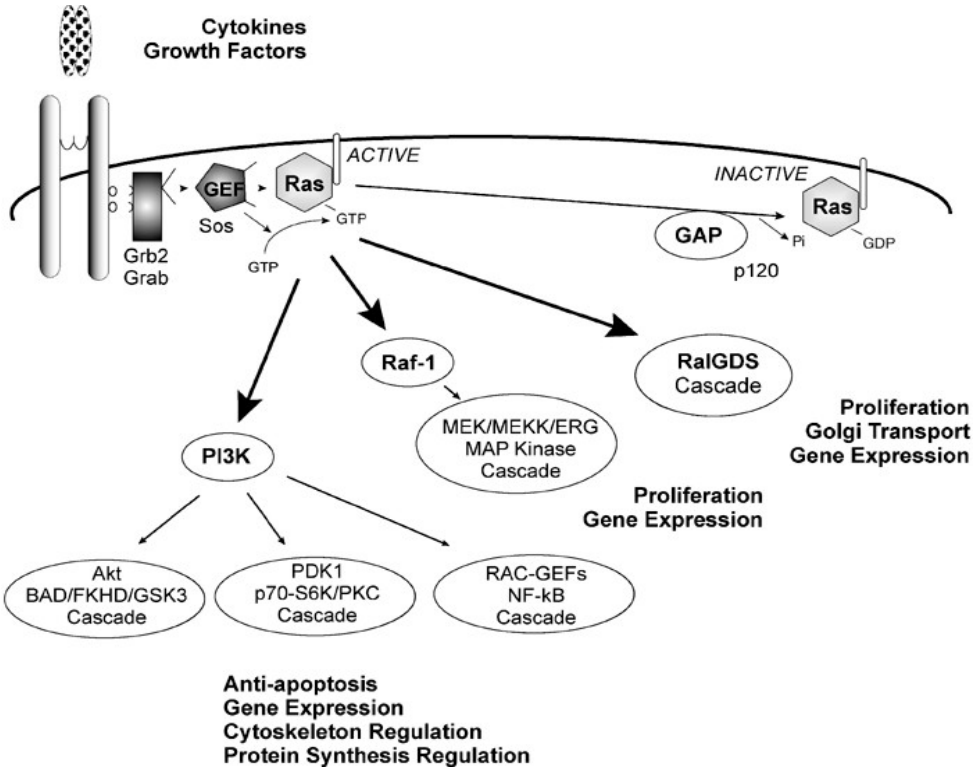


Figure 8

Perentesis JP, Bathia S, et al. **RAS oncogene mutations and outcome of therapy for childhood acute lymphoblastic** Leukemia 2004, 18, 685-692

Quantitative differences in the expression level of activated HRAS plays a crucial role in determining the time course of urothelial tumor development and high-level expression of an activated HRAS is sufficient to induce urothelial tumorigenesis along the low-grade, noninvasive phenotypic pathways (Mo L et al. 2007).

However, the measured expression levels in a tissue does not necessarily reflect the levels in cultured cells where the RAS mutation occurs (Bos JL 1989).

2.2.5. Strategies to inhibit RAS

Three main approaches to inhibit RAS-mediated signalling have been explored: a. prevention of membrane localization of RAS, b. inhibition of RAS protein expression via antisense nucleotides or RNAs and c. inhibition of kinases downstream of RAS (Lee JT et al. 2002).
In order for RAS to become active, it must undergo post-translational modifications that add farnesyl or geranylgeranyl isoprenoid moieties to its carboxy-terminus. These modifications are generated by three different enzymes, namely farnesyltransferase, geranyltransferase type I and type II. By blocking the enzymatic action of these enzymes, RAS is unable to localize to the cellular membrane where it mediates proliferative signaling.
The antisense approach uses oligonucleotides, which are complementary to mRNA transcripts of the *RAS* oncogene. These oligonuleotides hybridize to the complementary mRNAs and inhibit their subsequent translation into proteins.
The kinase inhibition approach would involve Raf-1 by the use of RNA aptamers or synthesizing proteins that are able to antagonize the binding between RAS and Raf or others such as Protein Kinase C, Mitogen-activated protein Kinase-1/2 (MEK-1/2), Extracellular signal-Regulated Kinase (ERK1 and 2).

2.2.6. *HRAS* promoter

As 80% of bladder tumors harbour *HRAS* mutations (Dinney CP et al. 2004) and more than half of bladder tumors overexpress *HRAS* (Vageli D et al. 1996), both mutation and overexpression are important factors in the tumorigenesis of bladder cancer (Mo L et al. 2007). *HRAS* mutations are less common, but they have a high prevalence in skin papillomas and urinary bladder tumors (Schubbert S et al. 2007). Indeed, it has been recently shown that low-level expression of constitutively active HRAS induces simple urothelial hyperplasia, while the doubling of activated *HRAS* oncogene triggers rapidly growing and penetrating tumors throughout the urinary tract. *HRAS* promoter contains numerous copies of the GGGCGG element. This G-box has been shown to interact with the Sp1 transcription factor (Ishii S et al. 1985, Ishii S et al. 1986). Upstream of the TSS there are runs of guanines spanning over three Sp1 sites, which are potential sites for quadruplex formation. We thus hypothesized that the G-rich elements might play a role in transcription regulation.

For what concerns *HRAS* in its –1 intron there is a repetitive element that may contribute to mature *HRAS* mRNA but its precise role in regulation of transcription is unknown (Lownder NF et al. 1990). Within this region *HRAS* contains a short GC-rich promoter which is sufficient to express the transformation activity of the *HRAS* oncogene. This element located between position –1418 and –1368 upstream of the first coding ATG is the minimum element necessary to activate *HRAS* transcription (Honkawa H et al. 1987).

The promoter of the human *HRAS* gene lacks typical TATA, CAAT boxes and contains an extremely high G+C content (80%) and multiple copies of the heptanucleotide GGGCGGG repeat (G-box). Because of this G-richness, in our laboratory we focused on two sequences namely hras-1 (435-462, accession number J00277) and hras-2 (506-530, accession number J00277) the first of which locates upstream the TSS and the latter spanning over it. These two sequences are potentially capable to fold into intramolecular quadruplex structures (Membrino A et al. 2011). According to a recent study, quadruplex-forming sequences covering Sp1 binding elements are present in many genes (Todd AK et al. 2008).

We have obtained a first hint that the *HRAS* promoter is structurally polymorphic while sequencing expression vectors bearing the wild-type or the mutated version of it. When primer-extension reactions were performed with primers complementary to the G-rich strand, Taq polymerase unexpectedly arrested at the G-rich elements. But with primers complementary to the C-rich strand no impediment was detected. An insight into the quadruplexes formed by the *HRAS* G-elements was obtained by DMS-footprinting, Circular Dichroism (CD) and Fluorescence Resonance Energy Transfer (FRET) experiments (Membrino A et al. 2011).

Focusing on the role of these two sequences we built up a construct harbouring the luciferase gene driven by *HRAS* promoter sequence and two related mutants each mutated in one of the two quadruplex-forming sequences. We had strong evidences that in the presence of wild type sequences, transcription is strongly inhibited. This inhibition was even stronger in the presence of quadruplex binding molecules such as TmPyP4 (Sun D et al. 2008) and phtalocyanines DIGP and DIGP-Zn (Membrino A et al. 2010).

This was the first hint that made us think at *HRAS* promoter structures as negative regulators of transcription.

To explore whether hras-1 and/or hras-2 are involved in transcription regulation, we first asked if these G-rich elements are recognized by nuclear proteins. Using an on-line software called MATInspector available on genomatix website, we checked whether hras-1 and hras-2

sequences were recognized by any nuclear factor. And we obtained as a result two highly probable binders of these sequences such as Myc Associated Zinc-finger (MAZ) protein and Specificity protein (Sp) 1. Both of them are transcriptional factors. MAZ is a 48,6 KDa zinc-finger protein whose consensus sequence is GGGAGGG, contains six C2H2-type zinc fingers at the C-terminus, a proline rich region and three alanine repeats and it is ubiquitously expressed, albeit at different levels in different human tissues (Song J et al. 2001). Sp1 is a 80,7 KDa protein belonging to the Sp/KLF transcription factor family (Krüppel-like factor) divided into two different transcriptional parts: the Sp group that prefers GC rich sequences and the KLF group that prefers GT rich sequences. Sp1 consensus sequence is GGGCGGG but can bind, even if less strongly, to GGGA/TGGG sequences (Wiestra I et al. 2008). Sp1 is ubiquitously expressed, and contains three C2H2-type zinc fingers, two serine and threonine-rich domains and two glutamine-rich domains; the C-terminus of Sp1 is involved in synergistic activation and interactions with other transcriptional factors (Song J et al. 2001).

To confirm the predicted binding we checked by EMSA and ChIP assays if those proteins bound hras-1 and hras-2 sequences, by which degree and to which structural conformation given that previous studies have shown that MAZ and Sp1 often bind to the same sequence regulating transcription in a cooperative way (Leroy C et al 2004, Song J et al. 2001).

Based on the results obtained we suggest a model explaining what we inferred from our experimental data concerning *HRAS* transcription.

HRAS promoter quadruplex-forming sequences usually act as repressors of transcription (Figure 9).

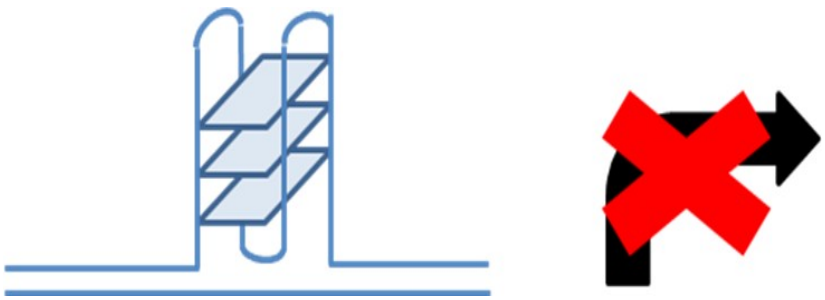


Figure 9

This effect can be enhanced by adding quadruplex stabilizers such as porphyrins or guanidinio phthalocyanines (Figure 10).

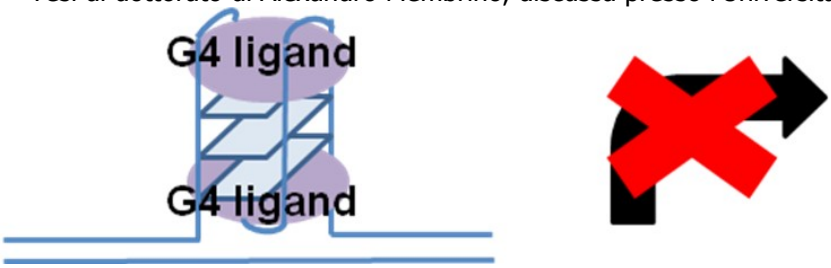


Figure 10

At the same time *HRAS* promoter quadruplexes are required to recruit essential transcription factors. Once these factors, such as MAZ and Sp1, bound the quadruplexes they can resolve the blocks and enhance transcription (Figure 11).

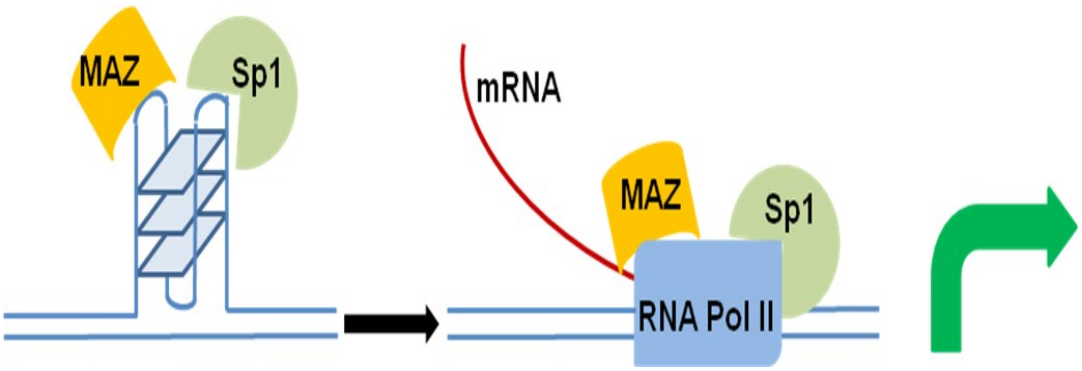


Figure 11

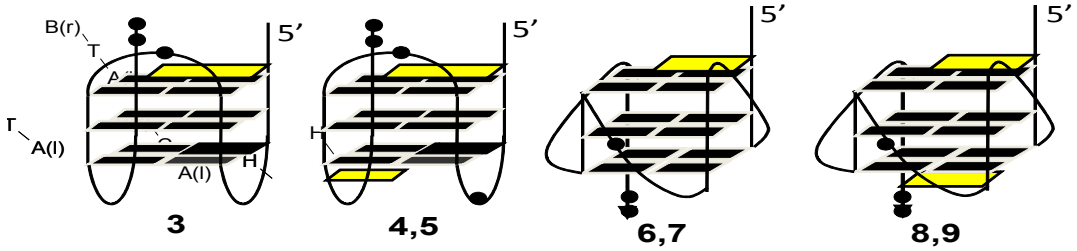
2.2.7. Affecting *HRAS* expression

Considering our model as a possible explanation of what happens in the cell we focused our attention on affecting *HRAS* transcription. Possible strategies are: a. stabilizing *HRAS* quadruplex structures or b. inhibiting nuclear quadruplex-binding factors.

The former option is achieved by the usage of porphyrins (TmPyP4) and phthalocyanines (DIGP and DIGP-Zn). They are able to bind and block *HRAS* promoter sequence into a quadruplex structure and this lead to inhibition of *HRAS* transcriptional levels (Figure 10).

The latter option can be achieved in two different ways: a. down-regulating MAZ and Sp1 genes by mean of siRNAs or b. competing the direct binding of MAZ and Sp1 to *HRAS*

quadruplexes using modified oligonucleotides (decoys) mimicking hras-1 and hras-2 sequences. For this purpose we designed G4-decoys oligonucleotides, mimicking the *HRAS* quadruplexes, with (R)-1-O-[4-(1-Pyrenylethynyl)phenyl-methyl] glycerol and LNA modifications to increase their stability and nuclease resistance (Figure 12). Decoy 3, 4 and 5 mimic hras-1 and decoy from 6 to 9 mimic hras-2.



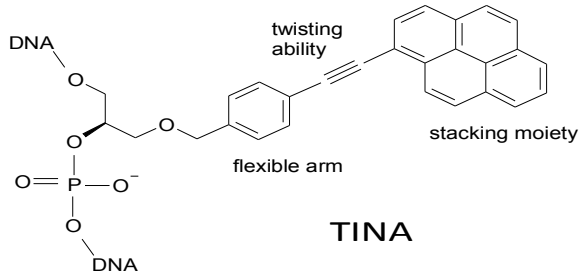


Figure 12

Exploiting the first way we saw that inhibiting MAZ and Sp1 led to *HRAS* down-regulation while using decoys MAZ and Sp1 proteins were sequestered from *HRAS* promoter. In this way transcription was abolished (Figure 13).

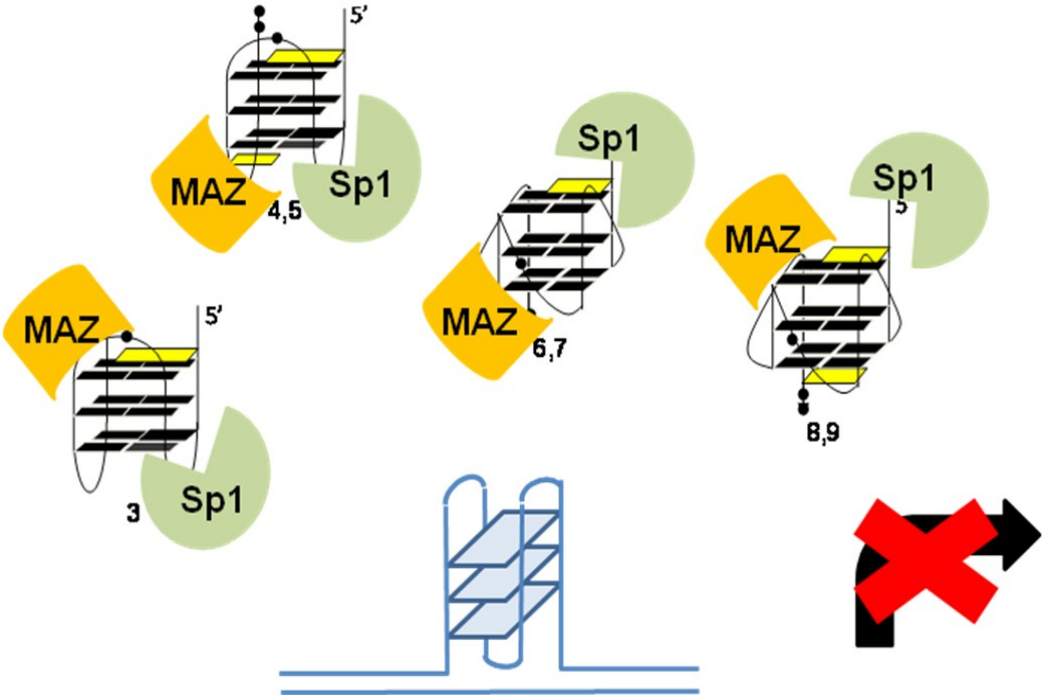


Figure 13

The designed decoys caused a potent antiproliferative effect in T24 bladder cancer cells, as they repressed transcription by sequestering MAZ and Sp1. Caspase and FACS assays showed that the G4-decoys promoted an antiproliferative effect mediated by apoptosis.

3. Aim of the work

The aim of my PhD was to study the impact of DNA quadruplex structures formed by critical G-elements located in the promoters of the *RAS* genes, in particular *KRAS* and *HRAS*.

The three-year work has mainly focused on the role played by quadruplex DNA in controlling transcription. In our laboratory, we found that *HRAS* and *KRAS* promoters show similar features: a. they lack TATA and CAAT boxes; b. they are rich in guanines; c. they contain runs of guanines in virtue of which the G-element can fold into an unusual quadruplex structure. Previous studies from our laboratory showed that the *KRAS* promoter is recognized by the nuclear factors hnRNP A1, Ku 70, Ku 80 and PARP-1.

In the first part of my PhD I have investigated the nature of the interaction between the proteins hnRNP A1 and UP1 (a proteolitic fragment of hnRNP A1) and quadruplex DNA. We setup a Fluorescence Resonance Energy Transfer (FRET)-melting technique to demonstrate that hnRNP A1/UP1 unfold the bound quadruplex.

In the second part of my PhD work I focused on *HRAS*: a proto-oncogene that is mutated in a big part of bladder cancers and that is for this reason a primary target of anticancer agents. We have studied for the first time how transcription is controlled in this gene and we have discovered that transcription is activated by the transcription factors MAZ and Sp1 and inhibited by quadruplex structures formed by promoter sequences overlapping the binding sites for MAZ and Sp1.

According to the transcription model proposed for *HRAS*, we have developed two novel strategies to down-regulate it in bladder cancer cells: a. use of decoy oligonucleotides that, mimicking the natural quadruplexes formed by the *HRAS* promoter sequences, bind and subtract transcription factors necessary to activate transcription; b. use of quadruplex ligands such guanidinio phthalocyanines that, stabilizing quadruplex DNA, lock the *HRAS* promoter in the folded and not active form.

With the results of this work four articles have been published in high impact-factor journals. Details of the work are given in the "Results" session of this thesis.

4. Results

4.1. Protein hnRNP A1 and its derivative Up1 unfold quadruplex DNA in the human *KRAS* promoter: implications for transcription.

Previous studies done in our laboratory involving human *KRAS* proto-oncogene showed that it contains a structurally polymorphic nuclease hypersensitive element (NHE) whose purine strand forms a parallel quadruplex structure (called 32R). From pull-down experiments published in previous work done by Xodo and co-workers (2008) our laboratory reported that quadruplex 32R can be recognized by three nuclear proteins: PARP-1, Ku70 and hnRNP A1. We focused on hnRNP A1 (A1) and its derivative Up1 producing them as recombinant proteins and we started working on their ability to interact with *KRAS* promoter quadruplex. We performed some CD analysis of the 32R sequence folded as a quadruplex in the presence of the two proteins. The results obtained demonstrate that these proteins strongly reduce the intensity of the 260 nm ellipticity. This means that these two proteins are able to unfold the *KRAS* quadruplex. FRET-melting experiments revealed that A1/Up1 completely abrogates the quadruplex structure of *KRAS* promoter. This was thought to be the initial step towards the annealing between the polypurine strand and its complementary sequence into the cell to inhibit transcription. When quadruplex 32R is stabilized by TMPyP4, hnRNP A1/Up1 brings about only a partial destabilization of the quadruplex structure. Based on this we suggested a possible regulatory model for *KRAS* transcription. In this model we stated that *KRAS* promoter is in equilibrium between the double strand and the quadruplex conformation. When the quadruplex is folded, transcription is allowed due to protein recognition of this unusual structure of DNA. Starting from this hypothesis we suggested two strategies to down-regulate *KRAS* transcription. This should be achieved: a. using quadruplex ligand molecules that stabilize *KRAS* promoter into the quadruplex conformation or b. using decoy molecules that should sequester quadruplex binding proteins.

Published online 12 March 2009 *Nucleic Acids Research, 2009, Vol. 37, No. 9* ***2841–2853***
doi:10.1093/nar/gkp138

Protein hnRNP A1 and its derivative Up1 unfold quadruplex DNA in the human KRAS promoter: implications for transcription

Manikandan Paramasivam[1], Alexandro Membrino[1], Susanna Cogoi[1], Hirokazu Fukuda[2], Hitoshi Nakagama[2,] and Luigi E. Xodo[1,*]

[1]Department of Biomedical Science and Technology, School of Medicine, P.le Kolbe 4, 33100 Udine, Italy and [2]Biochemistry Division, National Cancer Center Research Institute, 1-1, Tsukiji 5, Chuo-ku, Tokyo 104-0045, Japan
Received December 9, 2008; Revised and Accepted February 18, 2009

ABSTRACT

The promoter of the human KRAS proto-oncogene contains a structurally polymorphic nuclease hypersensitive element (NHE) whose purine strand forms a parallel G-quadruplex structure (called 32R). In a previous work we reported that quadruplex 32R is recognized by three nuclear proteins: PARP-1, Ku70 and hnRNP A1. In this study we describe the interaction of recombinant hnRNP A1 (A1) and its derivative Up1 with the KRAS G-quadruplex. Mobility-shift experiments show that A1/Up1 binds specifically, and also with a high affinity, to quadruplex 32R, while CD demonstrates that the proteins strongly reduce the intensity of the 260 nm-ellipticity—the hallmark for parallel G4-DNA—and unfold the G-quadruplex. Fluorescence resonance energy transfer melting experiments reveal that A1/Up1 completely abrogates the cooperative quadruplex-to-ssDNA transition that characterizes the KRAS quadruplex and facilitates the association between quadruplex 32R and its complementary polypyrimi-dine strand. When quadruplex 32R is stabilized by TMPyP4, A1/Up1 brings about only a partial desta-bilization of the G4-DNA structure. The possible role played by hnRNP A1 in the mechanism of KRAS transcription is discussed.

INTRODUCTION

The mammalian KRAS gene encodes for a guanine nucleotide-binding protein of 21 kDa that activates several cellular pathways controlling important events such as proliferation, dierentiation and signalling (1). The Ras proteins behave as a molecular switch cycling between inactive GDP-bound and active GTP-bound states. The state of nucleotide occupancy is regulated by specific pro-teins named guanine nucleotides exchange factors (GEFs) and GTPases activating proteins (GAPs) (1,2). The RAS genes are frequently mutated in solid and haematological neoplasias with single point mutations at exons 12, 13 and 61 (3,4). The most common mutated RAS gene in solid tumours is KRAS, with a 90% incidence in pancreatic adenocarcinomas (5,6). As the mutated Kras protein has a defective GTPase activity, it is not inactivated by GAPs

(7). It remains locked into the GTP-bound active state which continuously transmits to the nucleus mitotic signals that contribute to the neoplastic phenotypes in cancer cells (8–10). As pancreatic adenocarcinomas are refractory to conventional treatments, the discovery of new drugs capable to sensitize tumour cells to chemotherapy is being pursued in many laboratories. In our laboratory, we focused on KRAS and in order to design anti-KRAS drugs we investigated how the transcription of this proto-oncogene is controlled. Previous studies have shown that a nuclease hypersensitive element (NHE), located in the KRAS promoter upstream of the transcription start between –327 and –296, is responsible for most of the transcription activity (11). Earlier we reported that the purine strand of NHE is structurally polymorphic, as its tract of sequence recognized by nuclear proteins is able to fold into stable G-quadruplex structures (12,13). Using the purine strand of NHE (called 32R) in quadruplex confor-mation as a bait, we pulled down from a pancreatic nuclear extract three proteins with anity for the KRAS quadruplex. By SDS–PAGE and mass spectrometry, we identified these proteins as poly[ADP-ribose] polymerase 1 (PARP-1), ATP-dependent DNA helicase 2, subunit 1 (Ku70) and heterogeneous ribonucleoprotein A1 (hnRNP A1) (13). Protein hnRNP A1 (from now on A1) is a member of the heterogeneous ribonucleoprotein family, which is highly abundant in the nucleus of actively

*To whom correspondence should be addressed. Tel: +39 432 494395; Fax: +39 432 494301; Email: luigi.xodo@uniud.it

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

growing mammalian cells (14,15). All members of the hnRNP family are characterized by two highly conserved RNA recognition motifs (RRMs) at the N-terminus and by a glycine-rich domain at the C-terminus (16,17). Although a recent structure of a co-crystal of Up1 (a pro-teolitic portion of A1 retaining binding activity) bound to the telomeric repeat $(TTAGGG)_2$ suggests that both RRMs interact with DNA (18), it has been reported that only one motif (RRM1) is sucient for strong and specific binding to single-stranded telomeric DNA (19) and that its sub-element RNP11 mediates destabilization of quadru-plex $(CGG)_n$ (20). Proteins hnRNP play various roles in mRNA metabolism (14,15) and in the biogenesis of telo-meres (21). As protein A1 (and its derivative Up1) was reported to have a telomere-lengthening eect in erytro-leukemia cells (21,22), it is suspected to function as an auxiliary factor of the telomerase holoenzyme (23). Considering that the 3^0 G-rich repeats of the telomeres are folded in stable G-quadruplex structures, it has been hypothesized that A1 stimulates telomere elongation by disrupting high-order structures formed by the telomere repeats. Indeed, Up1 was reported to destabilize the bimo-lecular quadruplex formed by human telomere repeats d(TTAGGGTTAGGG), $d(TTAGGG)_4$ and the intramo-lecular quadruplex of $d(GGCAG)_5$ (23–25).

Since we discovered that A1 is associated to the KRAS promoter, in this study we have investigated the inter-action between recombinant A1/Up1 and the KRAS G-quadruplex. Electrophoretic mobility shift assay (EMSA) showed that A1/Up1 binds to the KRAS quad-ruplex with high anity and specificity, while CD and fluorescence resonance energy transfer (FRET) experi-ments revealed that A1/Up1 destabilizes this non B-DNA structure of the KRAS promoter. The results of our study support a transcription mechanism in which A1 should function as a G-quadruplex destabilizing protein, as it seems to occur in the G-rich 3^0 overhang strand of the telomeres (23). In conclusion, this study sheds some light on the mechanism of KRAS transcription regulation and may be useful for the rationale design of anticancer drugs specific for oncogenic KRAS.

MATERIALS AND METHODS

DNA and proteins hnRNP A1/Up1

The oligonucleotides used in this study (Table 1) were obtained from MWG (Germany) and Microsynth (Switzerland). They have been purified by 20% PAGE (acrylamide: bisacrylamide, 19:1) in TBE, under denatur-ing conditions (7 M urea, 558C). The bands were excised from the gel and eluted in water. The DNA solutions were filtered (Ultrafree-DA, Millipore) and precipitated. DNA concentration was determined from the absorbance at 260 nm of the oligonucleotides diluted in milli Q water, using as extinction coecients 7500, 8500, 15 000 and 12 500 M^{-1} cm^{-1} for C, T, A and G, respectively. Dual-labelled F-32R-T (5^0 end with FAM, 3^0 end with TAMRA) were HPLC-purified.

Recombinant proteins Up1 and A1 tagged to GST were expressed in Escherichia coli BL21 using plasmids pGEX-Up1 and pGEX-hnRNP A1. After transformation, the bacteria were grown for 2 h at 378C with 50 mg/ml ampicillin to an A_{600} of 0.5–2.0 prior to induction with IPTG (100 mM final concentration). Cells were allowed to grow for 7 h before harvesting. The cells were centrifuged at 5000 r.p.m., 48C. After centrifugation the supernatant was removed carefully and the cells washed twice with PBS. The pellet was re-suspended in a solution of PBS with PMSF 100 mM and DTT 1 M. The bacteria were lysed by sonication, added with Triton X-100 (1% final concentration) and incubated for 30 min on a shaker at room temperature. The lysate was then centrifuged for 10 min at 48C at 10 000 r.p.m. Glutathione Sepharose 4B (GE Healthcare) (50% slurry in PBS) was added to the supernatant and incubated for 30 min at 48C on a shaker. The mix was centrifuged for 5 min at 500 g and the pellet was washed 5 times in PBS and eluted with elution buer containing 20 mM NaCl, 20 mM reduced glu-tathione, 200 mM Tris–HCl, pH 9.5 for A1 elution and pH 7.5 for Up1 elution. Alternatively, to remove the GST tag, the mix was centrifuged for 5 min at 500g, washed with PreScission Cleavage buer (GE Healthcare) and centrifuged 5 min at 500g. The pellet was incubated for 4 h at 48C with PreScission protease to cleave the GST tag from the purified proteins. After PreScission cleavage, the A1 or Up1 moieties were detached from GST which remained bound to the Gluthatione Sephadex beads. The reaction mixtures were centrifuged for 5 min at 500g, 48C, and the untagged proteins collected from the supernatant. Finally, the purification of tagged and untagged Up1 and A1 proteins were checked by SDS–PAGE.

CD and fluorescence experiments

CD spectra have been obtained with a JASCO J-600 spec-tropolarimeter equipped with a thermostatted cell holder. CD experiments were carried out with oligonucleotides (3 mM) in 50 mM Tris–HCl, pH 7.4, 100 mM KCl. Spectra were recorded in 0.5 cm quartz cuvette. A ther-mometer inserted in the cuvette holder allowed a precise measurement of the sample temperature. The spectra were calculated with J-700 Standard Analysis software (Japan Spectroscopic Co., Ltd) and are reported as ellipticity (mdeg) versus wavelength (nm). Each spectrum was recorded three times, smoothed and subtracted to the baseline.

Fluorescence measurements were carried out with a Microplate Spectrofluorometer System (Molecular Devices) using a 96-well black plate, in which each well contained 50 ml of 200 nM dual-labelled F-32R-T in 50 mM Tris–HCl, pH 7.4 and KCl as specified in the figure captions. Before adding the protein, the samples were incubated for 24 h at room temperature in the speci-fied buer. The protein (Up1, A1 or BSA) was added 30 min before fluorescence analysis. The emission spectra were obtained by setting the excitation wavelength at 475 nm, the cut-o at 515 nm and recording the emission from 500 to 650 nm. Upon addition of KCl, F-32R-T assumes a folded quadruplex conformation and FRET is expected between the 5^0 and 3^0 fluorophores. The emission intensity of the donor (FAM) decreases while the intensity of the acceptor increases, correspondingly, as K^+ is added to the sample solution. The energy transfer from the donor to the acceptor and vice versa can be

empirically represented by the parameter P:

$$P = \frac{I_D}{(I_D + I_A)}$$

where I_D and I_A are the intensities of the donor and acceptor (26,27). Fluorescence melting experiments were performed on a real-time PCR machine (iQ5, BioRad), using a 96-well plate filled with 50 ml solutions of dual-labelled F-32R-T. The protocol used for the melting experiments is the following: (i) equilibration step of 5 min at low temperature (158C); (ii) stepwise increase of the temperature of 18C per min for 76 cycles to reach 958C. All samples in the wells were melted in 76 min.

Kinetic experiments were carried out using the iQ5 real-time machine. Oligonucleotide F-32R-T (200 nM) in 100 mM KCl, i.e. in the quadruplex conformation, was mixed with the complementary 32Y strand and the increase at 525 nm of the fluorescence was measured as a function of time. The experiment was also performed adding to F-32R-T a mixture containing 32Y (8-fold) and Up1 (400 nM). The increase of fluorescence $F = F - F_0$, where F_0 and F is the fluorescence at 525 nm (FAM) at t = 0 and at any time t, was best-fitted to a single or double-exponential curve. The half-life of the reaction is given by $t_{1/2} = 0.693/k$.

EMSA

Oligonucleotides 32R, HRAS-1, HRAS-2, CMYC, CKIT, VEGF, 32Y, Gmut1 and Gmut2 were end-labelled with [g-^{33}P]ATP and T4 polynucleotide kinase. Duplex dsNHE was prepared annealing (10 min at 958C, over-night at room temperature) a mixture containing equimo-lar amounts of radiolabelled 32R and complementary 32Y in 50 mM Tris–HCl, pH 7.4, 100 mM NaCl. Before EMSA, the quadruplex-forming oligonucleotides were allowed to form their structure in 50 mM Tris–HCl, pH 7.4, 100 mM KCl, 378C (overnight incubation). Radiolabelled oligonucleotides (35 nM) were treated for 30 min at room temperature with dierent amounts of A1/Up1, (r ([protein]/[oligonucleotide]) ratios are specified in Figure 3) in 20 mM Tris–HCl, pH 8, 30 mM KCl, 1.5 mM $MgCl_2$, 1 mM DTT, 8% glycerol, 1% Phosphatase Inhibitor Coktail I (Sigma, Milan, Italy), 5 mM NaF, 1 mM Na_3VO_4, 2.5 ng/ml poly [dI-dC]. After incubation, the reaction mixtures were loaded in 8% TBE (1_) polyacrylamide gel, thermostatted at 168C. After running the gel was dried and exposed to autoradiography (G E Healthcare, Milan) for 24–36 h at –808C.

Polymerase-stop assay

A linear DNA fragment of 87 nt, containing the G-rich element of NHE, was used as a template for Taq polymerase primer-extension reactions. This DNA sequence was purified by PAGE. The template (100 nM) was mixed with the labelled primer (50 nM) in 100 mM KCl, Taq buer 1_ and overnight incubated at 508C. The primer extension reactions were carried out for 1h, by adding 10 mM DTT, 100 mM dATP, dGTP, dTTP, dCTP and 3.75 U of Taq polymerase (Euro Taq, Euroclone, Milan). The reactions were stopped by adding an equal volume of stop buer (95% formamide, 10 mM EDTA, 10 mM NaOH, 0.1% xylene cyanol, 0.1% bromophenol blue). The products were separated on a 15% polyacrylamide sequencing gel prepared in TBE, 8 M urea. The gel was dried and exposed to autoradiography. Standard dideoxy sequencing reac-tions were performed to detect the points in which DNA polymerase I was arrested.

RESULTS

We previously demonstrated that the G-rich strand of NHE can form G-quadruplex structures (13,28). By means of CD and DMS-footprinting experiments we found that the G-tract called 32R forms a parallel G-quadruplex characterized by three G-tetrads (T_m of 708C in 100 mM KCl) (Figure 1). Pull-down assays with a pancreatic nuclear extract combined to mass spectrome-try showed that quadruplex 32R binds to three proteins: PARP-1 (116 kDa), Ku70 (72 kDa) and A1 (34 kDa) (13). Since A1 is involved in the biogenesis of the telomeres as a G4-DNA destabilizing protein (23) and is able to disrupt the secondary structures of the hypervariable minisatellite sequence $d(GGCAG)_5$ (24), we asked whether A1/Up1 can have a similar functional role in the human KRAS promoter. To address this question, recombinant A1 and its derivative Up1 were expressed in Escherichia coli as proteins fused to GST and purified by anity chromatog-raphy with glutathione sepharose 4B. The GST moiety was removed with a pre-scission protease and recombi-nant tagged and untagged proteins were obtained with a high purity level (Figure 2). Up1 is a proteolitic fragment (195 aa) of A1 (319 aa) that retains the two RNA-recogni-tion motifs (RRMs) responsible for binding to nucleic acids (18,22).

The interaction between A1/Up1 and a variety of DNA substrates, some of which were structured in G4-DNA and some not, was analysed by EMSA. ^{33}P-labelled 32R (35 nM) was first incubated for 24 h in 100 mM KCl to allow quadruplex formation, then incubated for 30 min with increasing amounts of Up1 or A1: r ([protein]/ [32R]) = 0, 0.5, 1, 2, 5, 10, 20, 50, 100. As preliminary experiments showed that GST-tagged and untagged pro-teins behave in the same way, we performed EMSA with the tagged proteins. Figure 3a and b shows that quadru-plex 32R forms with A1/Up1 a DNA–protein complex that, being detected even at r = 0.5, should have a 1 : 1 stoichiometry. In addition, for r > 20, another slow-migrating DNA–protein complex appears in the gel, most likely due to a 1:2 complex. When r was increased to 200 and the samples run in a longer gel, 32R migrated essentially as 1:2 complex (Figure 3c). The formation of two DNA–protein complexes by A1/Up1 is in keeping with the results of Zhang et al. (23) and the crystal struc-ture of $d(TTAGGG)_2$ bound to Up1 (18). Since a tract of 12 nt functions as a minimum binding unit, 32R has

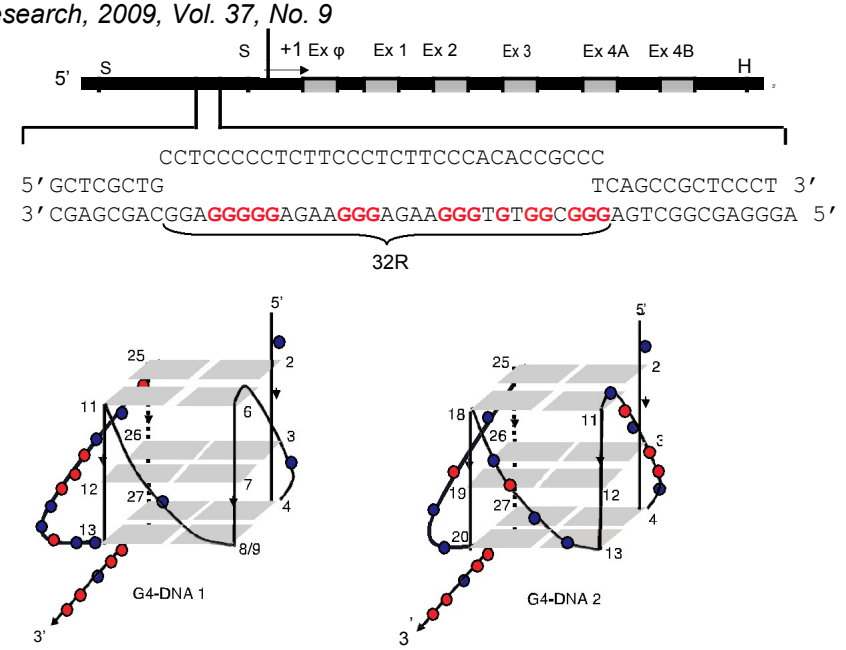
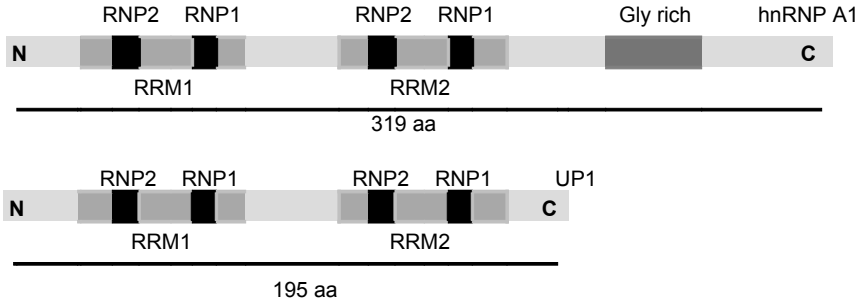


Figure 1. Sequence of the nuclease hypersensitive element (NHE) in the human KRAS promoter. The G-rich sequence 32R forms a G-quadruplex whose putative structure, consistent with CD and dimethyl sulfate footprinting, is G4-DNA1, which is characterized either by a flipped-out thymidine connecting G7 to G9 or a GGGT triad (13). The expected G4-DNA2 structure is not supported by dimethyl sulfate footprinting. The nucleotides of 32R (Table 1) are numbered from the 5^0-end.



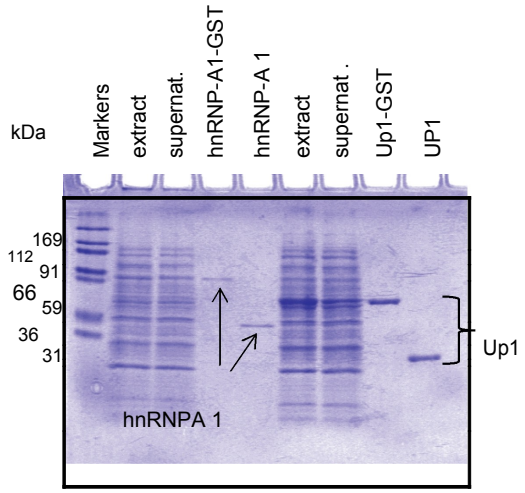


Figure 2. Schematic representations of proteins hnRNP A1 and Up1. The two RNA-recognition motifs (RRMs), that mediate ssDNA binding, contain each two conserved RNP2 and RNP1 submotifs. Up1 encompasses the amino-terminal two-third of the hnRNPA1 sequence. SDS–PAGE of GST-tagged and untagged hnRNP A1 and Up1, after gluthatione sepharose 4B purification. Lane 1, protein markers; lane 2, total extract (hnRNP A1); lane 3, supernatant; lane 4, purified GST-tagged hnRNP A1; lane 5, purified untagged hnRNP A1; lane 6, total extract (Up1); lane 7, supernatant; lane 8, purified GST-tagged Up1; lane 9, purified untagged Up1.

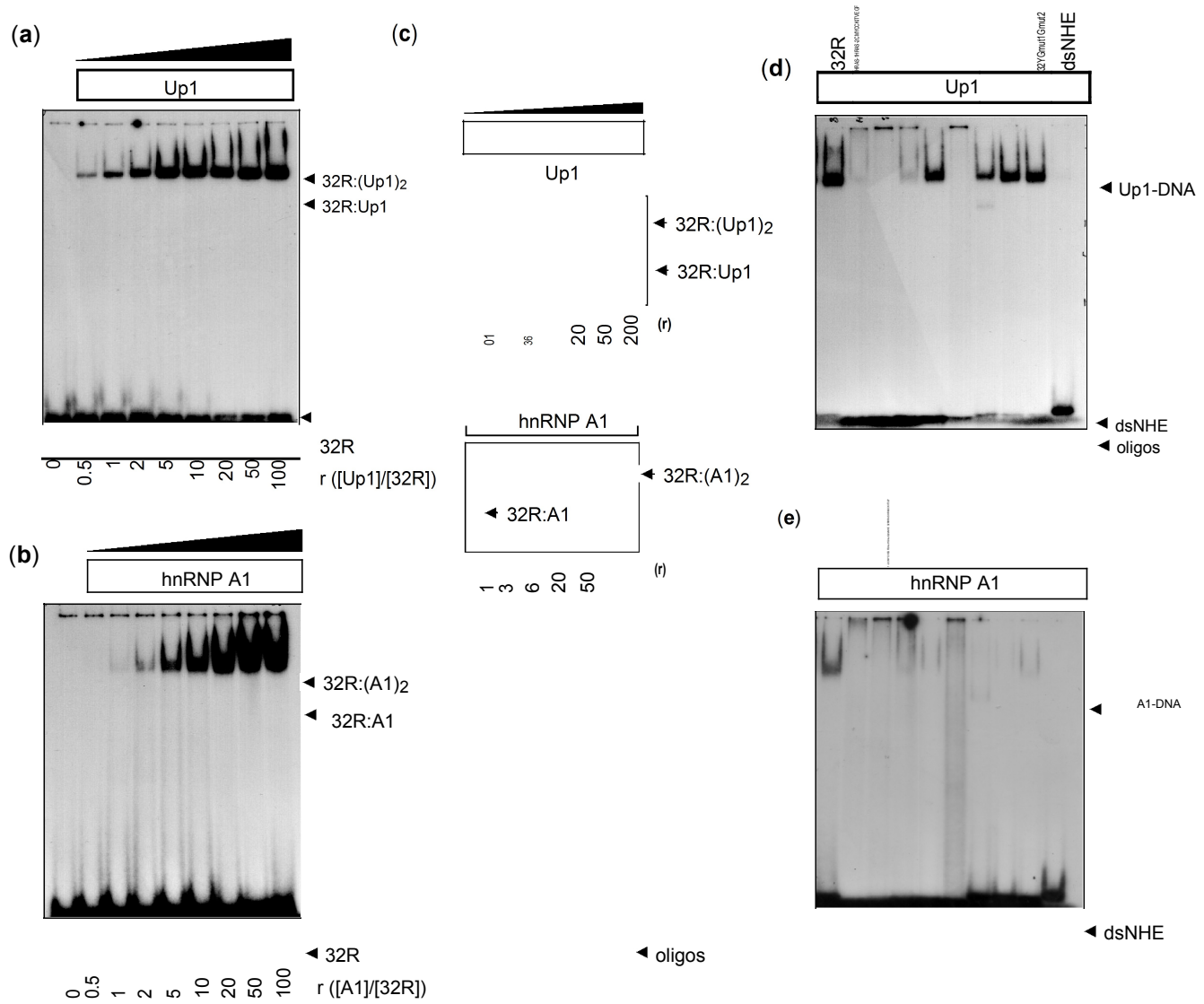


Figure 3. (a, b) EMSA of 35 nM ^{33}P-labelled quadruplex 32R after 30 min incubation with increasing amounts of Up1 or A1 at the specified r values, in 20 mM Tris–HCl, pH 8, 30 mM KCl, 1.5 mM $MgCl_2$, 1 mM DTT, 8% glycerol, 1% Phosphatase Inhibitor Coktail I (Sigma), 5 nM NaF, 1 mM Na_3VO_4, 2.5 ng/ml poly dI-dC, for 258C. The analyses were carried out in 8% polyacrylamide gel (29:1) in TBE (1_) at 168C. Before the EMSA, 32R was incubated overnight in 100 mM KCl to get it in the G-quadruplex conformation; (c) EMSA as in (a, b) but with r values up to 200; (d, e) EMSA of A1/Up1 mixed to various DNA substrates [G-quadruplexes 32R, HRAS-1, HRAS-2, CMYC, CKIT, VEGF, dsNHE (32R:32Y) and unstructured oligonucleotides Gmut1, Gmut2, 32Y]. PAGE carried out in 8% polyacrylamide gel (29:1) in TBE (1_) at 168C.

potentially two binding sites, which can in principle form two DNA–protein complexes by binding one or two protein molecules. By quantifying the intensity of the electrophoretic bands, we roughly estimated that the dissociation constant K_d of the 1:1 complex is about 50 nM for Up1 and 200 nM for A1. We also tested the binding specificity of A1/Up1 for a variety of well known G-quadruplex structures obtained from CMYC, CKIT, VEGF and HRAS promoter sequences (29–32) (for HRAS quadruplexes, see Supplementary Data S1) (Figure 3d and e). The various DNA substrates have been ^{33}P-labelled and treated with an excess of protein (r = 50). It can be seen that A1 shows good specificity for the KRAS quadruplex, as it does not bind to the other quadruplex-forming sequences, unstructured oligo-nucleotides Gmut1, Gmut2, 32Y (the complementary polypyrimidinic NHE strand) and dsNHE (32R:32Y) (Table 1). Instead, protein Up1, besides quadruplex 32R, shows affinity also for the CKIT quadruplex and unstructured oligonucleotides.

To analyse the effect of A1/Up1 on the KRAS G-quadruplex, we could not employ electrophoresis because the mobility between an intramolecular quadruplex and its unfolded form is not very different. Therefore, we used spectroscopic techniques such as circular dichroism (CD) and FRET. Figure 4 shows that in 100 mM KCl, 32R is characterized by a CD signature typical of a parallel G-quadruplex: a strong and positive ellipticity at 260 nm and a weak and negative ellipticity at 240 nm (33). When quadruplex 32R is denatured by increasing the temperature, the positive 260 nm band is dramatically reduced and its spectrum becomes similar to that of unstructured oligonucleotides (data not shown). Thus, the structural transition from quadruplex-to-ssDNA is accompanied by a strong reduction of the 260 nm ellipticity. A similar transition was obtained by adding to

quadruplex 32R increasing amounts of A1/Up1 (r = 1, 2, 4, 6). It can be seen that the protein causes a progressive reduction of the 260 nm ellipticity, indicating that the G4-DNA structure is unfolded by the protein. As a con-trol, we treated quadruplex 32R with an unrelated protein, the trypsinogen inhibitor, and found that the 260 nm ellip-ticity was not aected and remained constant at all protein concentrations used. The CD spectra of Up1 at increasing concentrations show that the protein between 240 and 320 nm does not have any negative band, but below 240 nm it shows a negative band typical of the polypeptide

Table 1. Oligonucleotides ($5^0 > 3^0$) used in this study

AGGGCGGTGTGGGAAGAGGGAAGAGGGGGAGG 32R F-
AGGGCGGTGTGGGAAGAGGGAAGAGGGGGAGG-T F-32R-T
AGGGAGGGCGCTGGGAGGAGGG CKIT
GGGCGGGCCGGGGGCGGGTCCCGGCGGGG VEGF
TGGGGAGGGTGGGGAGGGTGGGGAAGG CMYC
TCGGGTTGCGGGCGCAGGGCACGGGCG HRAS-1
CGGGGCGGGGCGGGGGCGGGGGCG HRAS-2
GCGGTGTGTGAAGAGTGAAGAGTGGGATGCAG Gmut1
GCATTCTGATTACACGTATTACCTTCACTCCA Gmut2
CCTCCCCCTCTTCCCTCTTCCCACACCGCCCT 32Y
GTACTACACTTGATA primer
ACCTTGATGAATCCAGGGCGGTGTGGGAAGAG template
GGAAGAGGGGGAGGAATCGCTACCGTTAAGCA
TCGATCATATCAAGTGATAGTAC

F: FAM; T: TAMRA.

backbone. The CD data showing G-quadruplex unfolding are in keeping with those previously obtained with the telomeric TTAGGG repeat (24,25) and the hypervariable minisatellite sequence $d(GGCAG)_5$ treated with Up1 (24).

The unfolding of the human KRAS quadruplex by A1/Up1 was also investigated by FRET, using the quad-ruplex-forming sequences tagged at the 5^0 and 3^0 ends with FAM (donor) and TAMRA (acceptor) (34). By exciting F-32R-T at 475 nm, the emission intensity of the donor at 525 nm decreases while the emission intensity of the acceptor at 580 nm increases, as the KCl concentration is increased from 0 to 140 mM (Supplementary Data S2). F-32R-T folded in the G-quadruplex conformation (T_m = 758C in 140 mM KCl) is characterized by a P-value of 0.52 (see 'Materials and methods' section). This P-value is higher than that observed for the quadru-plex formed by the human telomeric repeat d(GGGTTAG GGTTAGGGTTAGGG) (26), because F-32R-T forms a parallel quadruplex where the two fluorophores are at opposite ends of the structure (13). When the G-quadruplex is destabilized by scaling down the KCl concentration to zero or by adding the complementary 32Y strand, that transforms the G-quadruplex into a B-DNA duplex where the donor and acceptor are sepa-rated by about 115 A, the donor fluorescence significantly increases (for instance, from spectrum 2 to spectrum 1, Figure 5a) and the P-value becomes 0.75. This means that the unfolding of quadruplex F-32R-T is accompanied

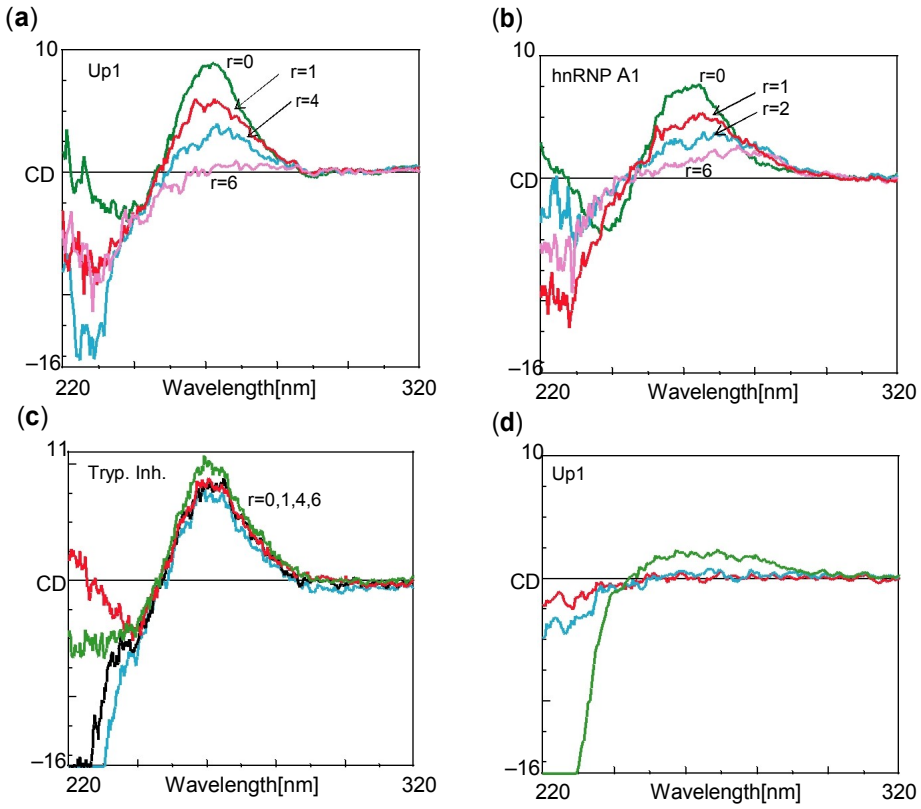


Figure 4. CD of 32R (2 mM) in 50 mM Tris pH 7.4, 100 mM KCl in the presence of increasing amounts of Up1 (r = 0, 1, 4, 6) (r = [protein]/[DNA]) (a); hnRNPA1 r = 0, 1, 4, 6, (b); trypsinogen inhibitor (TI) (r = 0, 1, 4, 6) (c). The CD of Up1 at three concentrations is reported (2, 4 and 8 mM) (d). Spectra have been recorded at room temperature with a path length cuvette of 0.5 cm. Ordinate reports ellipticity values in mdeg.

by a $\Delta P = 0.75 - 0.52 = 0.23$. We then asked if quadruplex F-32R-T is unfolded by A1/Up1. To choose at which ionic strength the FRET experiments in the presence of A1/Up1 should be performed, we measured the T_m of quadruplex F-32R-T in KCl and NaCl solutions (in 50, 100 and 140 mM KCl, T_m is 48, 70 and 75°C, respectively; in 100 mM NaCl, the T_m is 32°C). Figure 5a shows the effect on quadruplex F-32R-T in 50 mM KCl (T_m = 48°C,

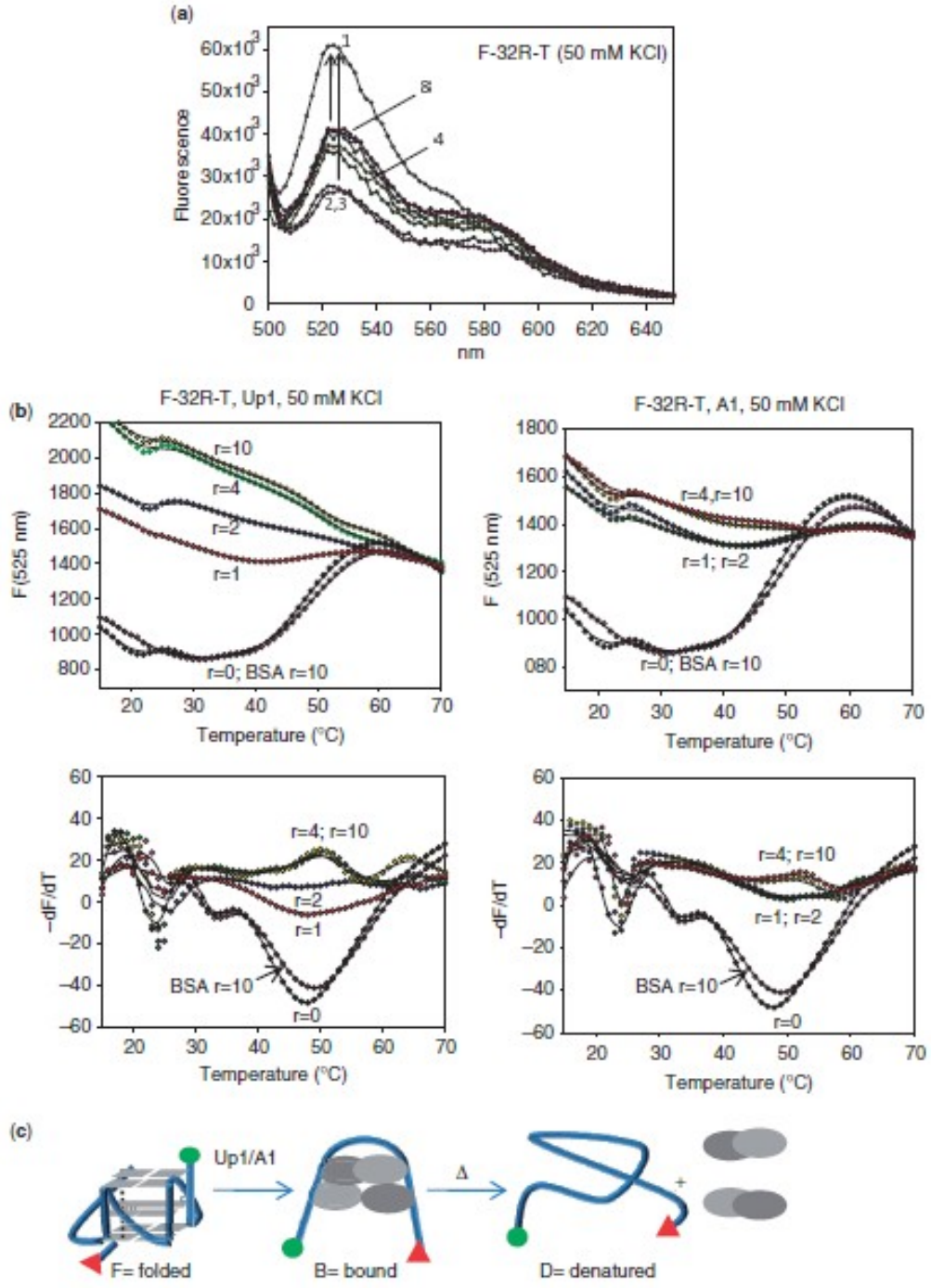


Figure 5. (a) Fluorescence spectra of 200 nM F-32R-T in water (spectrum 1) or 50 mM Tris–HCl, pH 7.4, 50 mM KCl in the absence (spectrum 2) or presence of BSA (r = 10, spectrum 3) or Up1 (r = 0.5, 1, 3, 6, 10, spectra 4–8); (b) row FRET-melting curves (F_{525} versus T) obtained with the iQ5 real-time PCR machine of quadruplex F-32R-T treated with A1/Up1 at various [protein]/[DNA] ratios, in 50 mM Tris pH 7.4, 50 mM KCl. As reference a melting curve of F-32R-T in the presence of BSA (r = 10) is reported. Bottom panels show the corresponding first derivative curves, $-dF_{525}/dT$ versus T. The G-quadruplex was incubated with the protein for 30 min prior to melting; (c) schematic representation of the U-shape structure of the DNA–protein complex between F-32R-T and Up1.

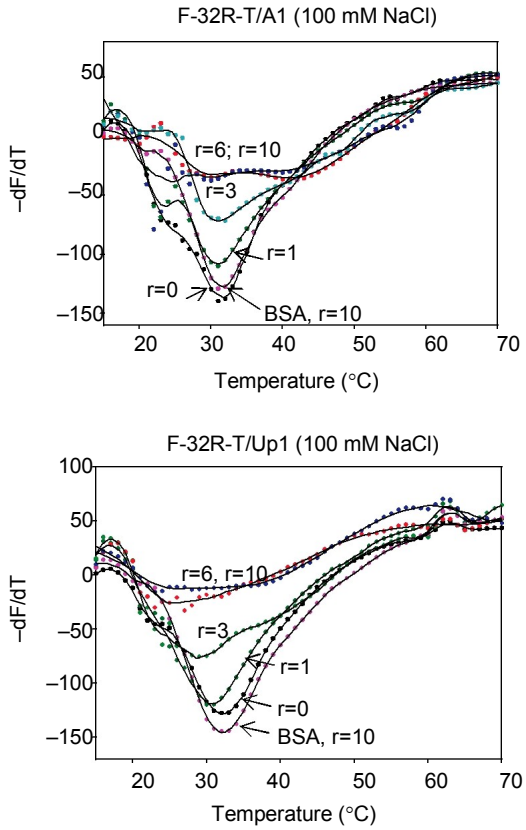


Figure 6. $-dF_{525}/dT$ versus T curves in 50 mM Tris pH 7.4, 100 mM NaCl of F-32R-T in the absence or presence of A1 (a), Up1(b), BSA at various [protein]/[DNA] ratios.

P = 0.62) of Up1 and A1 at r = 1, 2, 3, 6, 10. It can be seen that, compared to complementary 32Y, Up1 does not promote a significant increase of the donor emission, a behaviour that might suggest that Up1 has little eect on the quadruplex conformation [for instance the P-value is 0.62 at r = 0 (spectrum 2), 0.67 at r = 10 (spec-trum 8), P = 0.05]. If we assume that P = 0.23 reflects total opening of the G-quadruplex, P = 0.05 suggests that F-32R-T bound to Up1 is partially opened (20%). Alternatively, it is possible that F-32R-T in the DNA– protein complex is completely opened but with the 5^0 and 3^0 ends brought close to one another so that FRET takes place. To gain insight into this possibility we per-formed melting experiments. We reasoned that in case the quadruplex is partially unfolded, its T_m would be lowered, whereas in case it is completely opened by A1/Up1, the quadruplex-to-ssDNA transition should be abrogated. Figure 5b shows typical melting curves for quadruplex F-32R-T in 50 mM KCl, obtained with a real-time PCR machine, after the DNA was incubated for 30 min with A1/Up1 (r = 1, 2, 4, 10) or BSA (r = 10) just before melt-ing. It can be seen that an excess of BSA does not change the T_m of the G-quadruplex, as one expects with an unspe-cific protein which does not interact with DNA. In con-trast, when quadruplex F-32R-T is incubated with A1/ Up1, a strong change of the melting curves is observed. The cooperative transition relative to the denaturation of the G-quadruplex (T_m of 488C) is completely abrogated and replaced with a broad and non-cooperative curve, which reflects the disruption of the DNA–protein com-plex. The abrogation of the quadruplex-to-ssDNA transi-tion is clearly observed with both F versus T and –dF/dT versus T curves. A similar behaviour has been reported for the UV-melting of the virus type 1 nucleocapsid protein bound to the quadruplex formed by d(GGGTTGGTGTG GTTGG) (35). In 100 mM NaCl, where quadruplex F-32R-T shows a cooperative transition with a T_m of 328C, we also observed the abrogation of the cooperative transition by A1/Up1 (Figure 6). These data suggest that when F-32R-T is bound to A1/Up1, its secondary struc-ture is completely disrupted and F-32R-T in the DNA– protein complex is in the single-stranded form. The fact that the opening of the quadruplex by Up1 is accompa-nied by a P which is 20% of that observed with 32Y (0.05 against 0.23) can be rationalized on the basis of the crystal structure between Up1 and the telomeric repeat $(T\ TAGGG)_2$ (18). In the crystal, the two RRM elements within a Up1 molecule bind to two separate 12mer oligo-nucleotides, which are antiparallel and separated by an interstrand distance of 25–50 A. Thus, we expect that F-32R-T bound to A1/Up1 adopts a U-shape with the two fluorophores close enough to promote energy transfer (18) (Figures 5c).

It is well known that the cationic porphyrin TMPyP4 stabilizes quadruplex DNA by stacking externally to the G-tetrads and interacting with the loop nucleotides (36). We therefore tested whether TMPyP4 reduces the quadruplex destabilizing action of A1/Up1. Quadruplex F-32R-T (200 nM) was incubated for 12 h in 50 mM KCl, in the presence of 200 and 600 nM TMPyP4. Figure 7a shows that TMPyP4 enhances the T_m of quad-ruplex F-32R-T from 488C (curve 1) to 68 (curve 4) and 768C (curve 5). The mixtures were treated for 30 min with 1 mM A1 (r = 5) and then melted. While A1 at r = 5 is able to completely disrupt the KRAS quadruplex in 50 mM KCl (see Figure 5b), in the presence of the porphyrin it promotes only a partial destabilization of the G-quadru-plex: the T_m is reduced from 688C to 588C (in the presence of 200 nM TMPyP4, curve 2, Figure 7a) or from 768C to 638C (600 nM TMPyP4, curve 3, Figure 7a). So, the sta-bilizing eect of the porphyrin partially inhibits the capac-ity of the protein to unfold the G-quadruplex. To exclude the possibility that TMPyP4 directly interacts with and inhibits A1, we performed a control experiment with TMPyP2, the positional isomer of TMPyP4 showing lit-tle anity for quadruplex DNA (Figure 7b). As expected, TMPyP2 neither stabilizes appreciably quadruplex 32R, nor impairs the unfolding of the quadruplex structure by A1. These experiments provide a possible molecular mechanism that explains how TMPyP4 is found to repress the activity of the KRAS promoter (12,13).

Krainer and co-workers (23) showed that A1/Up1 binds to the single-stranded and structured human telo-meric repeat $(TTAGGG)_{n\ =\ 2,\ 4}$. They suggest that A1 is likely to function as an auxiliary factor of the telo-merase holoenzyme and propose that the protein stimulates telomerase elongation through unwinding of the G-quadruplex structures formed during the

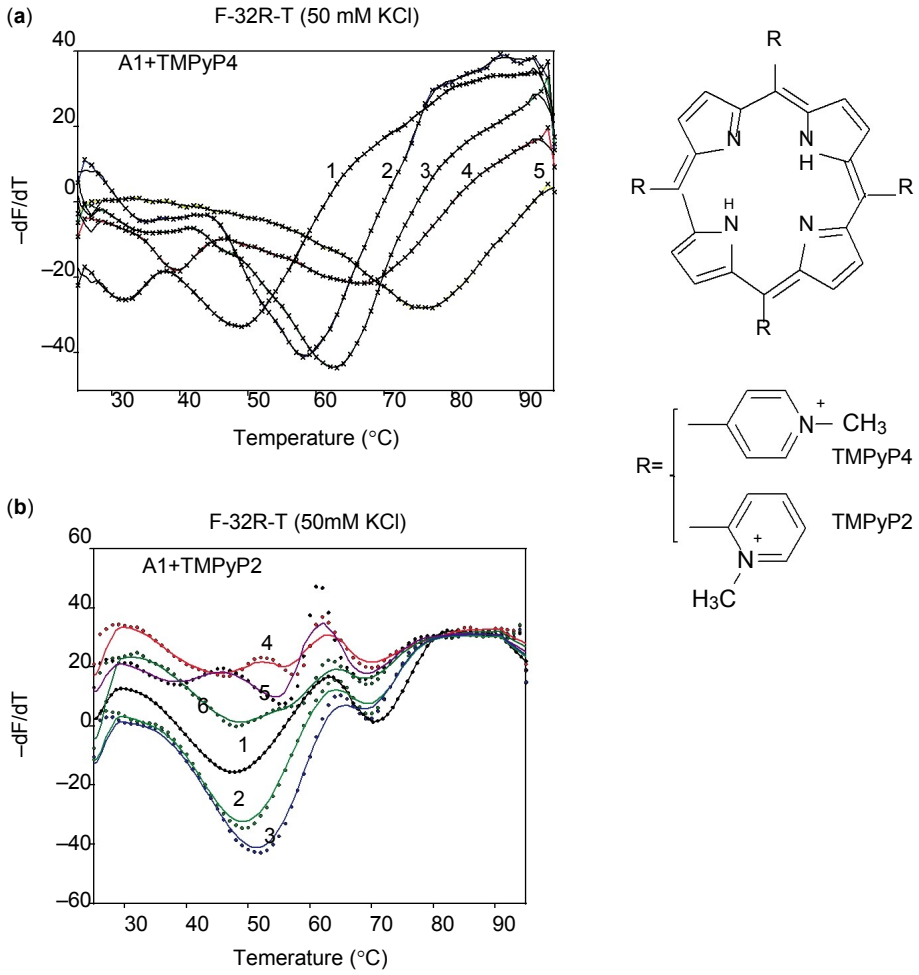


Figure 7. (a) $-dF_{525}/dT$ versus T melting curves of 200 nM F-32R-T in 50 mM Tris pH 7.4, 50 mM KCl (curve 1), in the presence of 200 nM (curve 4) or 600 nM (curve 5) porphyrin TMPyP4. Curves 2 and 3 show the melting curves obtained by F-32R-T treated with 200 nM TMPyP4+A1 (r = 5) or 600 nM TMPyP4+A1 (r = 5), respectively; (b) $-dF_{525}/dT$ versus T melting curves of 200 nM F-32R-T in 50 mM Tris pH 7.4, 50 mM KCl (curve 1), in the presence of 200 nM (curve 2) or 600 nM (curve 3) porphyrin TMPyP2. Curves 4, 5 and 6 show the melting curves obtained by F-32R-T treated with A1 (r = 5); 200 nM TMPyP2+A1 (r = 5); 600 nM TMPyP2+A1 (r = 5), respectively. Exc 475 nm, Em 525 nm; (c) Structures of TMPyP2 and TMPyP4.

translocation steps. Our study suggests that protein A1, being a component of a multiprotein complex formed within NHE (13), may have a similar function for the KRAS promoter: i.e. to resolve the folded quadruplex conformations. The destabilizing activity of A1 should facilitate a quadruplex-to-duplex transformation, that seems to be necessary to activate transcription (12,13). To test this hypothesis, we investigated whether the kinetic of hybridization between quadruplex F-32R-T and the complementary 32Y strand becomes faster in the presence of Up1. When quadruplex F-32R-T in 100 mM KCl (T_m = 708C) is mixed at 258C with the 32Y strand the quadruplex sequence is transformed into the more stable duplex (T_m = 788C) and the fluorescence of the donor increase as in the duplex it is separated from the acceptor (Figure 8a, from A to C). This assembly process can be monitored by measuring the increase of donor (FAM) fluorescence, F, as a function of time (F = F _ F_0, where F_0 is the FAM fluorescence at 525 nm at t = 0 and F the fluorescence at time t). The F versus t curve shows an exponential shape that was best-fitted to a double-exponential equation (37). For the slow phase a constant k_{slow} of 1.56 _ 10^{-3} _ 6 _ 10^{-5} s^{-1} was obtained (Figure 8b). The hybridization performed in the presence of Up1 occurs with a faster kinetic which was nicely best-fitted to a single-exponential equation: k = 5.2 _ 10^{-3} s^{-1}. In this case the assembly occurs between F-32R-T bound to Up1 and 32Y, the fluorescence increases from B to C (Figure 8b). The half-life $t_{1/2}$ for the hybridization of 32R to 32Y in the presence of Up1 is 133 s, while in the absence of Up1 is 444 s, i.e. more than 3 times higher. This demon-strates that Up1 is indeed a G4-DNA destabilizing protein that facilitates the quadruplex-to-duplex transformation within NHE.

Finally, by a primer extension assay using a template containing the KRAS G-rich element we tested whether A1/Up1 is able to remove the block to Taq polymerase caused by quadruplex formation (12,13,24). Figure 9 shows that when the template is incubated in 100 mM KCl prior to primer extension, Taq polymerase is arrested

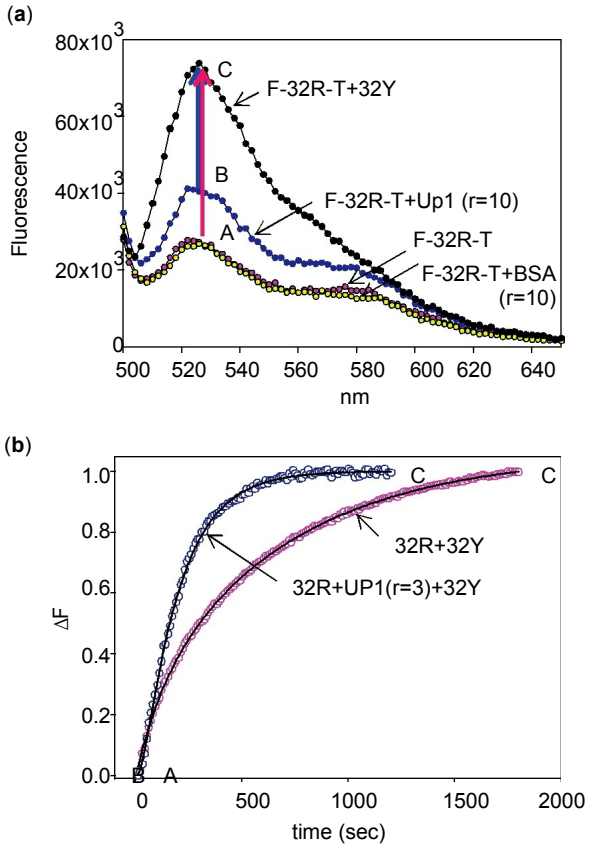


Figure 8. (a) FRET spectra of 200 nM F-32R-T in 50 mM KCl, in the presence of BSA (r = 10), Up1 (r = 10) and 6-fold complementary 32Y strand. (b) Increase of fluorescence F as a function of time following the addition to 200 nM F-32R-T of 6-fold complementary 32Y in 50 mM Tris pH 7.4, 100 mM KCl. Exc 475 nm; Em 525 nm. The experiment has been conducted in the absence and presence of Up1 (r = 3). The solid lines are the best-fits of the experimental points with an exponential equation (SigmaPlot 11, Systat Software Inc).

at the 3^0 end of the G-rich tract, as this element folds into a G-quadruplex structure. Contrarily to what we expected, the addition of increasing amounts of protein A1/Up1 strengthened the pause of Taq polymerase. The precise points at which Taq polymerase was arrested were determined by Sanger sequencing reactions and are indicated with arrows in the template sequence. This suggests that A1/Up1 forms with the DNA template a complex which is suciently stable to arrest the processivity of Taq polymerase. That's why A1/Up1 enhances the block of Taq polymerase at the G-rich element. However, to corroborate this hypothesis DNA footprinting experiments should be done to demonstrate direct binding of A1/ Up1 to the site of arrest. The complex between A1/Up1 and 32R is destabilized when the G-rich strand hybridizes to its complementary sequence to aord a B-DNA duplex for which A1/Up1 has no anity (see EMSA). Finally, in keeping with the results in Figure 3, the primer-extension assay shows that A1/Up1 binds to the G-rich tract of NHE with a high selectivity, since significant arrests of polymerase at other points of the template are not observed.

DISCUSSION

This work describes the ability of A1, and its derivative Up1, to destabilize the quadruplexes of the KRAS promoter and to facilitate their hybridization to the complementary polypyrimidine strand. In accord with pull-down experiments (13), EMSA confirmed that recombinant Up1 and A1 bind to the KRAS quadruplex with a high anity and sequence-specificity, as the binding to other G-quadruplex structures such as HRAS1, HRAS2, CMYC, VEGF appeared either weak or inconsistent. Only the quadruplex from the CKIT sequence (Table 1) is recognised by Up1. The association of A1 to the KRAS promoter is restricted to the polypurine strand, as EMSA shows that A1 does not bind to the complementary poly-pyrimidine strand, nor to NHE in duplex conformation. Considering that the minimum length for strong binding to Up1 is a stretch of 12 nucleotides (18), 32R, being com-posed by 32 nucleotides, has potentially two binding sites. In fact, EMSA shows that 32R forms two DNA–protein complexes that are expected to have a stoichiometry of 1:1 and 1:2 (DNA:Up1). This is in accord with the results of Zhang et al. (23) showing that Up1 forms with the telomeric repeats $(TTAGGG)_4$ two DNA–protein complexes.

In accord with previous observations (24,25), A1 and Up1 promote a significant reduction of the 260 nm ellipticity, typical of G4-DNA in the parallel conformation. This demonstrates that both proteins are able to unfold the quadruplex structures of the KRAS promoter. This conclusion is further supported by FRET-melting experiments showing that the quadruplex formed by F-32R-T is completely disrupted by A1 or Up1. When the KRAS G-quadruplex is incubated for 30 min with A1/Up1 before melting, the cooperative transition of the G-quad-ruplex is replaced by a non-cooperative transition. This suggests that when the KRAS sequence is bound to A1/Up1, it is open and in a single-stranded conformation, as shown by the crystal of Up1 with the telomeric repeat (18). In contrast, when a 10-fold excess BSA is added to the G-quadruplex, no change in the quadruplex transition is observed. We interestingly found that protein A1/Up1 facilitates the assembly into a duplex of the two complementary NHE strands. In fact, the half-life of renaturation is reduced from 444 to 133 s in the presence of Up1, 100 mM KCl. This is in accord with earlier studies report-ing that A1 promotes a rapid renaturation of nucleic-acid strands, probably by melting the secondary structures that are formed transiently during the annealing process (38). The finding that A1 resolves the KRAS quadruplexes has an important biological significance because previous studies supported the notion that the KRAS G4-DNA might behave as a transcription suppressor (12,13,28)

The role of A1 in vivo has been investigated in the context of the telomere biogenesis (21–23). One possible func-tion of the protein would be to disrupt the G4-DNA structures of the telomere G-repeats, allowing proper elongation by the telomerase (23). The data of our study suggest that A1 could have a similar function in the transcription of KRAS. This is in keeping with the fact that A1: (i) binds to the folded G4-DNA conformations of

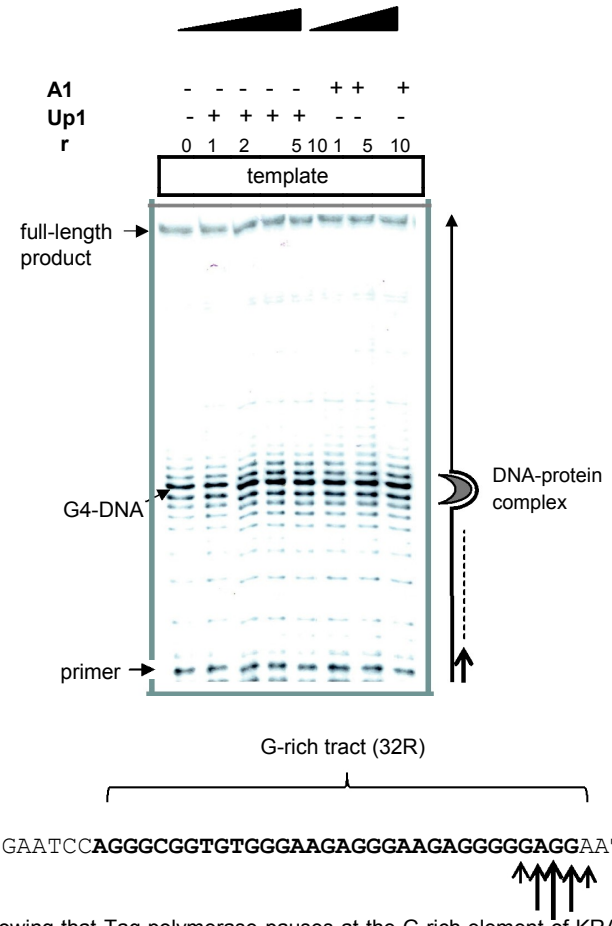


Figure 9. Primer extension assay showing that Taq polymerase pauses at the G-rich element of KRAS where the template forms G-quadruplex structures in the presence of KCl. The 87-mer DNA template (100 nM) was mixed with ^{33}P-labelled primer (50 nM) (Table 1) and incubated for 24 h in 140 mM KCl to allow quadruplex formation by the G-rich element. The mixtures were added with increasing amounts of Up1 (lanes 2–5) or A1 (lanes 6–8), r = [protein]/[DNA] as specified, and incubated for 30 min prior to primer extension. Taq polymerase is arrested at the G-rich element due to quadruplex formation. In the presence of Up1 or A1 the polymerase arrest is stronger. The points in which Taq polymerase is arrested, have been identified by standard Sanger sequencing reactions. Primer extension reaction performed at 378C for 1 h. Reaction products separated in a 12% Urea-TBE denaturing gel.

NHE but not to the complementary pyrimidinic strand or duplex NHE; (ii) disrupts G4-DNA and (iii) facilitates the assembly of the NHE strands into a duplex. A possible model for transcription regulation of KRAS is the following. NHE should exist in equilibrium between a folded (quadruplex) and a double-stranded conformation. In the folded form the promoter is locked into a form that might inhibit transcription (12,13). To activate transcrip-tion, the folded form of NHE should hybridize to the complementary strand in order to restore the duplex. As the quadruplex-to-duplex transformation is likely to be kinetically slow, the functions of A1 would be of desta-bilizing the quadruplex and allow the G-rich strand to hybridize to its complementary within a time compatible with a response of the cell to molecular stimuli. There are a number of genes with C+G-rich elements in the region surrounding the transcription start site that seems to be characterized by a transcription regulation mechanism involving G-quadruplex structures (12,13,39–45).

Several proteins from dierent organisms that interact with quadruplex DNA have been reported (46). They can be classified by function into five major groups: (i) proteins that increase the stability of DNA quadruplexes; (ii) proteins that destabilize quadruplex DNA in a non catalytic way; (iii) proteins that unwind catalytically quadruplex DNA in an ATP-dependent fashion; (iv) proteins that promote the formation of quadruplex DNA; (v) Nucleases that specifically cleave DNA at or adjacently to a quadruplex domain. Like other members of the hnRNP family such as hnRNP A2 (20) and CBF-A (20,47) that destabilise the G-quadruplex formed by the d(CGG)n fragile X expanded sequence, protein A1 acts on DNA in a non-catalytic way, i.e. remaining bound to the DNA substrate. Another protein with a similar property is POT-1 which binds to the telomere G-rich DNA over-hangs and disrupts G4-DNA structures (48,49). However, contrarily to A1/Up1, POT-1 causes a significant increase of the P-value of the quadruplex from the human telo-meric repeat, because in the DNA-protein complex the telomeric repeat assumes an extended conformation in which the donor–acceptor are separated by a distance that is too long for FRET (26). Similarly, A1 disrupts the G4-DNA structures assumed by NHE and its remain-ing bound to the G-rich sequence prevents the DNA from assuming again the folded conformation. We were indeed surprised to observe by primer extension experiments that

at 378C, A1/Up1 did not remove the block to Taq polymerase and the protein even enhanced the polymerase arrest. This clearly indicates that after interaction the pro-tein remains bound to the template, and the resulting DNA–protein complex is suciently strong to arrest the processivity of the polymerase.

Finally, the proposed transcription regulation model suggests two strategies to downregulate the KRAS onco-gene and sensitize pancreatic cancer cells, which are refractory to conventional treatment, to chemotherapy. First, use of G4-ligands that lock the promoter in the non-transcriptable form by stabilizing the G-quadru-plexes; second, use of decoy molecules specific for the proteins that recognize the G4-DNA structure of NHE (28). Work is in progress in our laboratory along this direction.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

The Italian Association for Cancer Research (AIRC-2008, Associazione Italiana per la Ricerca Contro il Cancro), FVG–Region (Grant-2007); Italian Ministry of University and Research (Prin 2008). Funding for open access charge: AIRC 2008.

Conflict of interest statement. None declared.

REFERENCES

1. Malumbres,M. and Barbacid,M. (2003) RAS oncogenes: the first 30 years. Nat. Rev. Cancer, 3, 459–465.
2. Lowry,D.R. and Willumsen,B.M. (1993) Functional and regulation of ras. Ann. Rev. Biochem., 62, 851–891.
3. Reuther,G.W. and Der,C.J. (2000) The ras branch of small GTPases: Ras family members don't fall far from the tree. Curr. Opin. Cell Biol., 12, 157–165.
4. Bos,J.L. (1989) Ras oncogenes in human cancer: a review. Cancer Res., 49, 4682–4689.
5. Barbacid,M. (1990) Ras oncogenes: their role in neoplasia. Eur. J. Clin. Invest., 20, 225–235.
6. Burmer,G.C. and Loeb,L.A. (1989) Mutations in the KRAS2 oncogene during progressive stages of human colon carcinoma. Proc. Natl Acad. Sci. USA, 86, 2403–2407.
7. Almoguerra,C., Shibata,D., Forrester,K., Martin,J., Arnheim,N. and Perucho,M. (1988) Most human carcinomas of the exocrine pancreas contain mutant c-K-ras genes. Cell, 53, 549–554.
8. Shirasawa,S., Furuse,M., Yokoyama,N. and Sasazuki,T. (1993) Altered growth of human colon cancer cell lines disrupted at acti-vated Ki-ras. Science, 260, 85–88.
9. Schubbert,S., Shannon,K. and Bollag,G. (2007) Hyperactive Ras in developmental disorders and cancer. Nat. Rev., 7, 295–308.
10. Bardeesy,N. and DePinho,R. (2002) Pancreatic cancer biology and genetics. Nat. Rev., 2, 897–909.
11. Yamamoto,F. and Perucho,M. (1988) Characterization of the human c-K-ras gene promoter. Oncogene Res., 3, 125–138.
12. Cogoi,S. and Xodo,L. (2006) G-quadruplex formation within the promoter of the KRAS proto-oncogene and its eect on transcrip-tion. Nucleic Acids Res., 34, 2536–2549.
13. Cogoi,S., Paramasivam,M., Spolaore,B. and Xodo,L.E. (2008) Structural polymorphism within a regulatory element of the human KRAS promoter: formation of G4-DNA recognized by nuclear proteins. Nucleic Acids Res., 36, 3765–3780.
14. Dreyfuss,G., Matunis,S., Pinol-Roma,S. and Burd,C. (1993) hnRNP proteins and the biogenesis of mRNA. Annu. Rev. Biochim., 62, 289–321.
15. McAfee,J., Huang,M., Soltaninassad,S., Rech,J., Iyengar,S. and Lestougeon,W. (1997) The packaging of pre-mRNA. In Krainer,A.R. (ed.), Eukaryotic mRNA Processing. Vol. 17, IRL Press at Oxford University Press, New York, N.Y., pp. 68–102.
16. Cobianchi,F., SenGupta,D., Zmudzka,B. and Wilson,S. (1986) Structure of rodent helix-destabilizing protein revealed by cDNA cloning. J. Biol. Chem., 261, 3536–3543.
17. Shamoo,Y., Abdul-Manan,N., Patten,A., Crawford,J., Pellegrini,M. and Williams,K.R. (1994) Both RNA-binding domains in hetero-genous nuclear ribonucleoprotein A1 contribute toward single-stranded-RNA binding. Biochemistry, 33, 8272–8281.
18. Ding,J., Hayashi,M., Zhang,Y., Manche,L., Krainer,A. and Xu,R.-M. (1999) Crystal structure of the two-RRM domain of vhnRHP A1 (Up1) complexed with single-stranded telomeric DNA. Genes Dev., 13, 1102–1115.
19. Fiset,S. and Chabot,B. (2001) hnRNP A1 may interact simulta-neously with telomeric DNA and the human telomerase RNA in vitro. Nucleic Acids Res., 29, 2268–2275.
20. Khateb,S., Weisman-Shomer,P., Hershco,I., Loeb,L.A. and Fry,M. (2004) Destabilization of tetraplex structures of the fragile X repeat sequence (CGG)n is mediated by homolog-conserved domains in three members of the hnRNP family. Nucleic Acids Res., 32, 4145–4154.
21. LaBrance,H., Dupuis,S., Ben-David,Y., Bani,M.-R., Wellinger,R. and Chabot,B. (1998) Telomere elongation by hnRNP A1 and a derivative that interacts with telomeric repeats and telomerase. Nat. Genet., 19, 199–202.
22. Riva,S., Morandi,C., Tsoulfas,P., Pandolfo,M., Biamonti,G., Merill,B., Williams,K., Multhaup,G., Beyreuther,K., Werr,H. et al. (1986) Mammalian single-stranded DNA binding protein UP I is derived from the hnRNP core protein A1. EMBO J., 5, 2267–2273.
23. Zhang,Q., Manche,L., Xu,R.-M. and Krainer,A. (2006) hnRNP A1 associates with telomere ends and stimulates telomerase activity. RNA, 12, 1116–1128.
24. Fukuda,H., Katahira,M., Tsuchiya,N., Enokizono,Y., Sugimura,M., Nagao,M. and Nakagama,H. (2002) Unfolding of quadruplex structure in the G-rich strand of the minisatellite repeat by the binding protein UP1. Proc. Natl Acad. Sci. USA, 99, 12685–12690.
25. Fukuda,H., Katahira,M., Tanaka,E., Enokizono,Y., Tsuchiya,N., Higuchi,K., Nagao,M. and Nakagama,H. (2005) Unfolding of higher DNA structures formed by the d(CGG) triplet repeat by UP1 protein. Genes Cells, 10, 953–962.
26. Salas,T.R., Petruseva,I., Lavrik,O., Bourdoncle,A., Mergny,J.L., Favre,A. and Saintome´,C. (2006) Human replication protein A unfolds telomeric G-quadruplexes. Nucleic Acids Res., 34, 4857–4865.
27. Nagatoishi,S., Nojima,T., Galezowska,E., Juskowiak,B. and Takenaka,S. (2006) G quadruplex-based FRET probes with the thrombin-binding aptamer (TBA) sequence designed for the ecient fluorometric detection of the potassium ion. Chembiochem, 7, 1730–1737.
28. Cogoi,S., Paramasivam,M., Filichev,V., Ge´ci,I., Pedersen,E.B. and Xodo,L.E. (2009) Identification of a new G-quadruplex motif in the KRAS promoter and design of TINA-modified G4-decoys with antiproliferative activity in pancreatic cancer cells. J. Med. Chem., 52, 564–568.

29. Seenisamy,J., Rezler,E.M., Powell,T.J., Tye,D., Gokhale,V., Joshi,C.S., Siddiqui-Jain,A. and Hurley,L.H. (2004) The dynamic character of the G-quadruplex element in the c-MYC promoter and modification by TMPyP4. J. Am. Chem. Soc., 126, 8702–8709.

30. Phan,A.T., Modi,Y.S. and Patel,D.J. (2004) Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter. J. Am. Chem. Soc., 126, 8710–8716.

33. Rujan,I.N., Meleney,J.C. and Bolton,P.H. (2005) Vertebrate telomere repeat DNAs favor external loop propeller quadruplex structures in the presence of high concentrations of potassium. Nucleic Acids Res., 33, 2022–2031.

34. Clegg,R.M. (1992) Fluorescence resonance energy transfer and nucleic acids. Methods Enzymol., 211, 353–388.

35. Kankia,B.I., Barany,G. and Musier-Forsyth,K. (2005) Unfolding of DNA quadruplexes induced by HIV-1 nucleocapsid protein. Nucleic Acids Res., 33, 4395–4403.

36. Han,H., Langley,D.R., Rangan,A. and Hurley,L.H. (2001) Selective interactions of cationic porphyrins with G-quadruplex structures. J. Am. Chem. Soc., 123, 8902–8913.

37. Green,J.J., Ying,L., Klenerman,D. and Balasubramanian,S. (2003) Kinetics of unfolding the humna telomeric DNA G-quartet struc-ture using a PNA trap. J. Am. Chem. Soc., 125, 3763–3767.

38. Pontius,B.W. and Berg,P. (1990) Renaturation of complementray DNA strands mediated by puyrified mammalian heterogeneous nuclear ribonucleoprotein A1 protein: implications for a mechanism for rapid molecular assembly. Proc. Natl. Acad. Sci USA, 87, 8403–8407.

39. Siddiqui-Jain,A., Grand,C.L., Bearss,D.J. and Hurley,L.H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-CMYC transcription.
Proc. Natl Acad. Sci. USA, 99, 11593–11598.

40. Eddy,J. and Maizels,N. (2006) Gene function correlates with potential for G4 DNA formation in the human genome. Nucleic Acids Res., 34, 3887–3896.

41. Huppert,J.L. and Balasubramanian,S. (2007) G-quadruplexes in promoters throughout the human genome. Nucleic Acids Res., 35, 406–413.

31. Phan,A.T., Kuryavyi,V., Burge,S., Neidle,S. and Patel,D.J. (2007) Structure of an unprecedented G-quadruplex scaold in the human c-kit promoter. J. Am. Chem. Soc., 129, 4386–4392.

32. Guo,K., Gokhale,V., Hurley,L.H. and Sun,D. (2008) Intramolecularly folded G-quadruplex and i-motif structures in the proximal promoter of the vascular endothelial growth factor gene. Nucleic Acids Res., 36, 4598–4608.

42. Palumbo,S.L., Memmott,R.M., Uribe,D.J., Krotova-Khan,Y., Hurley,L.H. and Ebbinghaus,S.W. (2008) A novel G-quadruplex-forming GGA repeat region in the c-myb promoter is a
critical regulator of promoter activity. Nucleic Acids Res., 36, 1755–1769.

43. Shklover,J., Etzioni,S., Weisman-Shomer,P., Yafe,A., Bengal,E. and Fry,M. (2008) MyoD uses overlapping but distinct elements to bind E-box and tetraplex structures of regulatory sequences of muscle-specific genes. Nucleic Acids Res., 35, 7087–7095.

44. Todd,A.K. and Neidle,S. (2008) The relationship of potential G-quadruplex sequences in cis-upstream regions of the
human genome to SP1-binding elements. Nucleic Acids Res., 36, 2700–2704.

45. Sun,D., Liu,W.J., Guo,K., Rusche,J.J., Ebbinghaus,S., Gokhale,V. and Hurley,L.H. (2008) The proximal promoter region of the human vascular endothelial growth factor gene has a G-quadruplex structure that can be targeted by G-quadruplex-interactive agents.
Mol. Cancer Ther., 7, 880–889.

46. Fry,M. (2007) Tetraplex DNA and its interacting proteins. Front. Biosci., 12, 4336–4351.

47. Weisman-Shomer,P., Cohen,E. and Fry,M. (2002) Distinct domains in the CarG-box binding factor-A destabilize tetraplex forms of the fragile X expanded sequence d(CGG)n. Nucleic Acids Res., 30, 3672–3681.

48. Zaug,A.J., Podell,E.R. and Cech,T.R. (2005) Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro.
Proc. Natl. Acad. Sci. USA, 102, 10864–10869.

49. Wang,F., Podell,E.R., Zaug,A.J., Yang,Y., Baciu,P., Cech,T.R. and Lei,M. (2007) The POT1-TPP1 telomere complex is a telomerase processivity factor. Nature, 445, 506–510

Supplementary Information

Protein hnRNP A1 and its derivative Up1 unfold quadruplex DNA in the human *KRAS* promoter: implications for transcription

Manikandan Paramasivam[1,&], Alexandro Membrino[1,&], Susanna Cogoi[1], H Fukuda[2], H. Nakagama[2], Luigi E.Xodo[1]*

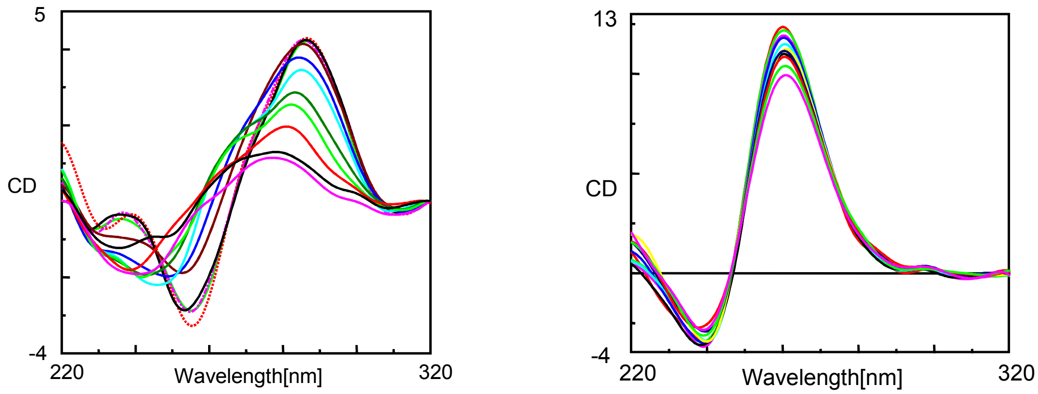
[1]Department of Biomedical Science and Technology, School of Medicine, P.le Kolbe 4, 33100 Udine, Italy; [2] Biochemistry Division, National Cancer Center Research Institute, 1-1, Tsukiji 5, Chuo-ku, Tokyo 104-0045, Japan.

Corresponding author: Luigi E. Xodo, Department of Biomedical Science and Technology, School of Medicine, P.le Kolbe 4, 33100 Udine, Italy; e-mail: luigi.xodo@uniud.it; Tel. +39-0432-494395; Fax. +39-0432-494301.

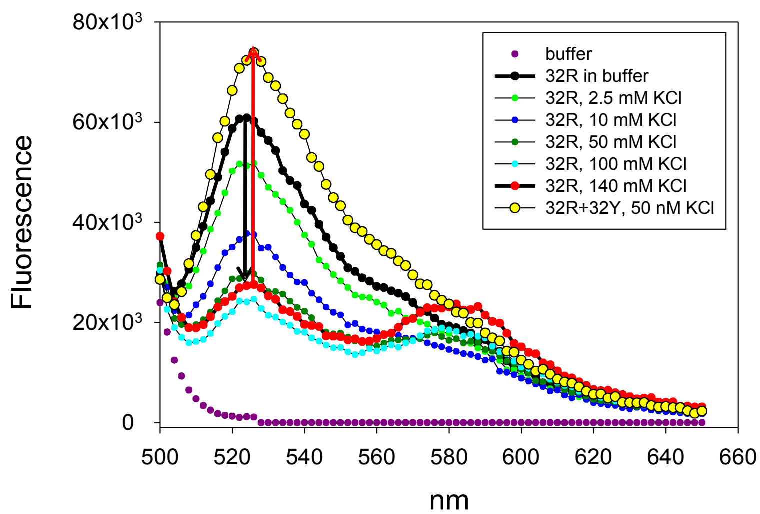
[&]MP and AM have equally contributed to this work;

HRAS promoter sequences forming G4-DNA structures

5’-TC**GGG**TTGC**GGG**CGCA**GGG**CAC**GGG**CG   5’-C**GGGG**C**GGGG**C**GGGGG**C**GGGGG**CG



Supplementary S1: G-rich sequences located in the *HRAS* promoter upstream the transcription start site. Both sequences are recognized by proteins and assume a G-quadruplex conformation as indicated by CD spectra obtained in 50 mM Tris-HCl, pH 7.4, 100 mM KCl. Sequence Hras-1 (left) forms an antiparallel quadruplex, while sequence Hras-2 (right) forms a parallel quadruplex. The CD spectra reported have been obtained as function of temperature in the 20-90°C temperature range.



Supplementary S2: Fluorescence spectra of 200 nM F-32R-T in 50 mM Tris-HCl, pH 7.4, at increasing concentration of KCl (from 2.5 to 140 mM). Exc= 475 nm, Em=550-660 nm.

4.2. Cellular uptake and binding of guanidine-modified phthalocyanines to *KRAS/HRAS* G-quadruplexes.

Focusing on *KRAS* and *HRAS* promoter sequences we identified their quadruplex structures and characterized them. Based on our results we recognized the role of these quadruplexes in transcription. This result was supported by the literature in which quadruplexes analyzed throughout the entire genome are found preferentially around the TSS of many genes.

Due to the potential to be cancer-specific targets, there is considerable interest in developing small ligands that stabilize quadruplexes. The most famous molecules used for quadruplex stabilization are porphyrins. Starting from this approach we decided to test some other quadruplex ligands that have shape and charge complementarity with the stacked G-tetrads that constitute the quadruplex DNA.

These new molecules are pyridinium and ammonium-containing porphyrazine derivatives which exhibited improved quadruplex specificity as compared to the widely studied, yet non-selective ligand 5,10,15,20-tetra(N-methyl-4-pyridyl) porphyrine (TMPyP4).

Here, we reported quadruplex binding from which DIGP showed its ability to stabilize both parallel and anti-parallel quadruplexes while its zinc-containing derivative and TMPyP4 exhibited selectivity for parallel quadruplexes such as hras-2 sequence; cellular uptake showed that DIGP phthalocyanine enters the nucleoli while the DIGP-Zn derivative remains mainly in the cytoplasm. By means of luciferase reporter assay we analyzed *KRAS* and *HRAS* promoter modulation driven by this new class of cationic phthalocyanines consistent with quadruplex-mediated promoter inhibition which provided interesting results that stimulated us to further explore the anticancer potential of guanidine-modified phthalocyanines (GPcs).

Cellular uptake and binding of guanidine-modified phthalocyanines to *KRAS*/*HRAS* G-quadruplexes†

Alexandro Membrino,‡[a] Manikandan Paramasivam,‡[a] Susanna Cogoi,[a] Jawad Alzeer,[b] Nathan W. Luedtke*[b] and Luigi E. Xodo*[a]

Received (in Cambridge, UK) 14th September 2009, Accepted 23rd October 2009
First published as an Advance Article on the web 20th November 2009
DOI: 10.1039/b918964e

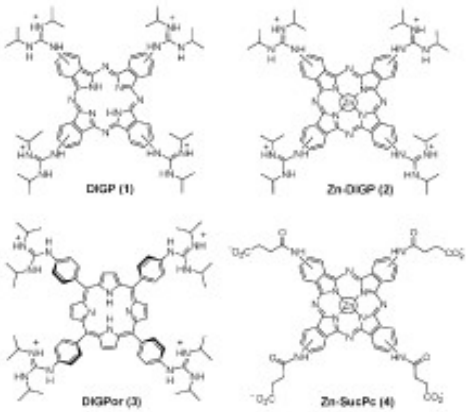
Guanidino-modified phthalocyanines are evaluated *in vitro* (polymerase-stop assays and FRET) and in cultured cells as G4-DNA ligands and modulators of gene transcription.

The hypothesis that G-quadruplex DNA (or G4-DNA) is involved in transcription regulation is gaining support.[1] Recent studies have shown that, in addition to the telomeres, G4-DNA motifs are found with a high frequency in the regions surrounding transcription start sites of many genes.[2]

G-quadruplex structures have been identified in proto-oncogenes as well as in 5′-untranslated regions of mRNA.[3] Several studies have suggested that G4-DNA affects the transcription of several genes including c-*MYC*,[4] c-*kit*,[5] *KRAS*,[6] *VEGF*,[7] c-*myb*[8] and ILPR (insulin gene).[9] Due to its potential as a cancer-specific target, there is considerable interest in developing small ligands that stabilize G4-DNA.[10] Structure-selective G-quadruplex ligands typically have shape and charge complementarity with the stacked G-tetrads that constitute G-quadruplex DNA. For example, pyridinium and ammonium-containing porphyrazine derivatives exhibited improved G-quadruplex specificity as compared to the widely studied, yet non-selective ligand 5,10,15,20-tetra(*N*-methyl-4-pyridyl) porphine (TMPyP4).[11,12] However, no information regarding the cellular uptake or promoter binding of these compounds was reported. Here, we report G4-DNA binding, cellular uptake, and promoter deactivation of a new class of cationic phthalocyanines called guanidino phthalocyanines (GPcs).[13] We have used polymerase-stop assays, CD spectroscopy, and a fluorescence quenching assay to characterize the G-quadruplex affinity and specificity of tetrakis-(diisopropyl-guanidine) phthalocyanine "DIGP" (**1**), and its Zn-containing derivative "Zn-DIGP" (**2**) (Scheme 1). To facilitate a direct comparison of porphyrin *versus* phthalocyanine scaffolds, a porphyrin containing four diisopropyl guanidinium groups at *meso* positions "DIGPor" (**3**) was synthesized and evaluated. To test the importance of charge–charge interactions, a phthalocyanine with four succinate groups "Zn-SucPc" (**4**) was also synthesized and characterized for G-quadruplex stabilization.

Polymerase-stop assays were conducted using an 80-mer DNA template containing the *GA*-element located in the promoter of the murine *KRAS* gene, which folds into a parallel G-quadruplex.[6] This promoter element is critical as its excision results in a total arrest of *KRAS* transcription.[14] To establish suitable experimental conditions for conducting the polymerase stop assays, we used a FRET assay to determine the lowest KCl concentration (25 mM) required by the *GA*-element to assume a stable G-quadruplex (T_M = 65 °C) (S_1†). Fig. 1a shows that in 25 mM KCl, but not in the presence of 25 mM LiCl (S_2†), Taq polymerase is partly arrested at the 3′ end of the *GA*-element, at the adenine adjacent to the first G-run (positions determined by sequencing Sanger reactions). As the murine GA-element comprises six G-runs (1–6) separated by "A" or "AAGGA", at least three topologically distinct G-quadruplexes can therefore be formed: Q_1 by the G-runs 1–2–3–4, Q_2 by the G-runs 2–3–4–5 and Q_3 by the G-runs 3–4–5–6. All three combinations have two single-nucleotide and one five-nucleotide loops (S_3†). In keeping with this expectation, a second polymerase stop is observed at the beginning of the second G-run (Fig. 1a), consistent with the formation of Q2.

When the primer-*KRAS* template was incubated with the GPcs at [ligand]/[DNA] ratios (r) between 10 to 40, compounds **1** and **2** induced a strong arrest of Taq polymerase, with a progressive reduction of the full-length product (*flp*)



Scheme 1 Structures of investigated compounds and common names.

[a] *Department of Biomedical Science and Technology, University of Udine, 33100 Udine, Italy. E-mail: luigi.xodo@uniud.it; Fax: +39 0432 494301; Tel: +39 0432 494395*
[b] *Institute of Organic Chemistry, University of Zürich, Winterthurerstrasse 190, CH-8057, Zürich, Switzerland. E-mail: luedtke@oci.uzh.ch; Fax: +41 44 635 6891; Tel: +41 44 635 4244*

† Electronic supplementary information (ESI) available: Experimental procedures, including synthesis, polymerase stop assays, FRET and CD data are available free of charge *via* the Internet at http://pubs.acs.org. See DOI: 10.1039/b918964e

‡ AM and MP have equally contributed to this work.

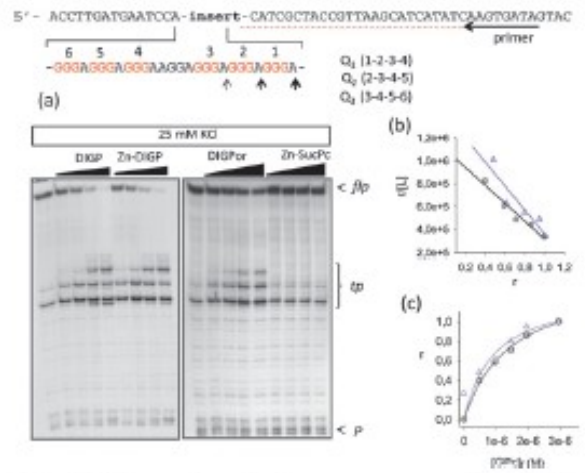


Fig. 1 (a) DNA template and polymerase stop assays where 50 nM of template was incubated with [GPc]/[DNA] ratios of 0, 10, 20, 30, 40 and subjected to primer extension with Taq polymerase; (b) Scatchard plots for compound **1** (diamonds) and **2** (circles) using data obtained from the gels; (c) the same binding data for **1** and **2** fit to $y = B{\cdot}L/(K_D + L)$.

and concomitant formation of truncated products (*tp*) (Fig. 1a). The porphyrin-based derivative DIGPor (**3**) stabilizes the murine G-quadruplexes, but to a much lower extent as compared to its phthalocyanine analogue **1** (*tp* ~ 50%). The anionic succinate phthalocyanine **4** did not show any stabilizing activity (*tp* ~ 0%). This suggests that both the cationic guanidinium groups and phthalocyanine scaffold are important for G-quadruplex binding. By polymerase stop assays we also compared GPcs with TMPyP4 and found that the latter showed for the *KRAS* quadruplex a slightly higher affinity than the former (S_4†). Assuming that the amount of truncated products is proportional to the relative amount of GPc-bound template, we estimated apparent K_D values of approximately 1.0×10^{-6} M for compounds **1** and **2** by both Scatchard and non-linear curve fitting (Fig. 1b and c). As an orthogonal analysis of ligand binding, we estimated K_D values by measuring the fluorescence quenching of a FRET construct Fam-GGGAGGGAGGGAAGGAGGGAGGGAGG-GA-Tamra (Fam = fluorescein, Tamra = tetramethylrhodamine) upon titration of the GPcs. Fig. 2a shows a representative titration in 50 mM Tris-HCl, pH 7.4, 100 mM KCl. The addition of GPcs **1** and **2** resulted in selective quenching of Tamra emission, as GPc absorbs in the region of Tamra emission (S_5†). The best-fit of this binding curve to a standard binding equation gave a K_D of 8.9×10^{-7} M, in keeping with the value obtained by polymerase stop assays (Fig. 2b). Compound **4**, which does not bind to quadruplex DNA according to the polymerase stop assay did not quench this construct (Fig. 2c). We extended this analysis to other G-quadruplexes from the *PTHR1* (S_6) and *HRAS* promoters and found K_Ds in the order of 10^{-6}–10^{-7} M (binding data are summarized in Table S_1, ESI†). These values are similar to those reported for the binding of tetramethylpyridinium-porphyrazines to a *Htelo* quadruplex.[11]

To ascertain whether structural changes occur upon GPc binding, the CD spectrum of the murine *KRAS* G-quadruplex

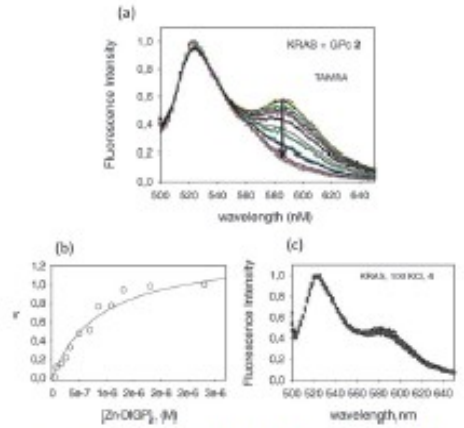


Fig. 2 (a) Titration of 200 nM F-28*R*-T in 50 m mM Tris-HCl, pH 7.4, 100 mM KCl with 50, 100, 200, 300, 400, 600, 800, 1000, 1250, 2000, 3000 nM GPc **2**; (b) Binding curve and best-fit to $y = B{\cdot}L/(K_D + L)$; (c) titration as in (a) but with compound **4**.

was measured as a function of GPc. The addition of increasing amounts of **1** up to 12 μM (r = 1,2,3,4) resulted in no changes of the CD spectrum, suggesting that the G4-DNA structure did not change upon GPc binding. Upon heating, the resulting GPc-quadruplex complex did not completely melt (T_M > 95 °C) (not shown). In contrast, no thermal stabilization of *KRAS* duplex was observed (Fig. 3a). Little if any duplex DNA binding was confirmed by (i) PAGE showing that the protein–DNA complexes C_1 and C_2, obtained by incubating NIH 3T3 nuclear extract with the *KRAS* duplex, were not disrupted upon addition of **1** (Fig. 3b); (ii) the fact that titrating quadruplex F-28*R*-T with **1** in the absence or presence of 5-fold salmon sperm DNA, results in similar fluorescence quenching curves (Fig. 3c, S_7). Together, this data shows that GPcs are structure-specific G4-DNA ligands, showing a good quadruplex/duplex selectivity.

To probe the selectivity of GPcs for parallel *versus* anti-parallel G-quadruplexes, we performed polymerase-stop assays with a DNA template containing a G-rich sequence from the *HRAS* promoter, called *hras*-1, that can form an antiparallel G-quadruplex (Fig. 4). Polymerase stop assays, performed with a template containing *hras*-1, showed that **1**

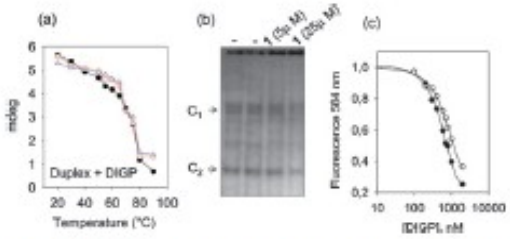


Fig. 3 (a) CD-melting curve (280 nm ellipticity *versus* T) of 3 μM murine *KRAS* duplex in 50 mM Tris-HCl (pH 7.4), 100 mM NaCl, in the absence (●) and presence of DIGP r = 1 (△), 2 (○); (b) Mobility-shift assay of ^{32}P-labelled *KRAS* duplex (20 nM) incubated with 5 μg NIH 3T3 nuclear extract, in the absence or presence of 5 and 25 μM DIGP; (c) Normalized fluorescence (584 nm) of quadruplex F-28*R*-T quenched upon addition of **1** in the absence (○) or presence of 5-fold salmon sperm DNA (●).

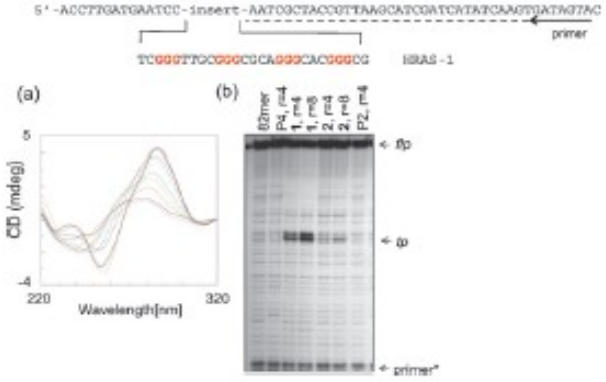


Fig. 4 Sequence of the *hras-1* template and (a) CD spectra of its quadruplex-forming element (3 μM) in 50 mM Tris-HCl, pH 7.4, 100 mM KCl from 20 to 90 °C (T_M = 60 °C). An antiparallel G4-DNA structure is indicated by a strong ellipticity at 288 nm and negative ellipticity at 257 nm; (b) Polymerase stop assays used 50 nM of the DNA template incubated with the primer and GPc at *r* of 0, 4 and 8. Primer extension with Taq polymerase was carried out for 1 h at 37 °C. P2 = TMPyP2, P4 = TMPyP4.

significantly stabilized this antiparallel G-quadruplex. In contrast, the zinc-containing phthalocyanine **2** caused little arrest at $r = 4, 8$, or 10 (Fig. 4 and S_8). This is an example of metal ion-mediated specificity for parallel *versus* antiparallel G-quadruplex DNA.

To determine the cellular uptake and the localization of the designed phthalocyanines, we utilized the fluorescence properties of GPcs to stain living NIH 3T3 and HeLa cells. Both GPcs were detected in the cells already 1 h after treatment (S_9). Typical images obtained with NIH 3T3 cells are shown in Fig. 5. It can be seen that Zn-DIGP (**2**) preferentially localizes in the perinuclear region, while DIGP (**1**) was more localized in the nuclei and nucleoli.

To evaluate whether GPcs are able to modulate gene expression, we employed a dual luciferase assay with a plasmid containing firefly luciferase driven by the *HRAS* promoter (pHRAS-luc). In this assay HeLa cells were treated for 24 h with 5 μM 5,10,15,20-tetra(*N*-methyl-2-pyridyl) porphine (TMPyP2) or DIGP and transfected with a mixture of pHRAS-luc and pRN-luc, a plasmid expressing renilla luciferase. Interestingly, DIGP was non-toxic and induced a strong down-regulation of the promoter activity to 35 ± 10% relative to untreated cells, while little if any change was observed for

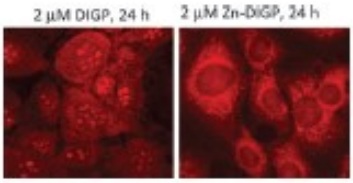


Fig. 5 NIH 3T3 cells treated for 24 h with 2 μM DIGP and Zn-DIGP, fixed on glass slides, and analysed by confocal microscopy. Excitation = 633 nm, emission = 680–800 nm.

TmPyP2 (103 ± 18% of the control). These results indicate that GPcs are potential tools for manipulating gene expression.

In summary, DIGP is a good stabilizing ligand for both parallel and antiparallel G4-DNA, while its zinc-containing derivative and TMPyP4 exhibited selectivity for parallel quadruplexes. GPcs also exhibited good cellular uptake into living cells and suppressed luciferase expression. These results are consistent with G-quadruplex-mediated promoter inhibition,[4–9] and provide motivation to further explore the anticancer potential of GPcs.

We thank the Italian Ministry of Scientific Research (PRIN 2007), AIRC 2008, FVG for financial support. The Swiss National Science Foundation (grant #116868), the Dr Helmut Legerlotz Stiftung, and the University of Zürich are also gratefully acknowledged.

Notes and references

1 (*a*) J. Eddy and N. Maizels, *Nucleic Acids Res.*, 2006, **34**, 3887; (*b*) Y. Qin and L. H. Hurley, *Biochimie*, 2008, **90**, 1149.
2 (*a*) J. L. Huppert and S. Balasubramanian, *Nucleic Acids Res.*, 2007, **35**, 406; (*b*) P. Rawal, V. B. R. Kummarasetti, J. Ravindran, N. Kumar, K. Halder, R. Sharma, M. Mukerji, S. K. Das and S. Chowdhury, *Genome Res.*, 2006, **16**, 644.
3 (*a*) S. Kumari, A. Bugaut, J. L. Huppert and S. Balasubramanian, *Nat. Chem. Biol.*, 2007, **3**, 218; (*b*) M. J. Morris and S. Basu, *Biochemistry*, 2009, **48**, 5313; (*c*) S. Lipps and D. Rhodes, *Trends Cell Biol.*, 2009, **19**, 414.
4 A. Siddiqui-Jain, C. L. Grand, D. J. Bearss and L. H. Hurley, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 11593.
5 (*a*) K. Jantos, R. Rodriguez, S. Ladame, P. S. Shirude and S. Balasubramanian, *J. Am. Chem. Soc.*, 2006, **128**, 13662; (*b*) Z. A. E. Waller, S. A. Sewitz, S. D. Hsu and S. Balasubramanian, *J. Am. Chem. Soc.*, 2009, **131**, 12628–33.
6 (*a*) S. Cogoi and L. E. Xodo, *Nucleic Acids Res.*, 2006, **34**, 2536; (*b*) S. Cogoi, M. Paramasivam, B. Spolaore and L. E. Xodo, *Nucleic Acids Res.*, 2008, **36**, 3765; (*c*) M. Paramasivam, A. Membrino, S. Cogoi, H. Fukuda, H. Nakagama and L. E. Xodo, *Nucleic Acids Res.*, 2009, **37**, 2841.
7 D. Sun, W. J. Liu, K. Guo, J. J. Rusche, S. Ebbinghaus, V. Gokhale and L. H. Hurley, *Mol. Cancer Ther.*, 2008, **7**, 880.
8 S. L. Palumbo, R. M. Memmott, D. J. Uribe, Y. Krotova-Khan, L. H. Hurley and S. W. Ebbinghaus, *Nucleic Acids Res.*, 2008, **36**, 1755.
9 A. Lew, W. J. Rutter and G. C. Kennedy, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 12508.
10 (*a*) S. Balasubramanian and S. Neidle, *Curr. Opin. Chem. Biol.*, 2009, **13**, 345; (*b*) S. Neidle, *Curr. Opin. Struct. Biol.*, 2009, **19**, 239; (*c*) N. Luedtke, *Chimia*, 2009, **63**, 134; (*d*) D. Monchaud and M. P. Teulade-Fichou, *Org. Biomol. Chem.*, 2008, **6**, 627; (*e*) T. M. Ou, Y. J. Lu, J. H. Tan, Z. S. Huang, K. Y. Wong and L. Q. Gu, *ChemMedChem*, 2008, **3**, 690; (*f*) P. Ragazzon and J. B. Chaires, *Methods*, 2007, **43**, 313.
11 (*a*) D. P. Gonçalves, R. Rodriguez, S. Balasubramanian and J. K. Sanders, *Chem. Commun.*, 2006, 4685; (*b*) L. Zhang, J. Huang, L. Ren, M. Bai, L. Wu, B. Zhai and X. Zhou, *Bioorg. Med. Chem.*, 2008, **16**, 303.
12 (*a*) I. Haq, J. O. Trent, B. Z. Chowdhry and T. C. Jenkins, *J. Am. Chem. Soc.*, 1999, **121**, 1768; (*b*) J. Ren and J. B. Chaires, *Biochemistry*, 1999, **38**, 16067; (*c*) M. W. Freyer, R. Buscaglia, K. Kaplan, D. Cashman, L. H. Hurley and E. A. Lewis, *Biophys. J.*, 2007, **92**, 2007–15; (*d*) H. Han, D. R. Langley A. Rangan and L. H. Hurley, *J. Am. Chem. Soc.*, 2001, **123**, 8902; (*e*) G. N. Parkinson, R. Ghosh and S. Neidle, *Biochemistry*, 2007, **46**, 2390.
13 J. Alzeer, P. Roth and N. W. Luedtke, *Chem. Commun.*, 2009, 1970.
14 E. Hoffman, S. Trusko, M. Murphy and D. George, *Proc. Natl. Acad. Sci. U. S. A.*, 1990, **87**, 2705.

Supporting Information

Cellular uptake and binding of guanidine-modofied phthalocyanines to *KRAS*/*HRAS* G-quadruplexes

Alexandro Membrino,[a] Manikandan Paramasivam,[a] Susanna Cogoi,[a] Jawad Alzeer,[b] Nathan W. Luedtke*[b], and Luigi E. Xodo*[a]

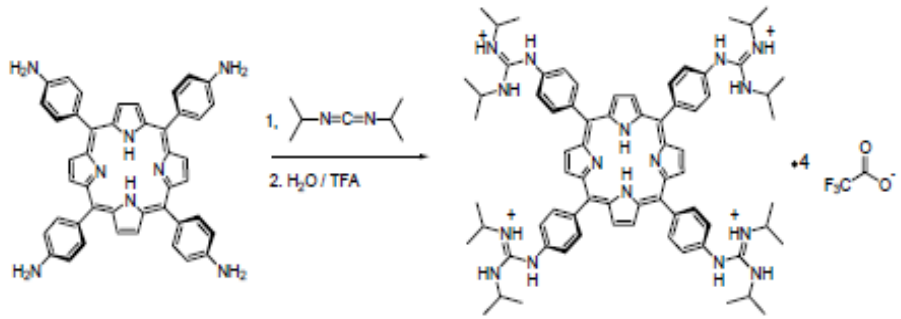
[a] *Department of Biomedical Science and Technology , University of Udine, 33100 Udine, Italy;* [b] *Institute of Organic Chemistry, University of Zürich, Winterthurerstrasse 190, CH-8057, Zürich, Switzerland.*

E-mail:luigi.xodo@uniud.it;luedtke@oci.uzh.ch

Materials and Methods

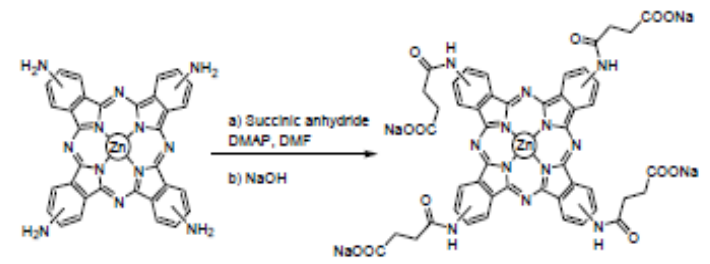
Compounds 1 – 4

The synthesis and characterization of tetrakis-(diisopropyl-guanidine) phthalocyanine "DIGP" (1), and its Zn-containing derivative "Zn-DIGP" (2) have been reported (reference 13 in the main manuscript). Compounds (3) and (4) were synthesized at the University of Zürich according to the following procedures:



5,10,15,20-tetrakis (4-aminophenyl)-21H,23H-porphine — 5,10,15,20-tetrakis (diisopropyl-guanidine)-21H,23H-porphine . (TFA)4 salt (3)

5,10,15,20-tetrakis(diisopropyl-guanidine)-21H,23H-porphine . $(TFA)_4$ salt (DIGPor, 3): The starting material 5,10,15,20-tetrakis (4-aminophenyl)-21H,23H-porphine was obtained from TCI Europe and 35 mg (52 µmoles) was combined with pyridine (4 mL), pyridine-HCl (2 g), and diisopropylcarbodiimide (500 µL, 3.3 mmoles, 59 equiv) and stirred under N_2 at 110 °C for 18 h. The reaction was removed from the heat, and 15 mL of H_2O was used to transfer the hot mixture into a polypropylene centrifuge tube. TFA (1 mL) was added, mixed, and the resulting precipitate was collected by centrifugation at 6'500 r.p.m. The precipitate was suspended into water (2 mL), sonicated, and TFA (40 µL) added. The resulting precipitate was collected by centrifugation at 6'500 r.p.m. This was repeated a total of three times. The resulting precipitate was dissolved into 1.2 mL of 1:1 acetonitrile / water and lyophilised to yield 55 mg (65 %) of a red powder. 1H-NMR (400 MHz, d_6-DMSO / d_4-methanol, 10 : 1 mixture) δ 8.89 (br s, 8H), 8.15 (d, J = 8.4 Hz, 8H), 7.55 (d, J = 8.4 Hz, 8H), 4.05 (m, J = 6.5 Hz, 8H), 1.26 (d, J = 6.5 Hz, 48H), ESI MS (m/z): [M+H]+ calcd for $C_{72}H_{91}N_{16}$, 1180; found 1180.



tetraamino-zinc-phthalocyanine — tetrakis(succinamic acid)-zinc-phthalocyanine • Na_4 salt (4)

Tetrakis(succinamic acid)-zinc-phthalocyanine . Na_4 salt (4): At room temperature and N_2 atmosphere, succinic anhydride (196 mg, 1.96 mmol) and dimethylaminopyridine (138 mg, 0.3 mmol) was added, at room temperature, to a solution of tetraamino-zinc-phthalocyanine (50 mg, 0.078 mmol) in DMF (6 mL). After 7 days, the reaction mixture was diluted with EtOAc (66 mL) and the resulting precipitate was collected by centrifugation. The dark green precipitate was washed repeatedly with H_2O. The precipitate was dissolved in TFA (4 mL) and then mixed water (20 mL) and centrifuged. The resulting ppt was dissolved in 1 N NaOH (50 mL), precipitated with MeOH (100 mL) and dried to afford 4 (76 mg, 86%) as a dark green solid. 1H-NMR (300 MHz, d_6-DMSO) δ 8.93 (br. s, 4H), 8.13 (br. m, 4H), 8.01 (br. s, 4H), 6.66 (br. m, 4H), 2.92 (br. s, 8H), 2.81 (br. s, 8H). MALDI TOF MS (*m/z*): $[M+H]^+$ calcd for $C_{48}H_{36}N_{12}O_{12}Zn$, 1037.19; found 1037.2. UV-Vis (DMSO) λmax (nm) and ε ($cm^{-1}M^{-1}$): 360 (6.1×10^4), 630 (2.57×10^4), and 690 (1.32×10^3).

Oligonucleotides and porphyrins

The oligonucleotides were obtained from Microsynth (Switzerland). They were purified by PAGE using a denaturing 20% gel (acrylamide: bisacrylamide, 19:1) in TBE, 7 M urea, 55°C. The bands were excised from the gel and eluted in water. The DNA solutions were filtered (Ultrafree-DA, Millipore) and precipitated. The concentration of each DNA was determined from the absorbance at 260 nm in milli Q water, using as extinction coefficients 7500, 8500, 15000 and 12500 $M^{-1}cm^{-1}$ for C, T, A and G, respectively. Dual-labeled F-28R-T (5'end with FAM, 3'end with TAMRA) was HPLC purified. Porphyrins TMPyP2 (P2), TMPyP3 (P3) were purchased from Porphyrin Systems (Lübeck, Germany), TMPyP4 (P4) from Sigma (Milan, Italy).

Polymerase stop assay

Single-stranded DNA fragments with a number of nt between 79-82, containing in the middle a quadruplex forming G-rich element from the murine *KRAS* or human *HRAS* promoters, were used as templates in the Taq polymerase primer-extension reactions. The DNA sequences have been purified by PAGE under denaturing conditions. The template (25 nM) was mixed with the ^{32}P-labelled primer (25 nM), in the presence or absence of porphyrins (P2, P3, P4, PP4 or 4) or phthalocyanines (**1-4**), in 25 mM KCl, Taq buffer 1X and incubated overnight at 37°C. The primer extension reactions have been carried out for 1h, by adding 10 mM DTT, 100 mM dATP, dGTP, dTTP, dCTP and 3.75U of Taq polymerase (Euro Taq, Euroclone, Milan). The reactions were stopped by adding an equal volume of stop buffer (95% formamide, 10mM EDTA, 10mM NaOH, 0.1% xylene cyanol, 0.1% bromophenol blue). The products were separated on a 12% polyacrylamide sequencing gel prepared in TBE1X, 8 M urea. The gel was dried and exposed to autoradiography. Standard dideoxy sequencing reactions were performed to detect the exact positions in which DNA polymerase was arrested.

Circular Dichroism

CD spectra were obtained using a JASCO J-600 spectro-polarimeter equipped with a thermostatted cell holder. The oligonucleotides used for the CD experiments were at a concentration of 3 μM, in 50 mM Tris-HCl, pH 7.4 and 25 mM KCl (or 100 mM NaCl when duplex DNA was used). Spectra were recorded in 0.5 cm quartz cuvette. A thermometer placed in the cuvette holder allowed a precise measurement of the sample temperature. The spectra were calculated with J-700 Standard Analysis software (Japan Spectroscopic Co., Ltd) and are reported as ellipticity (mdeg) versus wavelength (nm). Each spectrum was recorded three times, smoothed and subtracted from the baseline.

FRET spectroscopy

Fluorescence measurements were carried out with a Microplate Spectrofluorometer System (Molecular Devices) using a 96-well black plate, in which each well contained 50 μl of 200 nM dual-labelled Fam-GGGAGGGAGGGAAGGAGGGAGGGAGGGA-Tamra in 50 mM Tris-HCl, pH 7.4 and increasing amounts of KCl. In the presence of KCl, DNA assumes a folded quadruplex conformation and FRET is expected between the 5' and 3' fluorophores.

The interaction of the G-quadruplex with the phthalocyanines can be studied by following the quenching of FAM emission at 584 nm by the added GPcs, as they absorb at 584 nm (see spectrum of DIGP reported below). For the fluorescence quenching data we obtained the fraction of bound GPc at increasing ligand concentrations. Plotting the fraction of bound GPc (moles of bound GPc divided by total number of DNA moles) against the free ligand concentrations we obtained experimental points that were best-fitted to a standard binding equation $y=B_{max}\ L/(K_D+L)$ by using SigmaPlot 10.0.1.

The binding data obtained from polymerase stop assays were analysed according to Scatchard equation $r/[L]=n\ K_A - rK_A$ where r is the ratio of the moles of bound ligand divided by the total available binding sites, [L] is the concentration of free ligand, and *n* is the number of binding sites per DNA molecule.

Table S_1: Apparent dissociation constants of GPcs **1** and **2** binding to quadruplex DNA[a]

	KCl, mM	K_D (**1**), M	K_D (**2**), M
m*KRAS*	100	$8.6\ (\pm 2,1) \times 10^{-7}$	$8.9(\pm 1.7) \times 10^{-7}$
PTHR1	100	$5.8\ (\pm 1.3) \times 10^{-7}$	$3.2(\pm 0.8) \times 10^{-7}$
HRAS-1	100	$3.7\ (\pm 3.4) \times 10^{-7}$	-

[a] Data obtained using FRET constructs in 50 mM Tris-HCl, pH 7.4, 100 mM KCl. Higher apparent affinities were observed in 50 mM KCl (Table S_1).

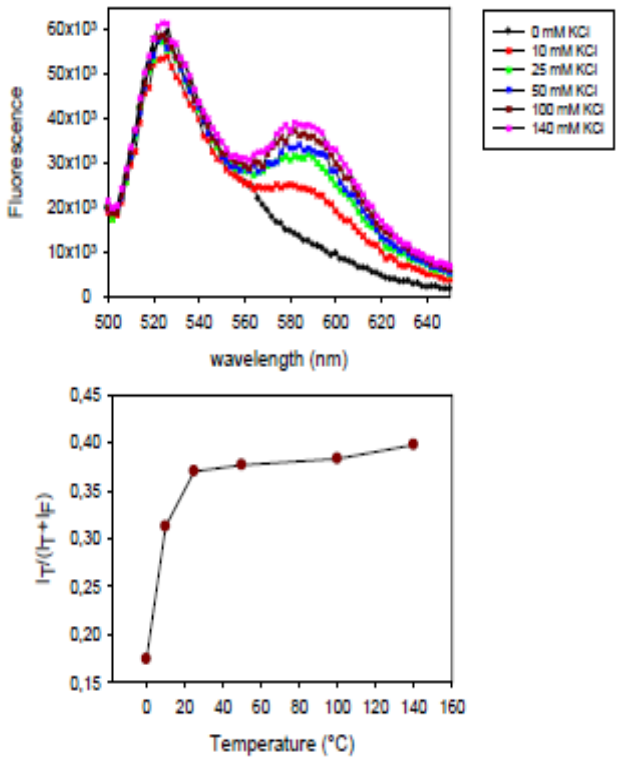


Figure S_1: Folding of murine GA-element followed by FRET as a function of KCl concentration. At 25 mM, the sequence is folded into its characteristic quadruplex structures. GA-element has been doubled labelled: 5'Fam-GGGAGGGAGGGAAGGAGGGAGGGAGGGA-tamra

Figure S_2. Taq polymerase stop assay of the DNA template containing the murine KRAS GA-element in the absence (lane 1) and presence of 0.5, 1 and 2.5 μM phthalocyanine **1** (lanes 2-4) or **2** (lanes 5-7). Buffer: 50 mM Tris-HCl, pH 7.4, 25 mM LiCl. Reaction conditions: template 50 nM, primer 50 nM, 1 h reaction at 37°C. Reaction products run in 12% PAGE in TBE-urea.

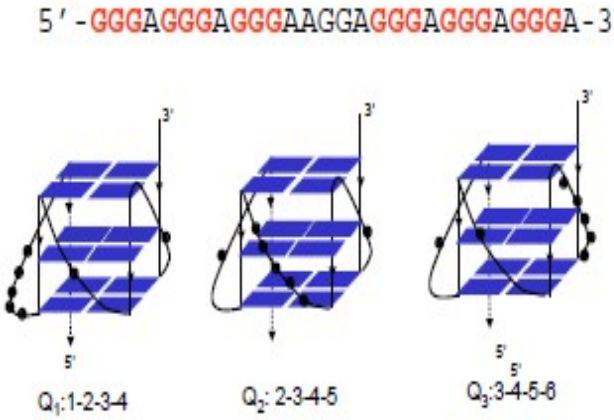


Figure S_3: Putative structure of the G4-DNA formed by the murine GA-element, as determined by DMS-footoprinting (Nucleic Acids Res, 2006, 34, 2536).

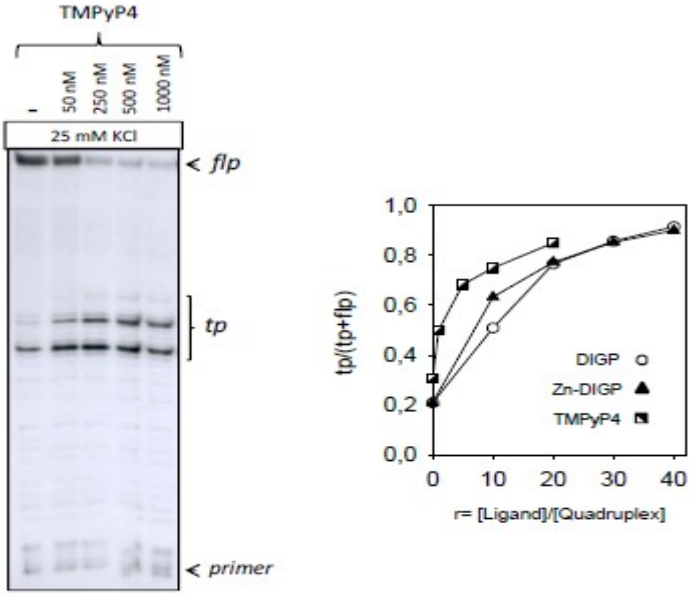


Figure S_4: (Left) Taq polymerase stop assay of the DNA template containing the murine KRAS GA-element in the absence (lane 1) and presence of 50, 250, 500 and 1000 nM TMPyP4. Buffer: 50 mM Tris-HCl, pH 7.4, 25 mM KCl. Reaction conditions: template 50 nM, primer 50 nM, 1 h reaction at 37°C. Reaction products run in 12% PAGE in TBE-urea. (Right) Plots showing the fraction of truncated products as a function of r for DIGP, Zn-DIGP and TMPyP4.

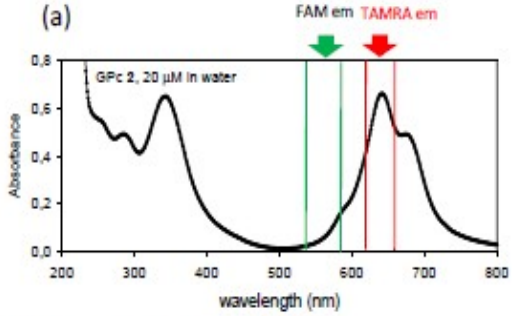


Figure S_5: (a) Absorption spectrum of phthalocyanine **2** in water;

Human parathyroid hormone PTH/PTH-related peptide receptor (PTHR1) gene

FAM-CCCGGGAGGGCGCCGGGGGAGGGAAGA-TAMRA

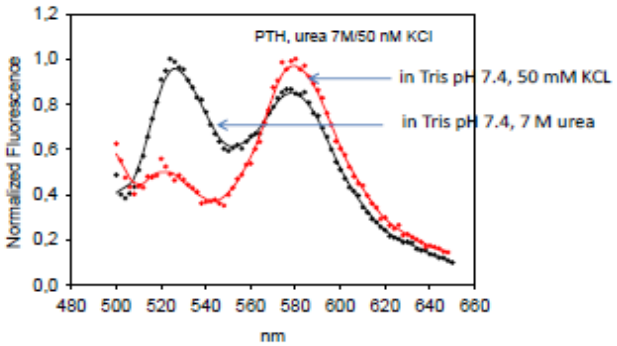


Figure S_6: Sequence of the human parathyroid hormone PTH/PTH-related peptide receptor (PTHR1) gene, forming a parallel G-quadruplex. The figure shows that the PTHR1 sequence folds in KCl, even in the presence of 7 M urea. In 50 mM Tris-HCl, pH 7.4, 50 mM KCl, the energy transfer $P=I_T/(I_T+I_F)$ is 0.66, indicating the folding of the sequence into a G-quadruplex with a T_M=67°C. As expected, in CsCl or LiCl this transition is not observed.

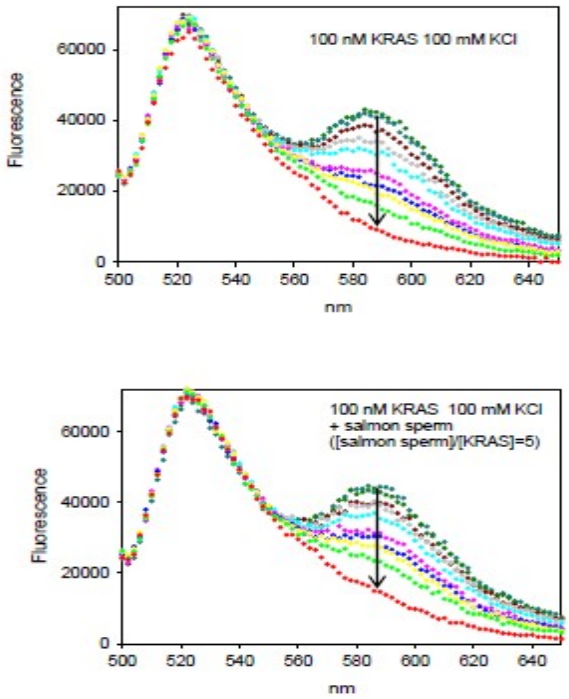


Figure S_7: Titration of 100 nM F-28-T in 50 mM Tris-HCl, pH 7.4, 100 mM KCl with increasing amounts of DIGP **1** in the absence (top) and presence (bottom) of salmon sperm DNA ([ssDNA]/[F-28R-T]=5).

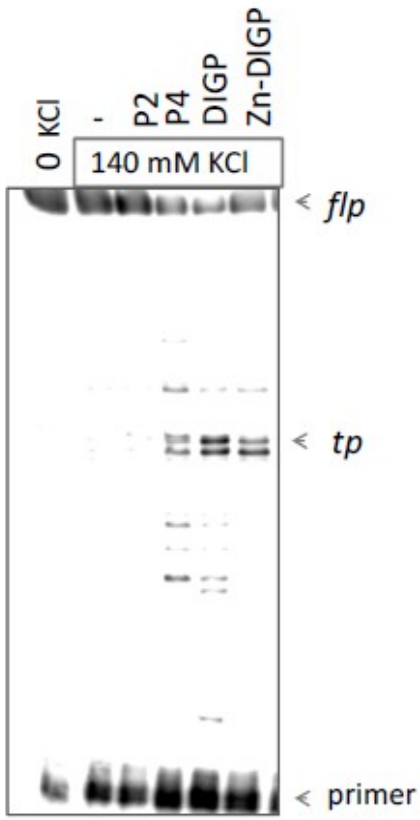


Figure S_8. Polymerase stop assays of 50 nM 82-mer template containing the *hras*-1 sequence forming an antiparallel G-quadruplex incubated with primer and GPc at r=10. Primer extension with Taq polymerase was carried out for 1 h at 37°C. P2=TMPyP2, P4= TMPyP4; flp= full length product, tp= truncated products.

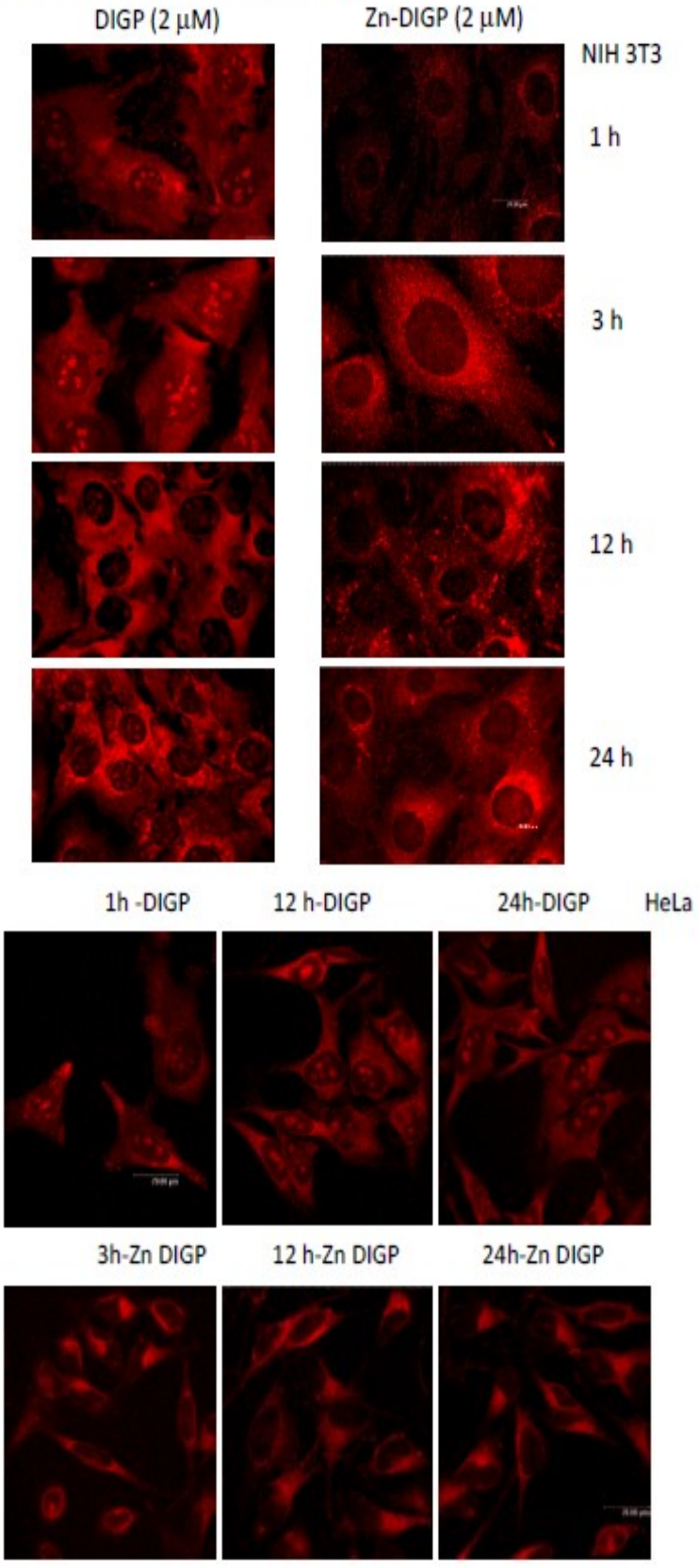


Figure S9. NIH 3T3 (top) and HeLa (bottom) cells treated for 1, 3, 12 and 24 h with 2 μM DIGP and Zn-DIGP, fixed on glass slides, and analysed by confocal microscopy. Excitation= 633 nm, emission= 680-800 nm.

4.3. The *KRAS* promoter responds to Myc-associated zinc finger and poly(ADP-ribose) polymerase 1 proteins, which recognize a critical quadruplex-forming GA-element.

KRAS promoter contains a G-rich nuclease hypersensitive element upstream of the TSS essential for transcription; by means of pulldown and chromatin immunoprecipitation assays we showed that the quadruplex is recognized by the Myc Associated Zinc finger (MAZ) and Poly(ADP-ribose) Polymerase-1 (PARP-1) proteins. These proteins are involved in transcription because, when down-regulated by shRNA, *KRAS* transcription is inhibited. On the basis of FRET-melting and polymerase stop assays we saw that MAZ stabilizes the *KRAS* quadruplex. When this folding capacity is abrogated by specific point mutations, *KRAS* transcription is down-regulated. On the other hand guanidine-modified phthalocyanines, which specifically interact with and stabilize the *KRAS* quadruplex push the promoter activity up to more than double.

These data support a transcription mechanism for murine *KRAS* that involves MAZ, PARP-1 and duplex-quadruplex conformational changes showing how *KRAS* promoter is able to behave as an enhancer of transcription when folded into quadruplex structure recruiting nuclear factors essential for transcription activation.

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 285, NO. 29, pp. 22003–22016, July 16, 2010

The *KRAS* Promoter Responds to Myc-associated Zinc Finger and Poly(ADP-ribose) Polymerase 1 Proteins, Which Recognize a Critical Quadruplex-forming *GA*-element*[S]

Received for publication, January 13, 2010, and in revised form, May 8, 2010 Published, JBC Papers in Press, May 10, 2010, DOI 10.1074/jbc.M110.101923

Susanna Cogoi[‡], Manikandan Paramasivam[‡], Alexandro Membrino[‡], Kazunari K. Yokoyama[§], and Luigi E. Xodo[‡1]

From the [‡]Department of Biomedical Science and Technology, University of Udine, School of Medicine, Piazzale Kolbe 4, 33100 Udine, Italy and the [§]Center of Excellence for Environmental Medicine, Kaohsiung Medical University, 100 Shih-Chuan 1st Road, 80708 Kaoshsiung, Taiwan

The murine *KRAS* promoter contains a G-rich nuclease hypersensitive element (*GA*-element) upstream of the transcription start site that is essential for transcription. Pulldown and chromatin immunoprecipitation assays demonstrate that this *GA*-element is bound by the Myc-associated zinc finger (MAZ) and poly(ADP-ribose) polymerase 1 (PARP-1) proteins. These proteins are crucial for transcription, because when they are knocked down by short hairpin RNA, transcription is down-regulated. This is also the case when the poly(ADP-ribosyl)ation activity of PARP-1 is inhibited by 3,4-dihydro-5-[4-(1-piperidinyl) butoxyl]-1(2H) isoquinolinone. We found that MAZ specifically binds to the duplex and quadruplex conformations of the *GA*-element, whereas PARP-1 shows specificity only for the *G*-quadruplex. On the basis of fluorescence resonance energy transfer melting and polymerase stop assays we saw that MAZ stabilizes the *KRAS* quadruplex. When the capacity of folding in the *GA*-element is abrogated by specific G → T or G → A point mutations, KRAS transcription is down-regulated. Conversely, guanidine-modified phthalocyanines, which specifically interact with and stabilize the *KRAS G*-quadruplex, push the promoter activity up to more than double. Collectively, our data support a transcription mechanism for murine *KRAS* that involves MAZ, PARP-1 and duplex-quadruplex conformational changes in the promoter *GA*-element.

Guanine-rich sequences have the potential to fold into *intra*molecular *G*-quadruplex (or G4-DNA) structures that are stabilized by planar arrays of four guanines paired by Hoogsteen hydrogen bonds (G-tetrad) (1). Quadruplex-forming sequences (QFS)[2] are present in prokaryotic and eukaryotic genomes, promoter regions, micro- and mini-satellite repeats, telomeres, rDNA, and the vertebrate immunoglobulin heavy chain switch regions (2). Recent bioinformatic search analyses have shown a surprisingly high presence in the human genome of QFS, on the order of $3–4 \times 10^5$ (3, 4). The gene distribution of QFS is highly skewed because tumor suppressor genes have a very low level of QFS, whereas proto-oncogenes have a high level of such sequences (5). There seems to be a correlation between QFS and genomic instability; a low level of QFS in tumor suppressor genes is associated with genomic stability, and a high level is associated with genomic instability (5). Furthermore, the observation that QFS are often located in the region surrounding the transcription start sites of the genes and within *cis*-elements suggests that they may be involved in transcription regulation. This hypothesis has been formulated for a number of genes including *CMYC, KRAS, C-MYB, VEGF, PDGFA, CKIT,* and human insulin (6–13). The best studied G-rich sequence folding into a *G*-quadruplex, whose function has been correlated with a mechanism of transcription regulation, is the one found in the promoter of the *CMYC* gene (6). Upstream of the P1 promoter, controlling about 80% of transcription, there is a QFS that can fold into a *G*-quadruplex. When this quadruplex is destabilized by a G-to-A mutation, the basal transcription increases 3-fold, but when the G-quadruplex is stabilized by cationic porphyrins, transcription is repressed. From these data it has been suggested that the *G*-quadruplex formed in the *CMYC* promoter may function as a transcription repressor. A comprehensive review on the mechanism controlling transcription in *CMYC* has been recently reported (14). Interestingly, G-rich elements do not seem to be a unique feature of mammalian genomes, as they are also over-represented in *Saccharomyces cerevisiae*. For instance, it has been reported that the gene promoters in this organism contain a level of G-rich sequences, which is 14-fold higher than the entire genome (15). A mutant of *S. cerevisiae* lacking the SGS1 gene, encoding for a helicase specific for quadruplex DNA, was characterized by a reduced expression of the genes in which open reading frames have a high potential to form quadruplexes (16). In addition, the human pathogen *Neisseria gonorrhoeae* contains a quadruplex-forming 16-base pair G-rich sequence that is required to promote pilin antigenic variation of

* This work was supported by Associazione Italiana per la Ricerca sul Cancro 2008 and Italian Ministry of Scientific Research (PRIN 2007).

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Table S1 and Figs. S1–S8.

[1] To whom correspondence should be addressed. Tel.: 39-0432-494395; Fax: 39-0432-494301; E-mail: luigi.xodo@uniud.it.

[2] The abbreviations used are: QFS, quadruplex-forming sequence; MAZ, myc-associated zinc finger protein; PARP-1, poly(ADP-ribose) polymerase; DPQ, (3,4-dihydro-5-[4-(1-piperidinyl)butoxyl]-1(2*H*)-isoquinolinone); DIGP, tetrakis-(diisopropyl-guanidine) phthalocyanine; shRNA, short hairpin RNA; FRET, fluorescence resonance energy transfer; CAT, chloramphenicol acetyltransferase; FAM, 6-carboxyfluorescein; GPc, guanidine-modified phthalocyanine; ChIP, chromatin immunoprecipitation; CMV, cytomegalovirus; EMSA, electrophoretic mobility shift assay; GST, glutathione *S*-transferase; PBS, phosphate-buffered saline; DTT, dithiothreitol; TBE, Tris borate EDTA; BSA, bovine serum albumin; SucPc, tetrakis-(succinyl)-phthalocyanines.

surface structures to avoid immune detection (17). Evidence that G4-DNA formed within the regulatory elements of promoters may be involved in transcription regulation has also been produced by other research groups (6–13). Nevertheless, whether quadruplex DNA is formed inside the cell or whether it has any biological function in cellular events such as transcription or recombination is still a matter of debate. The role of *G*-quadruplex DNA on transcription is essentially based on the observation that *G*-quadruplex destabilizing point mutations or deletions bring about a significant enhancement of transcription, whereas G4-DNA ligands repress transcription (6–13). Although these findings support the notion that G4-DNA may be a transcription repressor, they do not provide a conclusive answer about the cellular function of quadruplex DNA. Indeed, it can be argued that the mutations introduced into the guanine element of the promoter could abrogate the binding of proteins to DNA and consequently alter the level of transcription. It could also be that the quadruplex-stabilizing ligands affect transcription simply because they compete with the binding of the proteins to the promoter. To provide an answer to these questions, we have in the present paper extended our previous work (7) and analyzed the role played in transcription by a critical *GA*-element located in the promoter of the murine *KRAS* proto-oncogene (18). This promoter contains a 34-bp G-rich sequence between nucleotides −322 and −288 (−1 is the 3′ boundary of the exon 0), which is sensitive to nuclease S1 and able to fold into a parallel *G*-quadruplex conformation. This intramolecular quadruplex was characterized by circular dichroism and UV spectroscopy, dimethyl sulfate footprinting, polymerase stop assays, and electrophoresis (7). The *GA*-element contains two perfect copies of the consensus sequence for the myc-associated zinc finger (MAZ) transcription factor. By pulldown and chromatin immunoprecipitation (ChIP) assays, we demonstrated that not only MAZ but also PARP-1 is associated to the *GA*-element. Although MAZ binds to both duplex and quadruplex conformations of the *GA*-element, PARP-1 binds specifically only to the *G*-quadruplex. We provide here compelling evidence that MAZ and PARP-1 are transcription activators and propose that these proteins are recruited to the *KRAS* promoter by the *G*-quadruplex structure formed within the *GA*-element.

EXPERIMENTAL PROCEDURES

Site-specific Mutagenesis; Wild-type and Mutant Plasmid Construction—Plasmid pKRS413 harboring the CAT gene driven by the murine *KRAS* promoter was subjected to PCR using primers 5′-TGCAGCCGCTCCCTCTCTCTCCTTCTCTCTCCCGCGCG and 5′-GAGGGAGCGGCTGCAGCGCTGGGAG (to introduce four G → A point mutations) and 5′-TGCAGCCGCTCCCTCACTCACTCCTTCCCT and 5′-GAGGGAGCGGCTGCAGCGCTGGGAG (to introduce two G → T point mutations). The mutant plasmid with two T residues was amplified with 5′-CTCACTCACTCCTTCACTCACTCCCGCGCG and 5′-GAAGGAGTGAGTGAGGGAGCGGCTGCAG to obtain a mutant with four T residues. PCR was performed with 3 ng/μl DNA template, 0.1 μM each primer, 0.05 units/μl AccuPrime pfx DNA polymerase (Invitrogen) in 1× AccuPrime pfx reaction mix for 3 min at 95 °C, 30 cycles 1 min at 95 °C, 30 s at 65 °C, and 5 min at 68 °C. Bacteria DIH 101 were transformed with PCR product, and DNA was extracted and sequenced (primer 5′-CCTCTCGGCACCACCCTC, accession number U49448, Entrez Nucleotide database (NCBI), bases 403–420). A 300-bp XmaI-AvaII fragment containing the *KRAS* murine promoter from the wild-type and mutant plasmids was subcloned in pGL3–1B basic in XmaI-HindIII blunted site in order to have the reporter firefly luciferase gene driven by wild-type and mutant *KRAS* promoters. The inserts were sequenced with pGL3for primer (5′-CTAGCAAAATAGGCTGTCCC). The plasmids constructed in this way were pKRS413-luc, pKRS413-luc-2T, pKRS413-luc-4T, and pKRS413-luc-4A.

Cell Culture and Proliferation Assay—NIH 3T3 cells were maintained in exponential growth in Dulbecco's modified Eagle's medium containing 100 units/ml penicillin, 100 mg/ml streptomycin, 20 mM L-glutamine, and 10% fetal bovine serum (Euroclone, Milano, Italy). Cell viability was measured by resazurin assays following standard procedures.

Dual Luciferase Assay—NIH 3T3 cells were seeded the day before transfection at 8000 cells/well in a 96-well plate. When the experiment involved a phthalocyanine treatment, this was done on cells let for adhesion for 16 h. Transfection was performed by mixing each vector 250 ng/well with control plasmid pRL-CMV expressing *Renilla* luciferase under control of the CMV promoter 10 ng/well using JetPEI transfection reagent (Polyplus transfection) following the manufacturer's instructions. For cotransfection with pCMV-MAZ (19), 150 ng of pKRS413-luc or mutant with 10 ng of pRL-CMV were transfected with either 150 ng of pCMV-MAZ or pcDNA3 plasmid (empty vector) as the mass for control transfections. Each transfection was performed in triplicate. Luciferase assays were performed 48 h after transfection with the Dual-Glo Luciferase assay system (Promega, Milan, Italy) following instructions. Samples were read with Turner Luminometer and expressed as relative luciferase, *i.e.* $R_T/R_C \times 100$, where R_T and R_C are (firefly luciferase)/(*Renilla* luciferase) in phthalocyanine-treated and untreated cells.

Recombinant MAZ Expression and Electrophoretic Mobility Shift Assays (EMSA)—Recombinant MAZ protein tagged to GST was expressed in *Escherichia coli* BL21 using plasmid pGEX-hMAZ (19). The bacteria were grown for 1–2 h at 37 °C to an A_{600} of 0.5–2.0 before induction with isopropyl 1-thio-β-D-galactopyranoside (1.5 mM final concentration). Cells were allowed to grow for 7 h before harvesting. The cells were centrifuged at 5000 rpm, 4 °C, the supernatant was removed, and the cells washed twice with PBS. The pellet was resuspended in a solution of PBS with 100 mM phenylmethylsulfonyl fluoride, 1 M DTT, and protease inhibitor cocktails (for general use and for purification of histidine-tagged protein (Sigma) according to the manufacturer's protocols). The bacteria were lysed by sonication, added with Triton X-100 (1% final concentration), and incubated for 30 min on a shaker at 4 °C. The lysate was then centrifuged for 1 h at 4 °C at 20,000 rpm. Glutathione-Sepharose 4B (GE Healthcare) (50% slurry in PBS) was added to the supernatant from the previous step and incubated for 30 min at 4 °C on a shaker. The mix was centrifuged for 5 min at 500 × *g*, and the pellet was washed 5 times with PBS. The protein was eluted from the pellet with elution buffer containing 50 mM

Tris-HCl, pH 8, and 10 mM reduced glutathione by centrifugation for 5 min at 500 g, 4 °C, and the tagged proteins collected from the supernatant were checked by SDS-PAGE.

Protein-DNA interactions were analyzed by EMSA. 5 nM radiolabeled *GA*-duplex, GA-duplex (2T), GA-duplex (4T), or quadruplex 28R (see Table 1) were incubated with MAZ protein (or NIH 3T3 nuclear extract) as indicated in legends for Figs. 4, 5, and 9 for 30 min at room temperature in 20 mM Tris HCl, pH 8, 30 mM KCl, 1.5 nM $MgCl_2$, 1 mM DTT, 8% glycerol, 1% phosphatase Inhibitor Cocktail I (Sigma), 5 mM NaF, 1 mM Na_3VO_4, 2.5 ng/μl poly(dI-dC). The analyses were carried out in 5% polyacrylamide gels in 1× TBE at 20 °C.

Pulldown and Western Blotting Assays—One milligram of nuclear protein extract (0.25 mg/ml), prepared as described (8), was incubated for 1 h at 37 °C with 60 nM biotinylated G4-DNA (G4-biotin, prepared in 100 mM KCl) or biotinylated duplex (G4-biotin annealed with its complementary in 100 mM NaCl) (Table 1) in a solution containing 20 mM Tris-HCl, pH 8, 8% glycerol, 150 mM KCl, 25 ng/ml poly(dI-dC), 1 mM Na_3VO_4, 5 mM NaF, 1 mM DTT, 0.2 mM phenylmethylsulfonyl fluoride, 0.1 mM zinc acetate. Then 250 μg of Streptavidin MagneSphere Paramagnetic Particles (Promega) pretreated for 30 min with 0.25 mg/ml BSA were added and incubated for 30 min at 37 °C. Particles were captured with a magnet, and the proteins were eluted with the buffer containing 0.5 and 1 M NaCl. The eluted proteins were concentrated and desalted with Microcon YM-3 filters (Millipore, Billerica, MA) for further analyses. The eluted proteins were separated by 10% SDS-PAGE and blotted overnight in 25 mM Tris, 192 mM glycine, and 20% methanol at 4 °C on a nitrocellulose membrane. The membrane was incubated with different antibodies: MAZ H-50 and PARP-1 H-300 diluted 1:200 (Santa Cruz). The secondary antibody used was rabbit IgG peroxidase conjugate (1:10,000) (Calbiochem). The antibodies were diluted in 10 mM Tris, pH 7.9, 150 mM NaCl, 0.05% Tween, and 5% BSA. The signal was developed with Super-Signal West Pico or Femto (Pierce) and detected with ChemiDOC XRS, Quantity One 4.6.5 software (Bio-Rad).

ChIP Assay—NIH 3T3 cells were plated in 2 × 15-cm diameter plates, grown to 80% confluence (about 1.5×10^7 cells), and fixed in formaldehyde 1% in PBS for 2 or 5 min. Nuclei were isolated following the ChIP-IT™ Express kit (Active Motif, Rixensart, Belgium) and resuspended in radioimmune precipitation assay buffer. Sonication was performed with a Bioruptor™ sonicator (Diagenode, Liege, Belgium) following the manufacturer's instructions to obtain 200–500-bp fragments. Chromatin immunoprecipitation, washes, elution, reverse cross-linking, and proteinase K treatment were performed following the ChIP-IT™ Express kit manual. Antibodies used were MAZ H-50x and PARP-1 H-300 (Santa Cruz), 100 and 20 ng/μl, respectively. Control antibodies were RNA polymerase II mouse monoclonal antibody and negative control mouse IgG (ChIP-IT™ Control kit-mouse, Active Motif). About 15 μg of chromatin was used for each sample and fixed for 2 min for PARP-1 and 5 min for MAZ. The PCR reaction mixture was 1× PCR buffer (from the kit), 0.2 mM dNTPs, 0.4 μM primers, 0.04 units of Taq polymerase (EuroTaq Euroclone), and 1/10 of the ChIP samples. Primers used were EF1-α control primers (from control kit), which give a 233-bp product, and 5′-CCTCTCGGCACCACCCTC and 5′-GATGCGC-TCGGTGCTC (respectively, 403–420 and 561–546, sequence accession number U49448), which give a 160-bp product. PCR was performed as follows: 3 min at 94 °C, 40 cycles of 1 min at 94 °C, 30 s at 59 °C for EF1-α amplification and 61 °C for KRAS promoter amplification, and 30 s at 72 °C, with a final elongation 10 min at 72 °C. Amplification products were separated by a 10% acrylamide gel in TBE and visualized with a Gel-DOC apparatus (Bio-Rad).

shRNA Transfections and 3,4-Dihydro-5-[4-(1-piperidinyl) Butoxyl]-1(2H)-isoquinolinone (DPQ) Treatment—Cells were seeded 20–50,000/well in a 24-well plate. The day after plating they were either treated with DPQ at the concentrations indicated, or shRNA plasmids were transfected 0.5 μg/well. Plasmids used are Control shRNA Plasmid-A, PARP-1 shRNA Plasmid (m), and MAZ shRNA Plasmid (m) (Santa Cruz Biotechnology). Cells were collected 48 or 72 h after transfection.

RNA Extraction, Reverse Transcription, and Real-time PCR—RNA was extracted using TRIzol reagent (Invitrogen) following the manufacturer's instructions. For cDNA synthesis 5 μl of RNA in diethyl pyrocarbonate-treated water (extracted from about 25,000 cells) was heated at 55 °C and placed in ice. The solution was added to 7.5 μl of mix containing (final concentrations) 1× buffer, 0.01 M DTT (Invitrogen), 1.6 μM primer dT (MWG Biotech, Ebersberg, Germany; $d(T)_{16}$), 1.6 μM Random primers (Promega), 0.4 mM dNTPs solution containing equimolar amounts of dATP, dCTP, dGTP, and dTTP (Euroclone, Pavia, Italy), 0.8 units/μl RNase OUT, and 8 units/μl of Maloney murine leukemia virus reverse transcriptase (Invitrogen). The reactions were incubated for 1 h at 37 °C and stopped with heating at 95 °C for 5 min. As a negative control the reverse transcription reaction was performed with 5 μl of diethyl pyrocarbonate water. Real-time PCR reactions were performed with 1x iQ™ SYBR Green Supermix (Bio-Rad), 300 nm of each primer, 1 μl of reverse transcription reaction. The sequences of the primers used for amplifications are reported below. The PCR cycle was: 3 min at 95 °C, 40 cycles 10 s at 95 °C, 30 s at 60 °C with an iQ5 real-time PCR controlled by an Optical System software Version 2.0 or with CFX 96 controlled by Bio-Rad CFX Manager V1.5 (Bio-Rad). KRAS, MAZ, and PARP-1 expression are normalized with glyceraldehyde-3-phosphate dehydrogenase, hypoxanthine-guanine phosphoribosyltransferase, and β2-microglobulin. The PCR primers used are reported in supplemental Table S1).

Polymerase-stop Assay—A linear DNA sequence of 80 nucleotides, named wtR-Mur80 (Table 1), containing the G-rich element of murine *KRAS*, was used as a template for Taq polymerase primer-extension reactions. The primer used was an 18-mer sequence named pMur80 (Table 1). This DNA sequence was purified by PAGE. The template (50 nM) was mixed with the labeled primer (25 nM) in 25 mM KCl, 1× Taq buffer and incubated overnight at 37 °C. The primer extension reactions were carried out for 1 h by adding 10 mM DTT, 100 μM dATP, dGTP, dTTP, dCTP, and 3.75 units of Taq polymerase (EuroTaq, Euroclone, Milan). The reactions were stopped by adding an equal volume of stop buffer (95% formamide, 10 mM EDTA, 10 mM NaOH, 0.1% xylene cyanol,

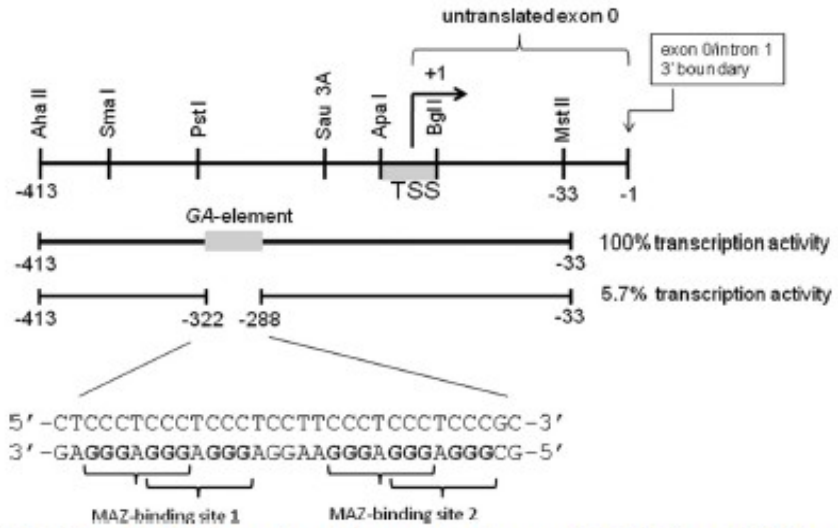


FIGURE 1. **The murine *KRAS* promoter contains a nuclease hypersensitive G-rich element (*GA*-element) that is essential for transcription.** The *GA*-element is characterized by six runs of guanines that can fold into an intramolecular *G*-quadruplex structure. The *GA*-element contains two perfect binding sites for the MAZ transcription factor. *TSS* means transcription start site. The sequence is numbered with the exon 0/intron 1 boundary taken as −1.

0.1% bromphenol blue). The products were separated on a 12% polyacrylamide sequencing gel prepared in 1× TBE, 8 M urea. The gel was dried and exposed to autoradiography. Standard dideoxy sequencing reactions were performed to detect the points where DNA polymerase I was arrested.

DNase I Footprinting—DNase I footprint was performed with the duplex obtained by annealing wtR-Mur80 with its complementary wtY-Mur80. The purine strand was end-labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase. The labeled duplex was preincubated at different ratios of MAZ-GST for 30 min in 20 mM Tris-HCl, pH 8, 30 mM KCl, 1.5 mM $MgCl_2$, 1 mM DTT, 8% glycerol, 50 μM zinc acetate, either in the presence or absence of 0.5 mM EDTA, and digested with DNase I (1 μl of a solution containing 0.002 unit of DNase I (Takara Biomedicals, Japan), 50 mM Tris-HCl, pH 7.4, 0.1 mg/ml BSA, 30 mM $MnCl_2$). The reaction was conducted for 1 min at room temperature and stopped by adding to the reaction mixture 10 μl of stop solution (90% formamide, 50 mM EDTA, bromphenol blue). The analyses were carried out in 15% polyacrylamide sequencing gel prepared in 1× TBE, 8 M urea. After running, the gel was fixed and exposed to autoradiography (Hyperfilm, GE Healthcare) at −80 °C for few hours. A standard dimethyl sulfate G-reaction was performed with wtR-Mur80 purine strand to locate the binding of MAZ within murine duplex.

Fluorescence Resonance Energy Transfer (FRET) Experiments—Fluorescence measurements were carried out with a Microplate Spectrofluorometer System (Molecular Devices, Concorde, Canada) using a 96-well black plate in which each well contained 50 μl of 200-nm dual-labeled 28R in 50 mM Tris-HCl, pH 7.4, KCl as specified in legends for Figs. 2 and 6. The samples were incubated overnight at 37 °C before measurements. The emission spectra were obtained setting the excitation wavelength at 475 nm, the cut-off was at 515 nm, and recording the emission was from 500 to 650 nm. FRET melting experiments were performed on a real-time PCR apparatus (CFX 96, Bio-Rad) using a 96-well plate filled with 50-μl solutions of dual-labeled 28R, called F-28R-T (Table 1) in 50 mM Tris-HCl, pH 7.4, and potassium chloride at different concentrations as specified on the figure. The protocol used for the melting experiments was (i) an equilibration step of 5 min at low temperature (20 °C) and (ii) a stepwise increase of the temperature of 1 °C/min for 76 cycles to reach 95 °C. All the samples in the wells were melted in 76 min. After melting, the samples were re-annealed then melted again to give the same melting curves.

RESULTS

The GA-element of the Murine KRAS Promoter Folds in One of Two Topologically Distinct G4-DNA Conformations—The promoter activity of murine *KRAS* resides within a 380-bp DNA fragment including a guanine-rich element (*GA*-element) between −322 and −288 (positions relative to exon 0/intron 1 boundary (accession number M16708)). The deletion of the *GA*-element drops the promoter activity to 5.7% that of the control (18) (Fig. 1). As it exhibits S1 nuclease sensitivity (18), the *GA*-element is likely to assume an unusual DNA structure under superhelical stress. A previous study has shown that the *GA*-element formed an intramolecular H-DNA structure, but under acidic conditions (20). We recently discovered that the polypurine strand of the *GA*-element (called 28R, Table 1), composed of six runs of guanines (G-runs), folds into an intramolecular *G*-quadruplex under physiological conditions (7). This structure is rather stable because it arrests the progression of Taq polymerase even at 50 °C, as demonstrated by polymerase stop assays with an 80-mer duplex containing the *GA*-element (Fig. 2*a*). Two polymerase arrests are observed at the adenines before the first and second G-run, at the 3′ end of the *GA*-element (supplemental Fig. S1). This suggests the formation of two *G*-quadruplexes, Q_1 or Q_2, the former involving G-runs 1–2-3–4 and the latter involving G-runs 2–3-4–5 (Fig. 2*b*). To provide further support that the *GA*-element folds in two topologically distinct *G*-quadruplex conformations, we performed FRET experiments. The polypurine strand of the *GA*-element (28R) was labeled at the 5′ and 3′ ends with a donor (6-carboxyfluorescein (FAM)) and an acceptor (tetramethylrodhamine) fluorophore (F-28R-T). In the presence of increasing amounts of KCl, F-28R-T folds into a *G*-quadruplex that upon FAM excitation at 475 nm gives rise to a FRET signal at 580 nm due to tetramethylrodhamine emission (21) (supplemental Fig. S2). The energy transfer between the two dyes is empirically measured by the *p* value, $p = I_T/(I_T + I_F)$, where I_F and I_T are the fluorescence intensities of the donor and

TABLE 1
Oligonucleotides used in this study
B, biotin.

5′-GGGAGGGAGGGAAGGAGGGAGGGAGGGA	28R
FAM-GGGAGGGAGGGAAGGAGGGAGGGAGGGA-TAMRA	F-28R-T
5′-GGGAGGGAGGGAAGGAGTGAGTGAGGGA	28R(2T)
5′-GGGAGTGAGTGAAGGAGTGAGTGAGGGA	28R(4T)
5′-TCCCTCCCTCCCTCCTTCCCTCCCTCCC	28Y
5′-TCCCTCACTCACTCCTTCCCTCCCTCCC	28Y(2A)
5′-TCCCTCACTCACTCCTTCACTCACTCCC	28Y(4A)
B-5′-CGCGCGGGAGGGAGGGAAGGAGGGAGGGAGGGAGCGGCT (accession number U49448, bases 490–528)	G4-biotin
B-5′-TACCACCGAAGAGTTAAACAGGGAGGCAAATAAGCGTGCG (accession number U49448, bases 11–50)	SRC-biotin
5′-ACCTTGATGAATCCAGGGAGGGAGGGAAGGAGGGAGGGAGGGACATCG-CTACCGTTAAAGATCATATCAAGTGATAGTAC	wtR-Mur80
5′-GTACTATCACTTGATATG	pMur80
5′-GTACTATCACTTGATATGATGCTTAACCGTAGCGATCTCCCTCCCTCC-CTCCTTCCCTCCCTCCCTGGATTCATCAAGGT	wtY-Mur80

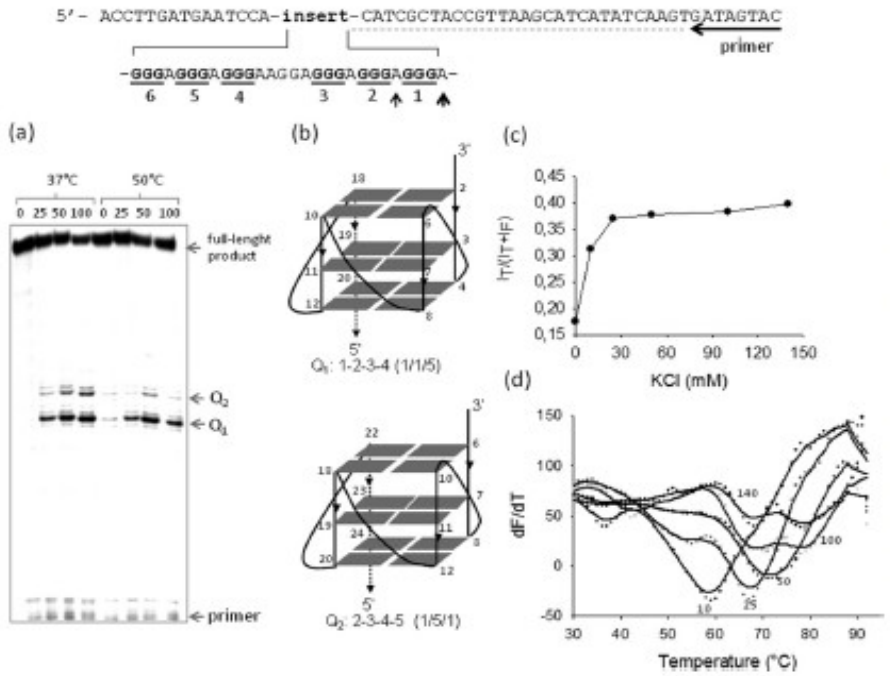


FIGURE 2. **Polymerase stop and FRET-melting assays.** The DNA template used for the polymerase stop assay is shown. *a*, the G-runs of the *GA*-element are *underlined*. Primer-extension reactions were performed at 37 and 50 °C. Two polymerase pauses were observed at the adenines (determined by Sanger sequencing, see Ref. 7 and supplemental Fig. S1) before the first and second G-run at the 3′ end of the *GA*-element. The numbers *above the gel* indicate the KCl concentrations in the samples. *b*, shown are structures of the putative *G*-quadruplexes determined by dimethyl sulfate footprinting and circular dichroism (see Ref. 7). *c*, energy transfer, $p = I_{580}/(I_{525} + I_{580})$-associated to the folding of F-28R-T as a function of KCl concentration, is shown. The experiment was performed in 50 mM Tris-HCl, pH 7.4, KCl as indicated in the figure (10, 25, 50, 100, and 140 mM). *d*, shown are FRET-melting curves obtained by measuring the increase of fluorescence at 525 nm emitted by F-28R-T as a function of temperature (excitation, 485 nm). The *numbers* indicate the KCl concentration. F-28R-T melts with a biphasic profile in the presence of 100 and 140 mM KCl.

acceptor at 525 and 580 nm, respectively. Fig. 2*c* shows that the *p* value increases with KCl concentration from 0.175 (open sequence) to 0.39 (folded sequence); in 140 mM KCl, the *p* value is not higher than 0.39 because F-28R-T forms a parallel *G*-quadruplex where the two fluorophores are located at the opposite ends of the structure and, therefore, not very close one to the other. When the *G*-quadruplex is melted, the two fluorophores increase their distance, and the FAM fluorescence at 525 nm (F_{525}) raises with temperature. Plotting dF_{525}/dt *versus* t, one obtains well resolved melting curves. In keeping with the polymerase stop assay, FRET-melting curves show that F-28R-T melts with a biphasic profile, confirming the formation of two *G*-quadruplexes with T_Ms of 68 and 80 °C in 100 mM KCl (Fig. 2*d*).

Functional Characterization of the GA-element—To functionally characterize how the *GA*-element regulates transcription, we first studied if the introduction of quadruplex-destabilizing point mutations affects the level of *KRAS* transcription. We constructed plasmid pKRS413-luc bearing the murine *KRAS* promoter driving firefly luciferase. From pKRS413-luc we obtained by site-directed mutagenesis three mutant plasmids, pKRS413-luc-2T, pKRS413-luc-4T, and pKRS413-luc-4A, in which two or four guanines located in the mid-G-tetrad of Q_1 or Q_2 have been replaced with thymines or adenines (Fig. 3, *a* and *b*). To explore whether the quadruplex-destabilizing point mutations change the level of transcription, wild-type or mutant plasmids were co-transfected with pRL-CMV, a vector encoding for *Renilla* luciferase, in NIH 3T3 cells. The levels of firefly and *Renilla* luciferases were measured with a luminometer 48 h after transfection. When firefly/*Renilla* ratio obtained with pKRS413-luc and pRL-CMV was set to 100, the same ratio given by the mutant plasmids pKRS413-luc-2T, pKRS413-luc-4T, and pKRS413-luc-4A turned out to be 60 ± 14, 44 ± 10, and 39 ± 12, respectively (Fig. 3*c*). This result shows that the abrogation of the capacity of quadruplex formation by the *GA*-element significantly decreases the expression of firefly luciferase, suggesting that the *G*-quadruplex could behave as a structural element favoring *KRAS* transcription.

MAZ Binds to the Duplex and G4-DNA Conformations of the GA-element—Considering that the *GA*-element is essential for the activation of transcription, we searched for putative tran-

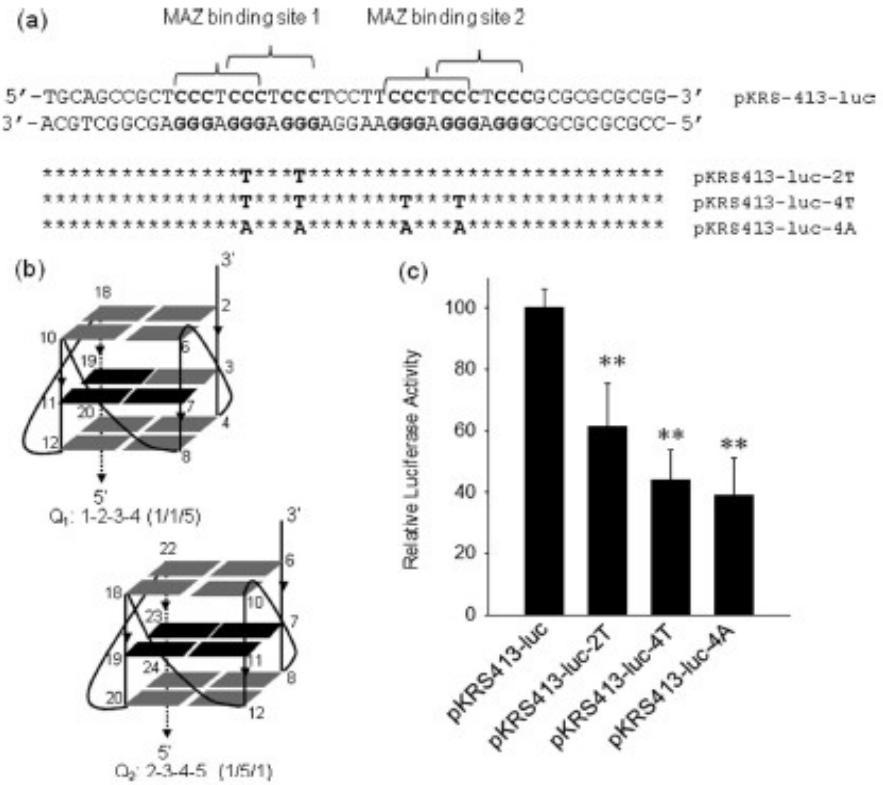


FIGURE 3. **Transcription activity of wild-type and mutant *KRAS* promoter.** *a*, shown is the sequence of the *KRAS* promoter in the region spanning over the *GA*-element. The two MAZ-binding sites are indicated by *brackets*. By site-directed mutagenesis the promoter sequence was modified in the *GA*-element either in one MAZ-binding site or in both MAZ-binding sites. The point mutations abrogated the capacity of the sequence to fold into a *G*-quadruplex. *b*, the structure of the putative *KRAS G*-quadruplexes shows that the point mutations fall in the mid G-tetrad of the structure (depicted in *black*). *c*, results of dual luciferase assay with wild-type and mutant plasmids show that the activity of the *KRAS* promoter is reduced by the introduction in the *GA*-element of the point mutations that destabilize *G*-quadruplex formation. All mutant expressions are different than wild-type expression by Student's *t* test; $p < 0,01$ (*two asterisks*).

scription factors binding to this critical sequence by using the MatInspector software. We found that the protein with the highest binding prediction is MAZ, whose consensus sequence is GGGAGGG. A tandem of two perfect MAZ-binding sites is indeed present in the *GA*-element. To find out whether MAZ actually binds to the *GA*-element, we expressed MAZ fused to GST in *E. coli* and purified the chimeric protein by affinity chromatography with glutathione-Sepharose 4B. The wild-type *GA*-element in duplex (28R:28Y, namely *GA*-duplex, Table 1) gave two retarded bands with MAZ-GST, in keeping with the binding of one protein to each of the MAZ-binding sites (Fig. 4*a*). This is consistent with the fact that (i) the introduction in the *GA*-duplex of two G → T point mutations in only one of the two MAZ-binding sites resulted in the abrogation of the DNA-protein complex of lower mobility, *i.e.* the complex where both protein sites are occupied by MAZ (slower band), and (ii) the introduction of four G → T point mutations, two in each MAZ-binding site, resulted in the abrogation of both DNA-protein complexes, as expected. This clearly demonstrates that the interaction of MAZ with the *KRAS* promoter is highly sequence-specific and that two MAZ molecules bind to the *GA*-element.

According to the luciferase transfection experiments, we cautiously hypothesized that a *G*-quadruplex in the *GA*-element could function as a transcription activator. However, it is also possible that the mutant plasmids expressed less luciferase because the G → T and G → A point mutations introduced in the *GA*-element abolished the binding of MAZ to the promoter. Therefore, mutational and EMSA studies did not allow us to understand whether repression of *KRAS* transcription was due to the loss of quadruplex formation by the *GA*-element or to the loss of the MAZ-binding sites. This suggests that the use of mutagenesis to investigate whether non-orthodox DNA secondary structures influence transcription should be considered with great caution. There is a correlation between the level of transcription and the number of mutations introduced in the *GA*-element; pKRS413-luc-4T and pKRS413-luc-4A with both MAZ-binding sites abrogated show a lower level of residual transcription (~40%) than pKRS413-luc-2T having only one MAZ-binding site abrogated (~60%). The fact that a residual transcription of ~40% is observed when the MAZ binding is abrogated suggests that *KRAS* transcription also depends on other nuclear proteins.

To confirm that recombinant MAZ binds to the *KRAS* promoter in a sequence-specific manner, we performed a DNase I footprinting assay on an 80-mer promoter fragment spanning over the *GA*-element (Fig. 4*b*). MAZ was found to protect the DNA portion containing the cluster of G-runs, confirming that both MAZ sites are indeed occupied by the protein.

As MAZ binds to a promoter sequence that can extrude a *G*-quadruplex, we asked whether it also recognizes the folded conformation of the *GA*-element. Previous studies have shown that MAZ binds to the *G*-quadruplexes formed by the G-rich region of the diabetes susceptibility locus *IDDM2* (13) and the GGA repeat region in the *CMYB* promoter (9). Fig. 5*a* shows that quadruplex 28R efficiently competed away the DNA-protein complexes between MAZ and the *GA*-duplex; a 5-fold excess of cold quadruplex 28R reduces the MAZ-duplex complexes to ~50%, whereas a 50-fold excess completely abrogates the complexes. From these data we roughly estimated for the interaction of MAZ with the *G*-quadruplex a K_D of about 0.5×10^{-9} M. Additionally, when MAZ was incubated with radiolabeled quadruplex 28R, a clear DNA-protein complex was observed by EMSA. Instead, the mutant oligonucleotide with

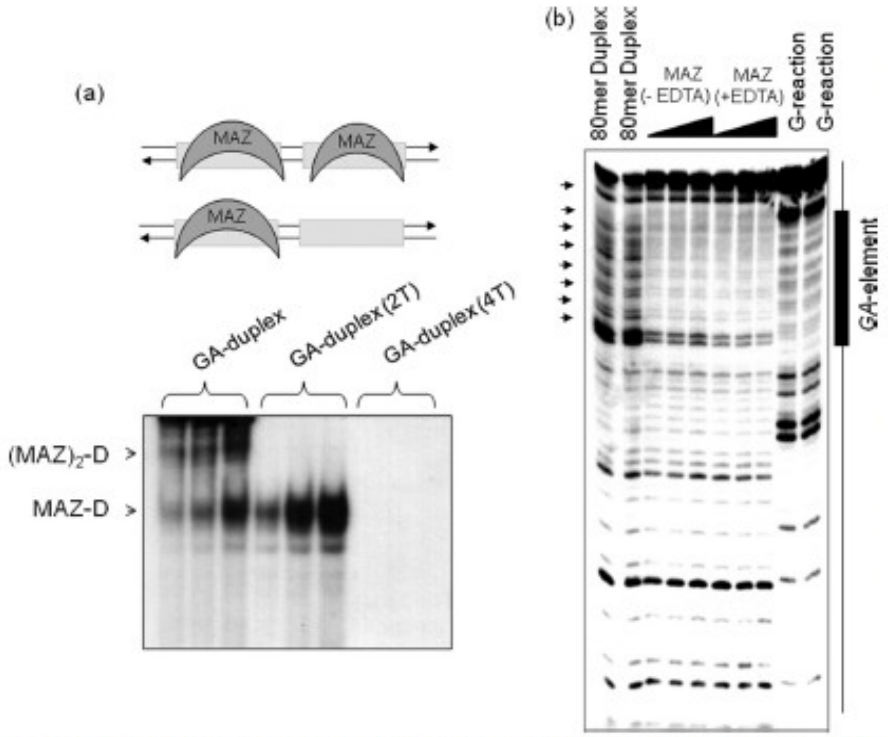


FIGURE 4. **Binding of MAZ to duplex of the *GA*-element.** *a*, an EMSA shows the formation of two DNA·MAZ complexes with a 1:1 and 1:2 stoichiometry. The targets used are the *GA*-duplex, the mutant duplexes with 2 and 4 G → T mutations (GA-duplex (2T) and GA-duplex (4T) were obtained annealing 28R(2T) and 28R(4T) (Table 1) to their complementary oligonucleotides). MAZ amounts of 1, 2, and 4 μg were used, whereas the target ^{32}P duplex was 20 nM. The experiment was performed with 5% PAGE in TBE. *b*, DNase I footprinting of an 80-mer promoter fragment containing the *GA*-element is shown. From *left*, *first* and *second lanes*, duplex digested with DNase I; *third-fifth lanes* and *sixth-eighth lanes*, DNase I digestion in the absence and presence of EDTA in the presence of 0.5, 1 and 2 μM MAZ-GST.

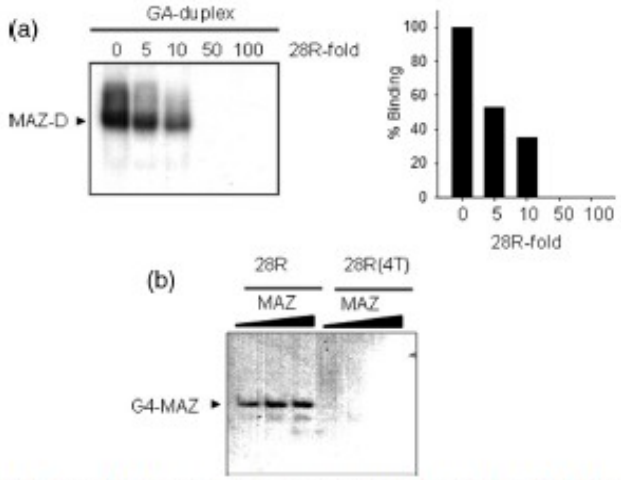


FIGURE 5. **EMSA competition experiments and binding of MAZ to G4-DNA.** *a*, a competition assay shows that the DNA-MAZ complex is competed away by 5-, 10-, 50-, and 100-fold of cold quadruplex 28R. Experiments performed with 20 nM GA-duplex and 3 μg of MAZ-GST. Histograms show the densitometric analysis of the gel. *b*, an EMSA shows the binding of quadruplex 28R to the MAZ protein. Radiolabeled quadruplex 28R and mutant 28R(4T) (20 nM) have been incubated with 2, 3, and 5 μg MAZ. 5% PAGE in TBE.

four G → T mutations, which is unable to fold into a *G*-quadruplex, did not form any DNA-protein complex (Fig. 5*b*).

MAZ Stabilizes the KRAS Quadruplex—The interaction between recombinant MAZ and the *KRAS* *G*-quadruplex has been investigated by FRET experiments, measuring the increase of FAM emission at 525 nm as a function of temperature (excitation, 485 nm). At *r* ([protein]/[DNA]) ratios of 0, 1, 2.5, and 5, MAZ does not change the emission spectrum of F-28R-T, suggesting that in the DNA-protein complex the *G*-quadruplex maintains its characteristic *p* value of 0.39 (Fig. 6*a*). To explore whether the binding to MAZ results in the stabilization of the *G*- quadruplex, we performed FRET-melting experiments. In 50 mM KCl, the folded DNA melts cooperatively, with a T_M of 70 °C (Fig. 6*b*). When an unrelated protein such as BSA is added to the *G*-quadruplex solution ($r = 5$), the T_M does not change, as expected. In contrast, when increasing aliquots of MAZ are added ($r = 1$, 2.5, and 5), the T_M increases up to 85 °C, indicating that the protein strongly stabilizes the DNA structure. The experiment was repeated in 100 mM KCl, and in this case the T_M increased from 80 to 90 °C (Fig. 6*c*). To provide further evidence about the stabilizing effect of MAZ on the *G*-quadruplex, we used a polymerase stop assay. In the presence of 25 mM KCl, Taq polymerase was arrested at the 3′ end of *GA*-element, before the first and second runs of guanines. The blocking of the polymerase increases in the presence of MAZ; Q_2 increases by 80% and Q_1 by 25%, confirming that MAZ facilitates the folding (Fig. 6*e*). Moreover, we expected that the stabilizing activity of MAZ should slow down the assembly rate between quadruplex F-28R-T and its complementary 28Y strand. When quadruplex F-28R-T in 50 mM KCl (T_M = 70 °C) was mixed with 28Y at 37 °C, the *G*-quadruplex was transformed into the more stable duplex, and the fluorescence of FAM at 525 nm increased (Fig. 6*d*). The assembly process could be monitored by measuring the increase of fluorescence, ΔF, as a function of time ($\Delta F = F - F_0$, where F_0 is the FAM fluorescence at 525 nm at $t = 0$, and F is the fluorescence at time t). The ΔF *versus* t curves showed an exponential shape that was best-fitted to a double-exponential equation (22). For the slow phase, a kinetic constant k of $5.1 \times 10^{-3} \pm 7 \times 10^{-6}$ s^{-1} was obtained. But when the hybridization was performed in the presence of MAZ at $r = 1$, the constant dropped to $1.7 \times 10^{-3} \pm 4 \times 10^{-5}$ s^{-1}. When a higher amount of MAZ was used ($r = 2$), the hybridization was completely inhibited. This demonstrates that MAZ, although stabilizing

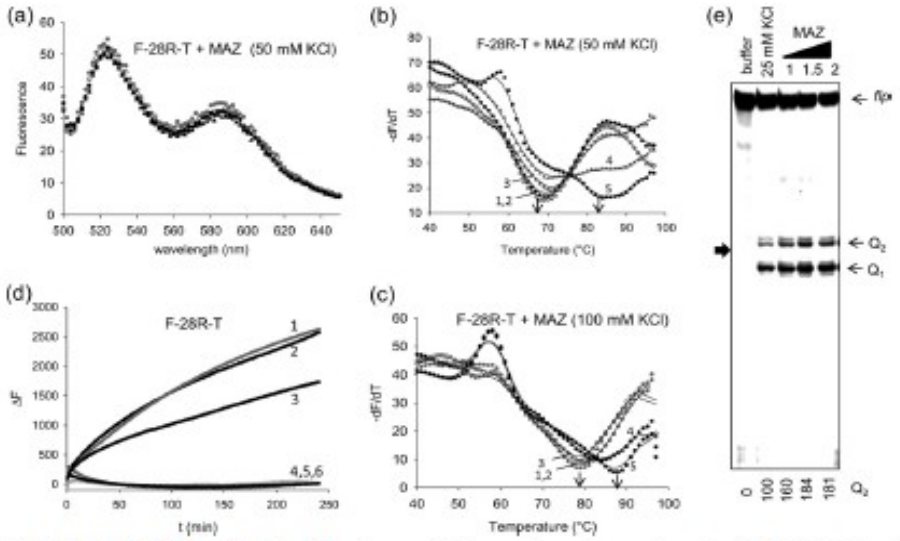


FIGURE 6. **MAZ stabilizes the *G*-quadruplex.** *a*, emission spectra are shown of quadruplex F-28R-T treated with BSA (r ([protein])/[quadruplex]) = 5) or increasing amounts of recombinant MAZ ($r = 0$, 1, 2. 5, 5). Excitation 475 nm. The experiment was carried out in 50 mM Tris-HCl, pH 7.4, 50 mM KCl, 50 μM zinc acetate; spectra were recorded using a fluorometer. *b*, shown is FRET melting of the same samples described in *a*. *Curves 1* and *2*, F-28R-T with and without BSA; *curves 3*, *4*, and *5*, F-28R-T with MAZ at $r = 1$, 2.5, and 5; the experiment was carried out in 50 mM Tris-HCl, pH 7.4, 50 mM KCl, 50 μM zinc acetate. *c*, the experiment was as in *b* but with 100 mM KCl. *d*, shown is the rate of assembly between quadruplex F-28R-T and 28Y. *Curve 1*, F-28R-T + 28Y (1:1); *curve 2*, F-28R-T + 28Y (1:1) + BSA ($r = 2$); *curve 3*, F-28R-T + 28Y (1:1) + MAZ ($r = 1$); *curve 4*, F-28R-T + 28Y (1:1) + MAZ ($r = 2$); *curve 5*, F-28R-T + BSA ($r = 2$); *curve 6*, F-28R-T MAZ ($r = 2$). The experiment was carried out in 50 mM Tris-HCl, pH 7.4, 50 mM KCl, 50 μM zinc acetate at 37 °C. The data shown in *panels b–d* have been collected by a real time apparatus (CFX96 Bio-Rad) measuring the FAM emission at 525 nm. *e*, shown is a polymerase-stop assay using the wtR-Mur80 template, 18-mer primer pMur80, and increasing amounts of MAZ (1, 1.5, and 2 μg) as indicated in the figure. The experiment was carried out in 50 mM Tris-HCl, pH 7.4, 25 mM KCl, 50 μM zinc acetate, 12% polyacrylamide gel in 1× TBE. *flp*, full-length product. *Numbers at the bottom* indicate the intensity of the Q_2 band.

the *G*-quadruplex, hinders the quadruplex-to-duplex transformation in the *GA*-element.

Selective Enrichment Strategy and ChIP Assay—To see if the *GA*-element is bound by endogenous MAZ present in NIH 3T3 nuclear extract, we used a selective enrichment strategy (Fig. 7*a*). Biotinylated oligonucleotides (Table 1) mimicking the *KRAS GA*-element (G4-biotin annealed to its complementary) or the *KRAS* quadruplex (folded G4-biotin) (Table 1) were used as bait in pulldown biotin-streptavidin assays. The pulled-down samples were separated by SDS-PAGE and analyzed by Western blot with specific antibodies. As shown in Fig. 7*a*, the *GA*-element enriched the pulled-down sample with MAZ, whereas a scramble duplex (scr-biotin annealed to its complementary) did not, indicating that this transcription factor binds to the *KRAS GA*-element with good affinity and specificity. In keeping with EMSA (Fig. 5), endogenous MAZ was efficiently pulled down by G4-biotin in the quadruplex conformation. Instead, the scramble oligonucleotide (scr-biotin) showed a much lower MAZ affinity than quadruplex G4-biotin.

In addition to MAZ, we focused our efforts on PARP-1, having found in a previous study that this protein binds to the human *KRAS* promoter at a sequence homologous to the *GA*-element (8) (supplemental Fig. S3). We found that PARP-1 interacts non-specifically with the *GA*-element, as it was pulled down by the *GA*-element and also by a scramble duplex (Fig. 7*a*). This is probably due to the fact that the protein recognizes the ends of the duplex (23). By contrast, PARP-1 showed binding specificity for quadruplex G4-biotin, as the unstructured oligonucleotide scr-biotin is weakly bound. The binding of PARP-1 to the *G*-quadruplex was confirmed by EMSA (supplemental Fig. S4).

To investigate whether MAZ and PARP-1 bind to the *GA*-element under *in vivo* conditions, we performed ChIP experiments. ChIP is a powerful tool for studying protein-DNA interactions under real physiological conditions. Living NIH 3T3 cells were treated with formaldehyde to cross-link and fix the protein-DNA complexes at the chromatin level. The chromatin was sheared into short fragments, and the DNA·MAZ and DNA·PARP-1 complexes were immunoprecipitated with specific anti-MAZ and anti-PARP-1 antibodies. The DNA present in the immunoprecipitated complexes was recovered and amplified by PCR. Fig. 7*b* shows that the amplification of sheared chromatin produced the expected band from the *KRAS* promoter region. It can be seen that although anti-RNA polymerase II and anti-IgG antibodies did not immunoprecipitate any DNA-protein complex involving the *KRAS* DNA (negative control), anti-MAZ and anti-PARP-1 antibodies effectively immunoprecipitated the MAZ·DNA and PARP-1·DNA complexes. The finding that antibody anti-MAZ was more efficient than antibody anti-PARP-1 may be explained by a different binding strength of the two antibodies or by the fact that the *GA*-element is prevalently occupied by MAZ. Together, the data demonstrate that under physiological conditions MAZ and, to a lower extent, PARP-1, is associated to the *GA*-element of murine *KRAS* promoter.

Role of MAZ and PARP-1 on KRAS Transcription—To investigate the role of MAZ in the activation of the murine *KRAS* promoter, NIH 3T3 cells were co-transfected with pKRS413-luc and pCMV-MAZ, a plasmid encoding for MAZ. The overexpression of MAZ increased luciferase by ~35%, suggesting that it activates the *KRAS* promoter (Fig. 8*a*). The fact that MAZ did not strongly stimulate the *KRAS* promoter may be due to a relatively high level of endogenous MAZ in NIH 3T3.

To further evaluate the impact of MAZ on *KRAS* transcription, we carried out shRNA knockdown experiments. NIH 3T3 cells were treated with a validated anti-MAZ shRNA, and the level of both MAZ and *KRAS* transcripts was measured by real-time PCR 48 h after shRNA treatment. As a control, we transfected the cells with an unrelated shRNA. An inhibition of *KRAS* transcription (~40%) was observed upon MAZ silencing even if this was only partial (Fig. 8*b*). This is in keeping with the dual luciferase assays showing that *KRAS* transcription is

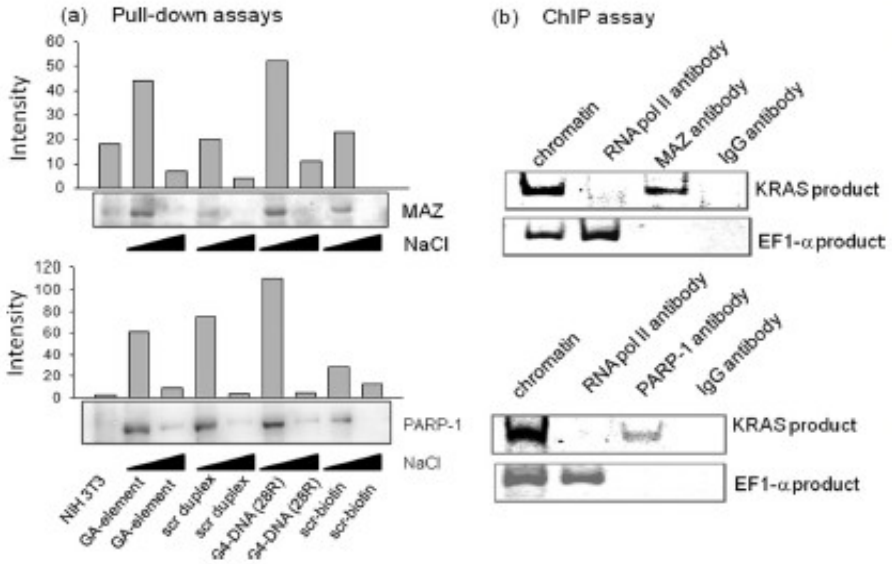


FIGURE 7. **Pulldown and ChIP assays.** *a*, biotinylated *GA*-element in duplex (G4-biotin hybridized to its complementary) or in quadruplex (G4-biotin) were used as bait in pulldown experiments with NIH 3T3 extract. The concentrations of NaCl in the elution buffer used to elute the protein fractions were 0.5 and 1 M. The panels show the Western blots of the pulled down fractions obtained with anti MAZ (*top panel*) and anti-PARP-1 (*bottom panel*) antibodies. The band intensities have been measured with ChemiDOC XRS apparatus (Bio-Rad). *b*, a chromatin immunoprecipitation assay was performed with anti-MAZ, anti-PARP-1, anti- RNA polymerase II (positive control), and IgG (negative control) antibodies. PCR analysis was performed on DNA isolated from ChIP reactions using controls, anti-PARP-1, and anti MAZ antibodies. PCR was performed with KRAS primers (see "Experimental Procedures") and EF1-α control primers (EF1-α primers provided by the kit amplify a 233-bp fragment from the DNA immunoprecipitated with anti-RNA polymerase II, used as a positive control). The *KRAS* PCR amplification product obtained with anti MAZ and anti-PARP-1antibodies show that under *in vivo* conditions the *GA*-element is bound by PARP-1 and MAZ.

down-regulated by 40% when the MAZ-binding sites are abrogated. From these data we can conclude that MAZ is an essential factor for activating *KRAS* transcription.

Similarly, the functional role of PARP-1 at the *KRAS* promoter was determined by using a validated shRNA. Fig. 8, *c* and *d*, shows that the level of *KRAS* transcripts parallels the inhibition of PARP-1 at both 48 and 72 h post-transfection, showing a dependence of *KRAS* transcription from PARP-1. Considering that PARP-1 is a protein that catalyzes the poly(ADP-ribosyl)ation of target proteins (heteromodification) and itself (automodification) (24), we tested whether this activity is important for *KRAS* transcription by treating NIH 3T3 cells for 4 and 8 h with 10 and 30 μM DPQ, an inhibitor of PARP-1 activity (25). Fig. 8*e*, which reports the levels of *KRAS* transcript determined by real-time PCR in NIH 3T3 cells treated with DPQ for 4 and 8 h, shows that the inhibitor suppresses the transcription to ~30% of the control. This suggests that PARP-1 acts on *KRAS* in a complex way and at two distinct levels; the first through its poly(ADP-ribosyl)ation activity and the second by binding directly to the promoter like a transcription factor (24, 26) (see "Discussion").

Effect of G4-DNA Ligands on KRAS Transcription—An approach adopted by several authors to investigate the possible role played by G4-DNA in transcription regulation is to treat the cells with G4-DNA ligands and then measure the expression of the target gene. This approach assumes that the ligands have a higher affinity for G4-DNA than for duplex DNA and that they do not compete with the binding of proteins to the duplex or quadruplex forms of the promoter. The cationic porphyrin TMPyP4 has been widely used as a G4-DNA ligand, whereas its positional isomer TMPyP2, which does not bind to G4-DNA, has been normally used as a control. However, TMPyP4 is characterized by poor binding specificity and similar affinities for quadruplex and duplex DNAs (27, 28). Fig. 9, *a* and *c*, shows that TMPyP4 competes with the binding of MAZ to the *GA*-duplex (at 5 μM the porphyrin practically abrogates the DNA-protein complex) but not to the G-quadruplex (TMPyP4 and, to a lower extent, DIGP/Zn-DIGP, seem to enhance the binding of MAZ to the G-quadruplex). A similar result was obtained by using a NIH 3T3 nuclear extract (supplemental Fig. S5). In light of these findings we used guanidine-modified phthalocyanines (GPc), DIGP and Zn-DIGP, a new class of G4-DNA ligands that specifically binds to G4-DNA and poorly competes with the binding of proteins to duplex DNA (21, 29). Indeed, Fig. 9*b* shows that 10 μM DIGP and Zn-DIGP weakly compete with the binding of MAZ to the *GA*-duplex. In a recent study we found that DIGP and Zn-DIGP interact with the murine *KRAS* quadruplex with a K_D between 10^{-7} and 10^{-6} M (21). Polymerase stop assays showed that these ligands efficiently stabilize the *G*-quadruplex formed by the *GA*-element (21). Indeed, circular dichroism spectra obtained at various temperatures (20–90 °C) showed that the T_M of the *G*-quadruplex in the presence of DIGP ($r = 4$) increased from 68 to >90 °C in 50 mM KCl (supplemental Fig. S6). Furthermore, the assembly between F-28R-T and 28Y was practically inhibited in the presence of DIGP at $r = 5$ (supplemental Fig. S7). Because MAZ binds to the *KRAS* quadruplex, we analyzed if this also occurs in the presence of phthalocyanines and found that 10 μM DIGP and Zn-DIGP promoted only a weak competition of the MAZ-quadruplex complex (Fig. 9*d*). To test the effect of the guanidine phthalocyanines on *KRAS* transcription, we performed dual luciferase assays. We used ligand concentrations within the window where the molecules showed a relatively low cytotoxicity, *i.e.* at concentrations lower than IC_{50}. NIH 3T3 cells were first treated with 10 μM GPc for 16 h, then transfected with a mixture of pKRS413-luc and pRL-CMV. They were grown for a further 48 h before firefly and *Renilla* luciferase were measured with a luminometer. As a control, we treated the cells with porphyrins or phthalocyanines that do not have affinity for G4-DNA (Zn-SucPc, and pentaphyrin; see supplemental Fig. S8) (21). The results show that DIGP and Zn-DIGP stimulate *KRAS* transcription by more than 2-fold, suggesting that

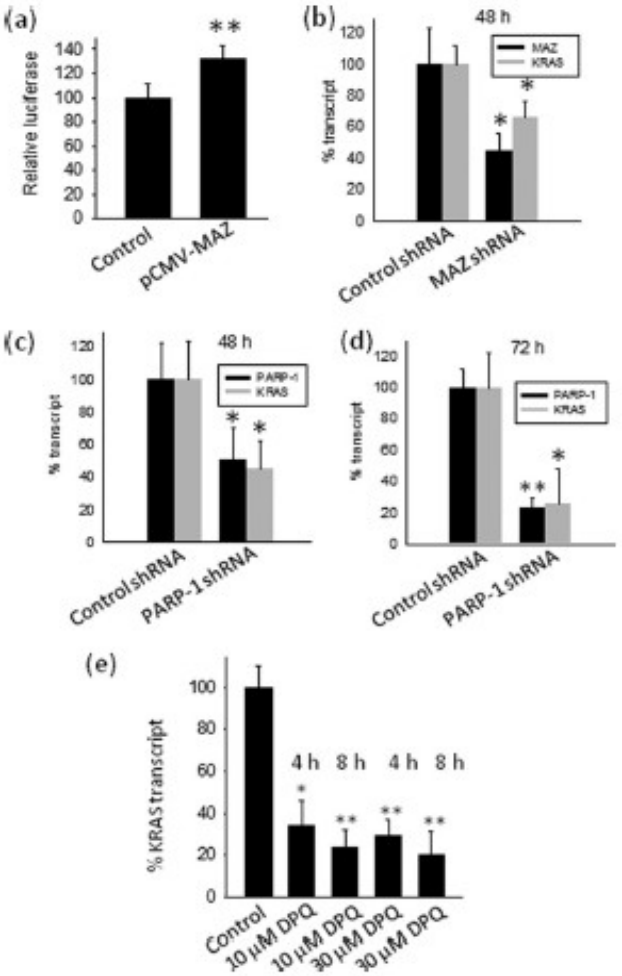


FIGURE 8. **Effect of MAZ and PARP-1 on transcription.** *a*, transient transfection experiments show the effect of MAZ overexpression on firefly luciferase driven by the wild-type *KRAS* promoter. *Left bar* (control), cells transfected with pKRS413-luc, pcDNA3 (empty vector), and pRL-CMV; *right bar*, cells transfected with pKRS413-luc, pCMV-MAZ, and pRL-CMV. The *ordinate* reports firefly luciferase normalized by *Renilla* luciferase; *b–d*, real-time PCR shows the effect on *KRAS* transcription of knocking down MAZ or PARP-1 with specific shRNAs at 48 and 72 h. The *ordinate* reports the % transcript (*KRAS* or MAZ or PARP-1), *i.e.* $R_T/R_C \times 100$, where R_T is (transcript)/(average transcripts from β2-microglobulin, glyceraldehyde-3-phosphate dehydrogenase, and hypoxanthine-guanine phosphoribosyltransferase) in shRNA-treated cells, and R_C is (transcript)/(average transcripts from β2-microglobulin, glyceraldehyde-3-phosphate dehydrogenase, and hypoxanthine-guanine phosphoribosyltransferase) in cells treated with unrelated shRNA. *e*, real time PCR shows the effect on *KRAS* transcription of 10 and 30 μM DPQ, an inhibitor of PARP-1 poly (ADP-ribose) activity. The *ordinate* reports the % *KRAS* transcript, *i.e.* $R_T/R_C \times 100$, where R_T is (*KRAS* transcript)/(β2microglobulin and hypoxanthine-guanine phosphoribosyltransferase transcripts) in untreated cells, and R_C is (*KRAS* transcript)/(β2-microglobulin and hypoxanthine-guanine phosphoribosyltransferase transcripts) in cells treated with DPQ. Differences from the control are supported by Student's *t* test, $p < 0.05$ (*one asterisk*), $p < 0.01$ (*two asterisks*).

the *G*-quadruplex behaves as a transcription activator element (Fig. 10). In contrast, the control molecules Zn-SucPc, DIGPor (tetrakis-(diisopropylguanidine) porphine), and pentaphyrin, which do not stabilize G4-DNA, have no effect on transcription. This result strengthens our hypothesis that quadruplex DNA in the *GA*-element is involved in the activation of transcription. Taken together, our data indicate that G4-DNA mediates the activation of *KRAS* transcription probably by recruiting MAZ to the *KRAS* promoter.

DISCUSSION

The *GA*-element of the murine *KRAS* gene promoter between −322 to −288, which contains six runs of guanines separated by adenine nucleotides, is essential for transcription; its deletion reduces transcription to 5.7% (18). This critical *GA*-element is structurally polymorphic, as in the presence of potassium and under physiological conditions it folds into one of two topologically distinct *G*-quadruplexes. The stability of the *G*-quadruplex is sufficiently high to arrest Taq polymerase even at 50 °C. We report here for the first time that the murine *KRAS GA*-element interacts with zinc finger proteins MAZ and PARP-1. These interactions are complex as both proteins recognize the duplex and/or folded conformations of the *GA*-element. This has been discovered by employing two independent techniques: a selective enrichment strategy (pull-down assay with biotinylated oligonucleotides) and ChIP assays.

MAZ was predicted to bind to the *GA*-element by using the MAT-Inspector bioinformatic software, which indicated the presence in the *GA*-element of two MAZ consensus sequences, GGGAGGG. The binding of this nuclear factor to the *GA*-element was analyzed by recombinant MAZ-GST. As the *GA*-element contains two sites for MAZ, it forms two DNA-protein complexes, one in which only one site is occupied by MAZ, the other in which both sites are occupied. The interaction of MAZ to the *GA*-element is highly sequence-specific, as it is sufficient to introduce in one binding site two point mutations to completely abrogate the interaction.

MAZ was first identified as a GA-box binding transcription factor in the *CMYC* promoter controlling transcription initiation and termination (30, 31). This nuclear protein is expressed in many tissues (32) and its functional role in the context of transcription seems to be complex, as some genes are activated by MAZ, whereas others are repressed. MAZ was reported to activate the serotonin 1a receptor gene promoter (33), the adenovirus major later promoter (34), the parathyroid hormone PTH/PTHrP promoter (35), the PNMT promoter (36), the insulin promoter (13) and the muscle creatine kinase promoter (37). Instead, MAZ was reported to repress the endothelial nitric-oxide synthase promoter (38) and the *CMYB* promoter (9). Here, we provide sound evidence that MAZ activates transcription of the murine *KRAS* gene. When MAZ was partially knocked-down with a specific shRNA, transcription dropped to ~60% of control. In contrast, when MAZ was overexpressed, transcription increased by 35%. These values are low probably because (i) endogenous MAZ is not limiting in NIH 3T3 cells; (ii) MAZ may require post-translational modifications, for instance phosphorylation, to become active and this cannot occur in all overexpressed MAZ molecules (39); (iii) MAZ may need to interact with other proteins to activate *KRAS* transcription.

An important finding of this study is that MAZ stabilizes the quadruplex structure formed by the *GA*-element. By means of FRET experiments we have discovered that in 50 mM KCl, the T_M of the *KRAS* quadruplex increased from 70 to 85 °C, in the presence of MAZ at $r = 5$. The ability of MAZ to recognize non

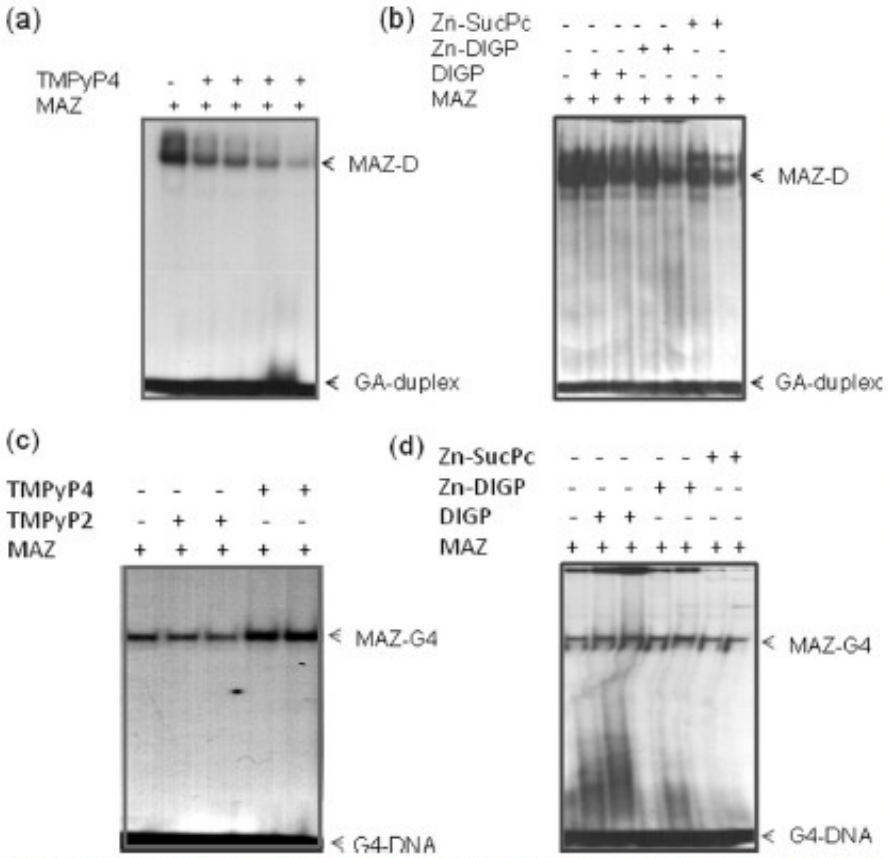


FIGURE 9. **Effect of G4-DNA ligands on *KRAS* transcription.** *a*, an EMSA shows that the complex between *GA*-duplex and MAZ-GST is competed by increasing amounts of TMPyP4 (0.5, 1, 2, and 5 μM). Duplex concentration is 20 nM, and MAZ is 4 μg (5% PAGE in TBE). *b*, shown is the same experiment as in *a* but with phthalocyanines DIGP, Zn-DIGP, and Zn-SucPc at concentrations 5 and 10 μM. *c*, an EMSA show that the *KRAS G*-quadruplex binds to MAZ-GST in the absence and presence of TMPyP2 or TMPyP4 (2.5 and 10 μM), 10% PAGE in TBE. *d*, shown is the same experiment as in *c* but phthalocyanines DIGP, Zn-DIGP, and Zn-SucPc at the concentrations of 5 and 10 μM. The streaking in the *second*, *third*, and *fifth lanes* from the *left* is due to the effect of phthalocyanines binding to the quadruplex.

B-DNA structures was previously reported. Indeed, MAZ binds to the *G*-quadruplex structures formed by the IDDM2 locus (13) and the GGA repeat sequence in the *CMYB* promoter (9). Because MAZ shows binding activities to both duplex and quadruplex DNA and is involved in the trans-activation of several genes, the activation of *KRAS* is likely to involve other cofactors. Several proteins have been reported to interact with MAZ. For instance, autoregulation of MAZ occurs through its interaction with histone deacetylase (40), whereas FAC1 has been shown to repress MAZ-mediated activation of the SV40 promoter and to co-localize with MAZ in plaque-like structures in the brains of Alzheimer's patients (41). Furthermore, DCC, a type I membrane protein and putative tumor suppressor, has been shown to associate with MAZ during neuronal differentiation of P19 cells, which correlates with the loss of *CMYC* expression (42).

The other zinc finger protein that was found to influence the *KRAS* transcription is PARP-1. In addition to its capacity of direct binding to DNA, it shows catalytic activity by attaching poly(ADP-ribose) units to target proteins including itself (22). The addition of negative charges on the modified proteins is expected to affect their interactions with DNA and other proteins. The intracellular level of PARP-1 and the removal of the ribose chains on the target proteins by specific glycohydrolases are tightly controlled (43). Data accumulated in the years suggest that the role of PARP-1 can be divided into emergency and housekeeping functions. The emergency role occurs after DNA damage, whereas the housekeeping role is the modulation of gene expression (44). Genome-wide localization studies (Chip-on-chip) have revealed that PARP-1 is associated with upstream of transcription start sites of actively transcribed genes (45). A significant number of genes (about 3.5% of the transcriptome) involve PARP-1 in transcription regulation (46). It occurs by heteromodification of histones, which promotes the decondensation of high order chromatin, and by binding to enhancer/promoter regulatory *cis*-elements (44). Previous studies have reported that PARP-1 recognizes unusual DNA conformations such as hairpins/cruciforms (47–49) and the *G*-quadruplex formed by a *cis*-element in the human *KRAS* promoter (8). Interestingly, Ladame and co-workers (50) have shown that PARP-1 undergoes automodification after binding to the *c-kit* quadruplex. For a review on the role of PARP-1 in transcription, we refer to Kraus and Lis (44). Many studies have shown that PARP-1 is implicated in the activation or repression of transcription (44). Our data show that PARP-1 acts as an activator because *KRAS* transcription decreases to 20% that of control when PARP-1 is knocked down by specific shRNA. Furthermore, as PARP-1 is an enzyme with poly(ADP-ribosyl)ation activity, we addressed the question of whether the inhibition of its catalytic activity affected *KRAS* transcription. We found that DPQ, a PARP-1 inhibitor, strongly down-regulates *KRAS* transcription. Combining these data, one can conclude that the action of PARP-1 on the *KRAS* promoter requires that the protein is catalytically active on other proteins and probably itself.

One approach to investigate the role of G4-DNA on transcription is to introduce point mutations in the quadruplex-forming sequence of the promoter and study the effect on transcription (6, 9, 51). It is assumed that the point mutations do not alter the binding of the proteins to DNA. Because we found that the insertion in the *GA*-element of two-point mutations com-

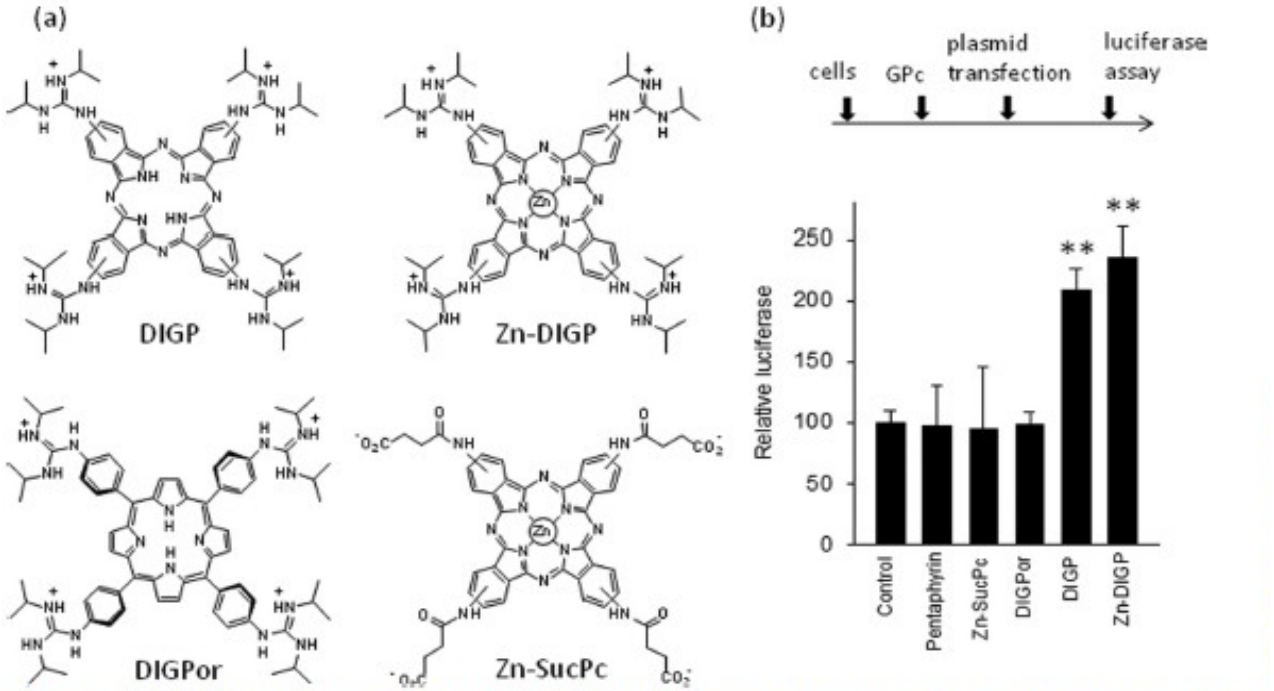


FIGURE 10. **Transcription assays.** *a*, structures of guanidine phthalocyanines used are shown. *b*, a dual luciferase assay was performed in NIH 3T3 cells treated with 10 μM DIGP, Zn-DIGP, Zn-SucPc, DIGPor, or pentaphyrin for 16 h and subsequently transfected with a mixture of pKRS413-luc and pRL-CMV. The signal was normalized to *Renilla* luciferase (control). Luciferase expressions in the presence of DIGP and Zn-DIGP are different from control, by Student's *t* test, $p < 0.01$ (*two asterisks*).

pletely abrogated the MAZ binding, we could not draw any conclusion about the role of G4-DNA on transcription by promoter mutational analysis. We, therefore, adopted another strategy to ascertain the role of G4-DNA within the *GA*-element; we used ligands that specifically interact with G4-DNA. If the ligands stabilize the quadruplex and shift the duplex-quadruplex equilibrium, their effects on transcription can give information about the functional role of G4-DNA. However, this rationale is valid if the ligands do not compete with the proteins binding to the duplex or quadruplex conformations of the *GA*-element. We employed for our functional studies a class of G4-DNA ligands showing a higher specificity for G4-DNA than TMPyP4. TMPyP4 is a well known cationic porphyrin that has been extensively used as G4-DNA ligand in transcription assays, although it shows a poor binding specificity for quadruplex DNA (27, 28). We found that the complex between MAZ and *GA*-duplex is strongly competed by TMPyP4 but not by GPcs. Furthermore, GPcs and also TMPyP4 did not compete the complex between MAZ and quadruplex 28R. By means of luciferase assays we found that GPcs enhanced the activity of the *KRAS* promoter by more than 2-fold, whereas the control ligands did not. This result is apparently in conflict with a previous study in which we showed that TMPyP4 down-regulated the promoter activity of murine *KRAS* (7). In that study we cotransfected TMPyP4 and pKRS413-CAT, a plasmid-bearing CAT driven by the murine *KRAS* promoter. Most likely, in the transfection mixture TMPyP4 saturated the *GA*-element, so that the observed down-regulation of transcription was due to the inability of MAZ (and other proteins) to bind to the promoter. Thus, our previous hypothesis that G4-DNA might be a repressor is not supported by the present study. On the contrary, the data here reported provide strong evidence that G4-DNA should be an activator element. Similarly to our study, Lew *et al.*, (13) reported that the insulin-linked polymorphic region located in the promoter of the insulin gene forms intra- and inter-molecular G-quartets that bind to Pur-1 (MAZ) and stimulate transcription activity.

In light of these new discoveries, how can MAZ, PARP-1 and G4-DNA regulate transcription of *KRAS*? A possible model that rationalizes the results of this study is the following. PARP-1 is likely to act primarily at chromatin level by promoting, through heteromodification of histones, a local decondensation of chromatin that increases the DNA accessibility (44). This is in keeping with the fact that when heteromodification is inhibited by DPQ, *KRAS* transcription is repressed. Under negative supercoiling, the tract of promoter containing the *GA*-element unwinds (this is a nuclease hypersensitive site (18)), and the unpaired G-rich strand is expected to fold into a *G*-quadruplex (52, 53). As MAZ stabilizes the *G*-quadruplex, the protein should shift the duplex-folded equilibrium toward the folded form. This may also be favored by the binding to the complementary polypyrimidine strand of proteins recognizing C-rich sequences such as heterogeneous nuclear ribonucleoprotein K (54, 55) and poly(C)-binding proteins PCBP1–4 (56). The binding of MAZ to the promoter results in the activation of transcription, as indicated by the fact that the overexpression of MAZ in NIH 3T3 cells up-regulates transcription, whereas the depletion of MAZ by shRNA down-regulates transcription.

Our data indicate that the murine *KRAS* transcription is favored by the folded form of the *GA*-element as (i) *KRAS* transcription is reduced to 50% of control when point mutations abrogating DNA folding are introduced in the *GA*-element, (ii) transcription is 2-fold up-regulated by guanidine phthalocyanines that stabilize the murine *KRAS* quadruplex (21), (iii) MAZ stabilizes the *G*-quadruplex and activates transcription, and (iv) hydridization of the *G*-quadruplex to its complementary strand is inhibited by MAZ. It is likely that the unwinding of the *GA*-element serves to recruit at the promoter the proteins forming a multiprotein complex that activates transcription. Therefore, the *G*-quadruplex should function as an attraction point to recruit MAZ and the other proteins of the transcription machinery.

In summary, in this study we show that the critical *GA*-element of murine *KRAS* interacts with MAZ and PARP-1 in a very complex way involving DNA conformation changes from duplex to quadruplex DNA. We provide compelling evidence that both MAZ and PARP-1 are activators of the *KRAS* promoter. Both proteins recognize the parallel quadruplex conformation adopted by the *GA*-element, which probably has the function of recruiting these proteins to the promoter. Our study provides the basis for the rationale design of anticancer drugs, for example, G-rich G4-aptamers specific for MAZ and PARP-1 to down-regulate the expression of oncogenic *KRAS* in cancer cells (57, 58).

Acknowledgments—Plasmids pGEX-hMAZ and pCMV-MAZ were provided by RIKEN BioResource Center, which participates in the National Bio-Resources Project of the Ministry of Education, Culture, Sports, Science, and Technology of Japan. We thank Nathan Luedtke (Zurich University) for the gift of the guanidine-modified phthalocyanines.

REFERENCES

1. Sen, D., and Gilbert, W. (1988) *Nature* **334**, 364–366
2. Matzels, N. (2006) *Nat. Struct. Mol. Biol.* **13**, 1055–1059
3. Huppert, J. L., and Balasubramanian, S. (2007) *Nucleic Acids Res.* **35**, 406–413
4. Todd, A. K., Johnston, M., and Neidle, S. (2005) *Nucleic Acids Res.* **33**, 2901–2907
5. Eddy, J., and Maizels, N. (2006) *Nucleic Acids Res.* **34**, 3887–3896
6. Siddiqui-Jain, A., Grand, C. L., Bearss, D. J., and Hurley, L. H. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11593–11598
7. Cogoi, S., and Xodo, L. E. (2006) *Nucleic Acids Res.* **34**, 2536–2549
8. Cogoi, S., Paramasivam, M., Spolaore, B., and Xodo, L. E. (2008) *Nucleic Acids Res.* **36**, 3765–3780
9. Palumbo, S. L., Memmott, R. M., Uribe, D. J., Krotova-Khan, Y., Hurley, L. H., and Ebbinghaus, S. W. (2008) *Nucleic Acids Res.* **36**, 1755–1769
10. Sun, D., Liu, W. J., Guo, K., Rusche, J. J., Ebbinghaus, S., Gokhale, V., and Hurley, L. H. (2008) *Mol. Cancer Ther.* **7**, 880–889
11. Qin, Y., Rezler, E. M., Gokhale, V., Sun, D., and Hurley, L. H. (2007) *Nucleic Acids Res.* **35**, 7698–7713
12. Rankin, S., Reszka, A. P., Huppert, J., Zloh, M., Parkinson, G. N., Todd, A. K., Ladame, S., Balasubramanian, S., and Neidle, S. (2005) *J. Am Chem. Soc.* **127**, 10584–10589
13. Lew, A., Rutter, W. J., and Kennedy, G. C. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 12508–12512
14. Brooks, T. A., and Hurley, L. H. (2009) *Nat. Rev. Cancer* **9**, 849–861
15. Hershman, S. G., Chen, Q., Lee, J. Y., Kozak, M. L., Yue, P., Wang, L. S., and Johnson, F. B. (2008) *Nucleic Acids Res.* **36**, 144–156
16. Lee, S. K., Johnson, R. E., Yu, S. L., Prakash, L., and Prakash, S. (1999) *Science* **286**, 2339–2342
17. Cahoon, L. A., and Seifert, H. S. (2009) *Science* **325**, 764–767
18. Hoffman, E. K., Trusko, S. P., Murphy, M., and George, D. L. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 2705–2709
19. Tsutsui, H., Sakatsume, O., Itakura, K., and Yokoyama, K. K. (1996) *Biochem. Biophys. Res. Commun.* **226**, 801–809
20. Pestov, D. G., Dayn, A., Siyanova, E. Yu., George, D. L., and Mirkin, S. M. (1991) *Nucleic Acids Res.* **19**, 6527–6532
21. Membrino, A., Paramasivam, M., Cogoi, S., Alzeer, J., Luedtke, N. W., and Xodo, L. E. (2010) *Chem. Comm.* **46**, 625–627
22. Green, J. J., Ying, L., Klenerman, D., and Balasubramanian, S. (2003) *J. Am. Chem. Soc.* **125**, 3763–3767
23. de Murcia, G., and Ménissier de Murcia, J. (1994) *Trends Biochem. Sci.* **19**, 172–176
24. Caiafa, P., and Zlatanova, J. (2009) *J. Cell. Physiol.* **219**, 265–270
25. Martin-Oliva, D., Aguilar-Quesada, R., O'valle, F., Muñoz-Gámez, J. A., Martínez-Romero, R., Garcia Del Moral, R., Ruiz de Almodóvar, J. M., Villuendas, R., Piris, M. A., and Oliver, F. J. (2006) *Cancer Res.* **66**, 5744–5756
26. Wacker, D. A., Ruhl, D. D., Balagamwala, E. H., Hope, K. M., Zhang, T., and Kraus, W. L. (2007) *Mol. Cell. Biol.* **27**, 7475–7485
27. Ren, J., and Chaires, J. B. (1999) *Biochemistry* **38**, 16067–16075
28. Freyer, M. W., Buscaglia, R., Kaplan, K., Cashman, D., Hurley, L. H., and Lewis, E. A. (2007) *Biophys. J.* **92**, 2007–2015
29. Alzeer, J., Roth, P. J., and Luedtke, N. W. (2009) *Chem. Comm.* **15**, 1970–1971
30. Bossone, S. A., Asselin, C., Patel, A. J., and Marcu, K. B. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 7452–7466
31. Izzo, M. W., Strachan, G. D., Stubbs, M. C., and Hall, D. J. (1999) *J. Biol. Chem.* **274**, 19498–19506
32. Song, J., Murakami, H., Tsutsui, H., Tang, X., Matsumura, M., Itakura, K., Kanazawa, I., Sun, K., and Yokoyama, K. K. (1998) *J. Biol. Chem.* **273**, 20603–20614
33. Parks, C. L., and Shenk, T. (1996) *J. Biol. Chem.* **271**, 4417–4430
34. Parks, C. L., and Shenk, T. (1997) *J. Virol.* **71**, 9600–9607
35. Leroy, C., Manen, D., Rizzoli, R., Lombès, M., and Silve, C. (2004) *J. Mol. Endocrinol.* **32**, 99–113
36. Her, S., Claycomb, R., Tai, T. C., and Wong, D. L. (2003) *Mol. Pharmacol.* **64**, 1180–1188
37. Himeda, C. L., Ranish, J. A., and Hauschka, S. D. (2008) *Mol. Cell. Biol.* **28**, 6521–6535
38. Karantzoulis-Fegaras, F., Antoniou, H., Lai, S. L., Kulkarni, G., D'Abreo, C., Wong, G. K., Miller, T. L., Chan, Y., Atkins, J., Wang, Y., and Marsden, P. A. (1999) *J. Biol. Chem.* **274**, 3076–3093
39. Tsutsui, H., Geltinger, C., Murata, T., Itakura, K., Wada, T., Handa, H., and Yokoyama, K. K. (1999) *Biochem. Biophys. Res. Commun.* **262**, 198–205
40. Song, J., Ugai, H., Kanazawa, I., Sun, K., and Yokoyama, K. K. (2001) *J. Biol. Chem.* **276**, 19897–19904
41. Jordan-Sciutto, K. L., Dragich, J. M., Caltagarone, J., Hall, D. J., and Bowser, R. (2000) *Biochemistry* **39**, 3206–3215
42. Ugai, H., Li, H. O., Komatsu, M., Tsutsui, H., Song, J., Shiga, T., Fearon, E., Murata, T., and Yokoyama, K. K. (2001) *Biochem. Biophys. Res. Commun.* **286**, 1087–1097
43. Bonicalzi, M. E., Haince, J. F., Droit, A., and Poirier, G. G. (2005) *Cell. Mol. Life Sci.* **62**, 739–750
44. Kraus, W. L., and Lis, J. T. (2003) *Cell* **113**, 677–683
45. Krishnakumar, R., Gamble, M. J., Frizzell, K. M., Berrocal, J. G., Kininis, M., and Kraus, W. L. (2008) *Science* **319**, 819–821
46. Ogino, H., Nozaki, T., Gunji, A., Maeda, M., Suzuki, H., Ohta, T., Murakami, Y., Nakagama, H., Sugimura, T., and Masutani, M. (2007) *BMC Genomics* **8**, 41–56
47. Potaman, V. N., Shlyakhtenko, L. S., Oussatcheva, E. A., Lyubchenko, Y. L., and Soldatenkov, V. A. (2005) *J. Mol. Biol.* **348**, 609–615
48. Chasovskikh, S., Dimtchev, A., Smulson, M., and Dritschilo, A. (2005) *Cytometry A.* **68**, 21–27
49. Lonskaya, I., Potaman, V. N., Shlyakhtenko, L. S., Oussatcheva, E. A., Ly-

ubchenko, Y. L., and Soldatenkov, V. A. (2005) *J. Biol. Chem.* **280**, 17076–17083
50. Soldatenkov, V. A., Vetcher, A. A., Duka, T., and Ladame, S. (2008) *ACS Chem. Biol.* **3**, 214–219
51. Thakur, R. K., Kumar, P., Halder, K., Verma, A., Kar, A., Parent, J. L., Basundra, R., Kumar, A., and Chowdhury, S. (2009) *Nucleic Acids Res.* **37**, 172–183
52. Sun, D., and Hurley, L. H. (2009) *J. Med. Chem.* **52**, 2863–2874
53. Kouzine, F., Sanford, S., Elisha-Feil, Z., and Levens, D. (2008) *Nat. Struct. Mol. Biol.* **15**, 146–154
54. Michelotti, E. F., Michelotti, G. A., Aronsohn, A. I., and Levens, D. (1996) *Mol. Cell. Biol.* **16**, 2350–2360
55. Tomonaga, T., and Levens, D. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5830–5835
56. Chot, H. S., Hwang, C. K., Song, K. Y., Law, P. Y., Wet, L. N., and Loh, H. H. (2009) *Biochem. Biophys. Res. Commun.* **380**, 431–436
57. Chot, E. W., Nayak, L. V., and Bates, P. J. (2010) *Nucleic Acids Res.* **38**, 1623–1635
58. Teng, Y., Girvan, A. C., Casson, L. K., Pierce, W. M., Jr., Qian, M., Thomas, S. D., and Bates, P. J. (2007) *Cancer Res.* **67**, 10491–10500

Supporting Information

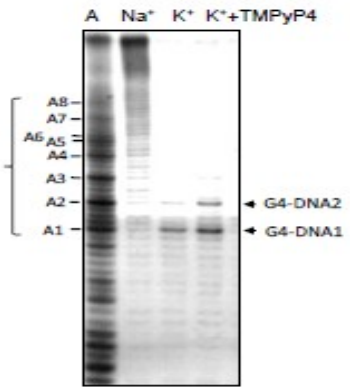
THE *KRAS* PROMOTER RESPONDS TO MYC-ASSOCIATED ZINC FINGER AND POLY[ADP-RIBOSE]POLYMERASE 1 PROTEINS WHICH RECOGNIZE A CRITICAL QUADRUPLEX-FORMING *G4*-ELEMENT*

Susanna Cogoi[1], Manikandan Paramasivam[1], Alexandro Membrino[1], Kazunari K. Yokoyama[2], and Luigi E. Xodo[1]*

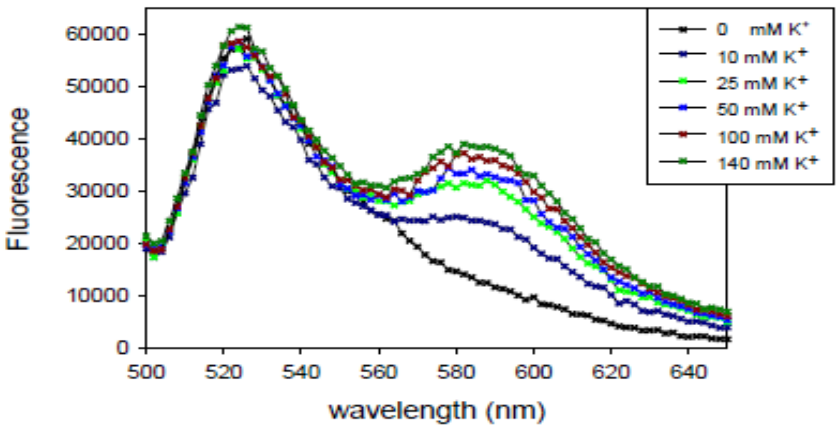
[1]Department of Biomedical Science and Technology, School of Medicine, P.le Kolbe 4, 33100 Udine, Italy; [2]Center of Excellence for Environmental Medicine, Kaohsiung Medical University, 100 Shih-Chuan 1st Road, 80708 Kaoshsiung, Taiwan.

Table S_1: Sequence of the primers used in this study

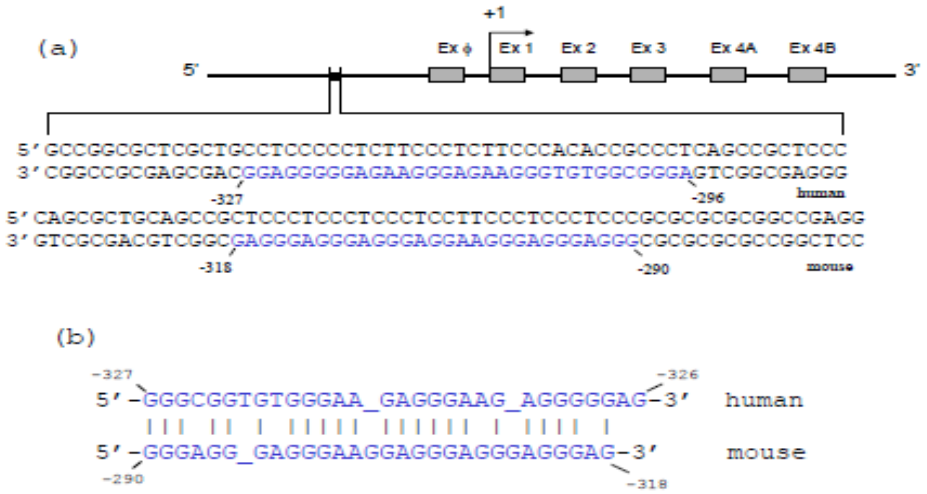
NAME	5'-3' sequence	Acccession n.	bases
mKRAS for	GCTCAGGAGTTAGCAAGGAG	BC 010202	570-89
mKRAS back	GTATTCACATAACTGTACACCTTG	BC 010202	730-753
mGAPDH for	GAAGCTTGTCATCAACGGGAAG	NM 008084	239-260
mGAPDH back	ACTCCACGACATACTCAGCAC	NM 008084	316-336
mPARP for	TGTTGAGGTGGATGGCTT	NM 007415	263-280
mPARP rev	GTCACCCAATGTCTTCTCC	NM 007415	377-395
mMAZ for	GGCTTATATTTCCGACCACA	NM_010772	1269-88
mMAZ rev	GGACAAACCTCACCAGTAC	NM_010772	1340-58
mβ2microglobulin for	GTCTCACTGACCGGCCTGTATG	NM_009735	91-112
mβ2microglobulin rev	CCCGTTCTTCAGCATTTGGATTTC	NM_009735	220-43
mHPRT for	GTGTTGGATACAGGCCAGACTTTG	NM_013556	599-622
mHPRT rev	ATCAACAGGACTCCTCGTATTTGC	NM_013556	765-88



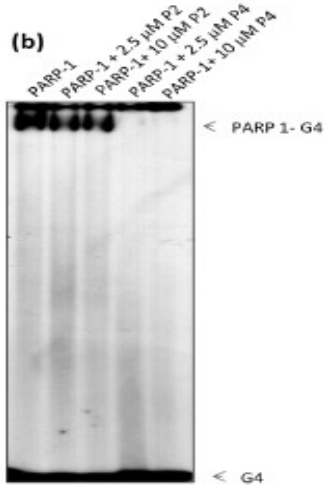
S_1: polymerase stop assay with 80mer wtR-Mur80 template containing the GA-element. Lane 1 shows the Sanger A-reaction of the template showing that Taq polymerase is arrested at the first two adenines at the 3' end of the GA-element.



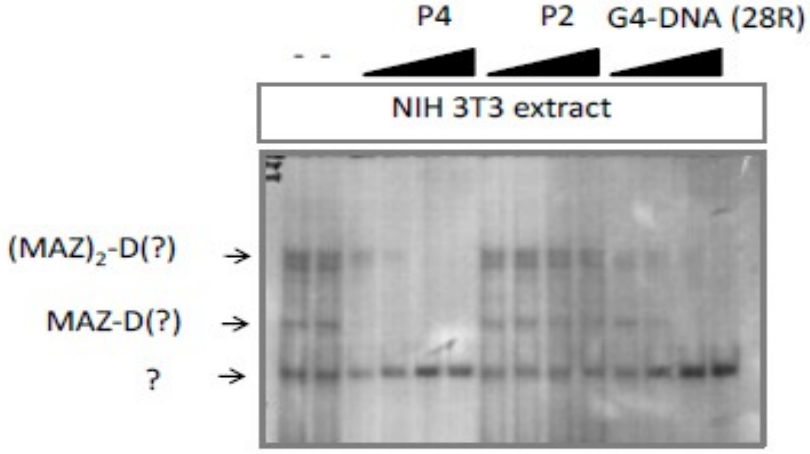
S_2: Fluorescence emitted by 200 nM F-28R-T in 50 mM Tris-HCl, pH 7.4 and increasing concentrations of KCl 0, 10, 25, 50, 100 and 140 mM KCl (Ex 475 nm).



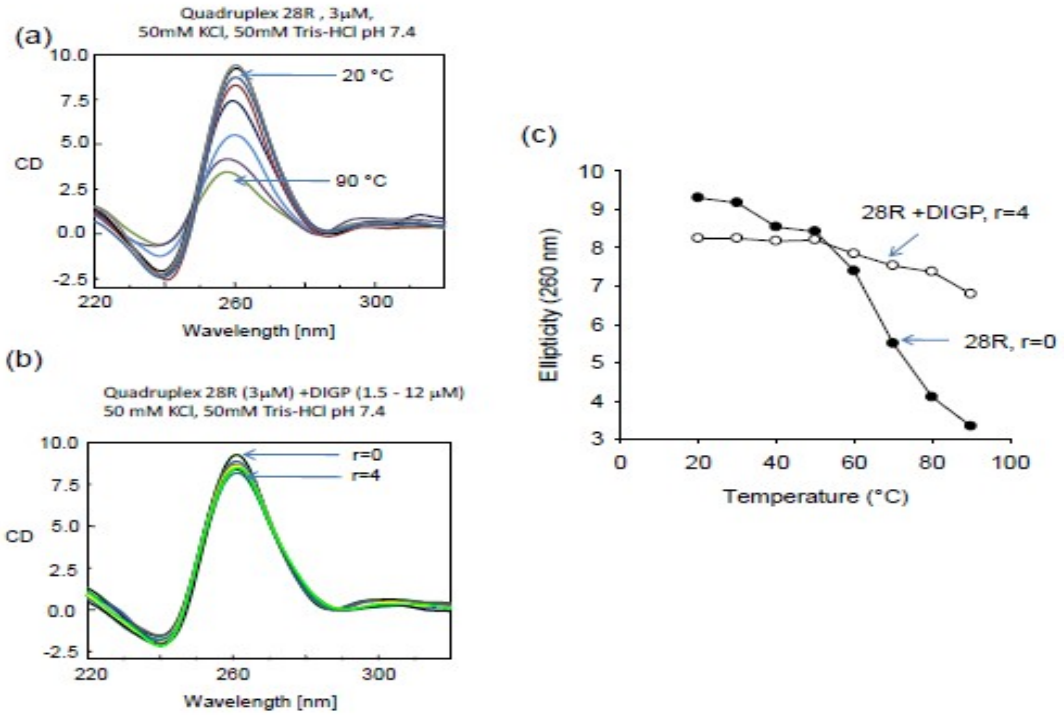
S_3: (a) Sequences of the *G4*-elements in the human and mouse *KRAS* promoters; (b) sequence homology between the mouse and human *G4*-element.



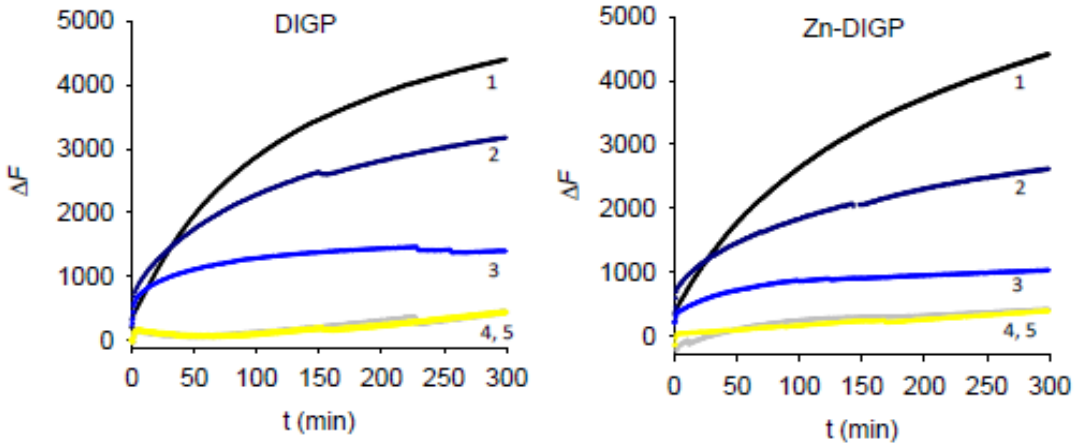
S_4: Mobility-shift experiment showing that recombinant PARP-1 (1 μg) forms a DNA-protein complex with 20 nM radiolabelled quadruplex 28R. PARP-1 and quadruplex DNA were incubated for 30 min in the presence of 30 mM KCl and 50 μM zinc acetate. EMSA performed in 10% polyacrylamide gel in TBE. The protein-DNA binding is not competed by TMPyP2, while it is by TMPyP4.



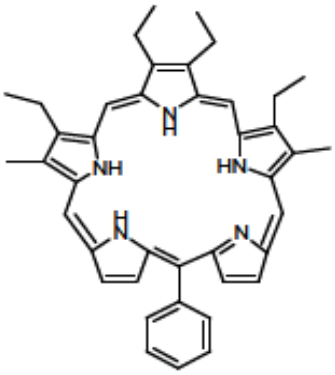
S_5: Porphyrin TMPyP4 (P4) competes the binding of nuclear proteins from NIH 3T3 extract to the *GA*-duplex. Each lane contains 25 nM radiolabelled *GA*-duplex, NIH 3T3 extract and competitor: lanes 1-2, 10 μg extract; lanes 3-6, 10 μg extract plus 5, 10, 25 and 50 μM P4; lanes 7-10, 10 μg extract plus 5, 10, 25 and 50 μM TMPyP2 (P2); lanes 11-14, 10 μg extract plus 25, 50, 100 and 200 fold quadruplex 28R. EMSA in 5% polyacrylamide gel in TBE 1X.



S_6: (a) Circular dichroism as a function of temperature (20, 30, 40, 50, 60, 70, 80, 90 °C) of quadruplex 28R in 50 mM KCl, 50 mM Tris-HCl, pH 7.4; (b) Circular dichroism as a function of temperature (20, 30, 40, 50, 60, 70, 80, 90 °C) of 3 μM quadruplex 28R in the presence of 12 μM DIGP in 50 mM KCl, 50 mM Tris-HCl, pH 7.4; (c) Ellipticity at 260 nm as a function of temperature for 28R (r=0) and 28R+DIGP (r=4) in 50 mM KCl, 50 mM Tris-HCl, pH 7.4. The ratio [DIGP]/[28R] is indicated with r. Experiments conducted in 0.2 cm cuvette containing 3 μM 28R in the buffer.



S_7 (a) Rate of assembly between quadruplex F-28R-T and 28Y in the absence (curve 1) or presence of DIGP at r=1, 3 and 12 (curves 2, 3 and 4). Grey curve is quadruplex F-28R-T alone (curve 5). Experiment carried out in 50 mM Tris-HCl, pH 7.4, 50 mM KCl; 37°C; (b) as in a, but using Zn-DIGP. The data have been collected by a real time apparatus (CFX96 BioRad) measuring the FAM emission at 525 nm.



Pentaphyrin

S_8: Structure of pentaphyrin. For synthesis see J Med Chem. 2006 Jan 12;49(1):196-204. According to polymerase stop assay, this expanded porphyrin does not bind to G4-DNA (not shown).

4.4. G4-DNA formation in the *HRAS* promoter and rational design of decoy oligonucleotides for cancer therapy.

We focused on *HRAS* G-rich promoter and we identified two runs of guanines, namely hras-1 and hras-2, able to fold respectively into an anti-parallel and parallel quadruplex (qhras-1, qhras-2). We demonstrated this by means of DMS Footprinting, CD and FRET-melting analysis. Given the abundance of the GGGCGGG element we checked if some transcriptional factors were able to recognize hras-1 and/or hras-2. From the literature and by means of MATInspector software MAZ and Sp1 proteins were the best candidates to bind hras-1 and hras-2 sequences. We confirmed this prediction *in vitro* by EMSA, using recombinant MAZ-gst and Sp1-gst proteins, and *in vivo* by ChIP experiments. Usually MAZ and Sp1 regulate transcription of several genes cooperatively or in opposite ways due to their similar consensus sequence. So we focused on *HRAS* transcription mediated by these proteins. First we silenced MAZ and Sp1 genes using specific shRNA and we saw an inhibition of *HRAS* transcription. Than we built up a reporter system where luciferase expression is driven by *HRAS* promoter (pHRAS-luc). This showed how MAZ and Sp1 overexpression synergistically increased luciferase transcription. This effect was stronger when we used mutated forms of pHRAS-luc in which hras-1 and hras-2 quadruplexes were abrogated. All these findings suggested that *HRAS* quadruplex-forming sequences act as repressors of transcription. We reasoned that sequestering MAZ and Sp1 by means of decoys resembling hras-1 and hras-2, bearing (R)-1-O-[4-(1-Pyrenylethynyl) phenyl-methyl] glycerol and LNA modifications that increase stability and nuclease resistance, would interfere with *HRAS* transcription. After transfection of the T24 urinary bladder cancer cell line we observed a repression of *HRAS* transcription and a strong antiproliferative effect. To go further into details of these phenomena we performed some competition assays by EMSA in which we competed the MAZ/Sp1 to hras-1/hras-2 binding with increasing amount of G4-decoys. The decoys with the stronger effect *in vivo* on *HRAS* transcription and cell survival were those most intensively competing the DNA–protein binding, especially those concerning hras-1 sequence. Furthermore after 24 h from decoy treatment we saw caspase 3 and 7 activation.

G4-DNA Formation in the *HRAS* Promoter and Rational Design of Decoy Oligonucleotides for Cancer Therapy

Alexandro Membrino[1], Susanna Cogoi[1], Erik B. Pedersen[2], Luigi E. Xodo[1]*

1 Department of Medical and Biological Science, School of Medicine, Udine, Italy, 2 Nucleic Acid Center, Institute of Physics and Chemistry, University of Southern Denmark, Odense, Denmark

Abstract

HRAS is a proto-oncogene involved in the tumorigenesis of urinary bladder cancer. In the HRAS promoter we identified two G-rich elements, hras-1 and hras-2, that fold, respectively, into an antiparallel and a parallel quadruplex (qhras-1, qhras-2). When we introduced in sequence hras-1 or hras-2 two point mutations that block quadruplex formation, transcription increased 5-fold, but when we stabilized the G-quadruplexes by guanidinium phthalocyanines, transcription decreased to 20% of control. By ChIP we found that sequence hras-1 is bound only by MAZ, while hras-2 is bound by MAZ and Sp1: two transcription factors recognizing guanine boxes. We also discovered by EMSA that recombinant MAZ-GST binds to both HRAS quadruplexes, while Sp1-GST only binds to qhras-1. The over-expression of MAZ and Sp1 synergistically activates HRAS transcription, while silencing each gene by RNAi results in a strong down-regulation of transcription. All these data indicate that the HRAS G-quadruplexes behave as transcription repressors. Finally, we designed decoy oligonucleotides mimicking the HRAS quadruplexes, bearing (R)-1-O-[4-(1-Pyrenylethynyl) phenylmethyl] glycerol and LNA modifications to increase their stability and nuclease resistance (G4-decoys). The G4-decoys repressed HRAS transcription and caused a strong antiproliferative effect, mediated by apoptosis, in T24 bladder cancer cells where HRAS is mutated.

Citation: Membrino A, Cogoi S, Pedersen EB, Xodo LE (2011) G4-DNA Formation in the HRAS Promoter and Rational Design of Decoy Oligonucleotides for Cancer Therapy. PLoS ONE 6(9): e24421. doi:10.1371/journal.pone.0024421

Editor: Mark Isalan, Center for Genomic Regulation, Spain

Received June 3, 2011; Accepted August 9, 2011; Published September 8, 2011

Funding: AIRC 2010, ''Associazione Italiana per la Ricerca sul Cancro''. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscruipt.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: luigi.xodo@uniud.it

Introduction

The ras gene family consists of three functional proto-oncogenes (HRAS, NRAS and KRAS) that encode for guanine-binding proteins sharing a high homology ($p21^{RAS}$) [1]. These proteins, located on the inner cell membrane through a farnesyl group [2], are active when they are bound to GTP, and inactive when GTP is hydrolyzed to GDP [3]. Ras proteins regulate cellular responses to many extracellular stimuli, including mitogens and differentiation factors [4]. The ras genes are expressed in a tissue-specific fashion: HRAS is highly expressed in skin and skeletal muscles, KRAS in colon and thymus and NRAS in male germinal tissue [1]. The ras genes have similar structures and sequences with five exons, the first of which is noncoding, and conserved splicing sites. The introns, instead, have different lengths and sequences [1]. The ras proto-oncogenes are converted to oncogenes by point mutations that decrease the capacity of the encoded protein to hydrolyze GTP to GDP, with the result that $p21^{RAS}$ remains constitutively active. Hyperactivated ras proteins stimulate phosphorylation cascades including the Raf/MEK/ERK pathway which leads to uncontrolled cell proliferation [5,6]. Mutations in the ras genes are frequently found in many human tumors [7,8]. HRAS mutations are less common, but they have a high prevalence in skin papillomas and urinary bladder tumors [9]. As 80% of bladder tumors harbor HRAS mutations [10] and more than half of bladder tumors overexpress HRAS [11], both mutation and overexpression are important factors in the tumorigenesis of bladder cancer [12]. Actually, it has been recently shown that low-level expressions of constitutively active HRAS induced simple urothelial hyperplasia, while the doubling of the activated HRAS oncogene triggered rapidly growing and penetrating tumors throughout the urinary tract. Given the crucial role of HRAS overexpression and mutations in the tumorigenesis of bladder cancer, one attractive therapeutic strategy could be to inhibit HRAS transcription with molecules that are able to impair the activity of the gene promoter. For this aim we asked how HRAS transcription is regulated. We observed that the promoter of HRAS contains numerous copies of the GGGCGGG element or its complement. This G-box has been shown to interact with the Sp1 transcription factor [13,14]. Upstream of the transcription start site (TSS) there are runs of guanines spanning over three Sp1 sites, which are potential sites for G-quadruplex formation. We thus hypothesized that the G-rich elements might play a role in transcription regulation. G4-DNA are unusual structures stabilized by planar arrays of four guanines (G-quartet) linked one to the other by Hoogsteen hydrogen bonds [15]. The edges of the terminal G-quartets are connected by loops that can vary both in length and topology, giving rise to a variety of different conformations [16]. Genome-wide analyses have revealed that runs of guanines are abundant in gene promoter regions surrounding TSS [17–20]. It has been theorized therefore that G4-DNA might be involved in transcription regulation [21–25]. Our study provides compelling evidence that HRAS transcription is regulated by the interplay between Sp1, MAZ and G4-DNA, which acts as a transcription repressor. On the basis of this discovery we have designed G-rich oligonucleotides specific for

HRAS which have a strong antiproliferative effect in urinary cancer cells bearing a mutant HRAS. Although the cytotoxicity of certain G-rich oligonucleotides has been previously reported, their mechanism of action is not yet fully understood [26,27]. Our study shows that the designed quadruplex-forming oligonucleotides may act by sequestering MAZ and thus impairing HRAS transcription. For their potent antiproliferative effect in T24 urinary bladder cancer cells, G4-decoys seem to be very promising effector drugs for urinary bladder cancer therapy.

Results

The HRAS promoter is structurally polymorphic

The promoter of the human HRAS gene lacks typical TATA and CAAT boxes, contains a high G+C content (80%) and multiple copies of GGGCGGG (G-box), recognized by the transcription factor Sp1 [13,14]. The three G-boxes closest to the RNA start sites overlap quadruplex-forming sequences, namely hras-1 (435–462, accession number J00277) and hras-2 (506–530, J00277) (Figure 1). According to a recent study, quadruplex-forming sequences covering Sp1 binding elements are present in several genes [28]. We have obtained a first hint that the HRAS promoter is structurally polymorphic while sequencing the expression vectors specially constructed for this study. When sequencing primer-extension reactions were performed with primers complementary to the G-rich strand, Taq polymerase unexpectedly arrested at the hras-2 or hras-1 G-rich elements. In contrast, with primers complementary to the C-rich strand we did not observe any impediment. This suggested that both hras-1 and hras-2 sequences formed unusual structures. Our hypothesis was confirmed by polymerase-stop assays. We designed two linear wild-type templates containing hras-1 or hras-2 and one mutant template in which four GRT mutations were introduced into hras-2 to prevent quadruplex formation. Primer-extension reactions showed that Taq polymerase in the presence of potassium arrested at the 39 end of hras-2 or hras-1, just before the first run of guanines, in keeping with the formation of a G-quadruplex structure by each G-rich element (Figure S1).

Promoter sequences hras-1 and hras-2 form stable G4-DNA structures in vitro

An insight into the G-quadruplexes formed by the HRAS G-rich elements was obtained by DMS-footprinting, circular dichroism (CD) and fluorescence resonance energy transfer (FRET) exper-iments. Figure 2 shows the results obtained with sequence hras-1. The DMS-footprints of 27-mer hras-1 in water or buffer containing 100 mM Li^+ or Cs^+ (lanes 1, 4, 5) show that all guanines react with DMS. Instead, in the presence of 50, 100 or 140 mM KCl (lanes 2, 3, 9), the guanines of the G-runs A–D are progressively protected, while guanines G8 and G14 in the intervening ''TTGC'' and ''CGCA'' sequences are not. This cleavage pattern suggests that hras-1 folds into a G-quadruplex (qhras-1). In the presence of 50 nM TMPyP4, hras-1 gives a strong footprint either at low (1 mM) or high (100 mM) KCl concentration (lanes 7, 8) [29]. As expected, the mutant sequence h1-mut, bearing hras-1 with four GRT mutations that abolish the folding, does not give any footprint. To determine the strand orientation of qhras-1 we used CD (Figure 2c,d). The CD spectrum of qhras-1 in 100 mM KCl is characterized by positive and negative ellipticities at 287 and 260 nm, respectively, typical of an antiparallel conformation [30]. Heating the sample we obtained a melting curve (287 nm ellipticity versus temperature) that was not perfectly superimposable with the cooling curve, as the first was slightly biphasic while the second was monophasic. A more sensitive method of analysis was obtained by FRET-melting experiments with sequence hras-1 end-labelled with FAM and TAMRA at the 5 and 39 end, respectively. We obtained a well resolved biphasic melting curve with T_M's at 53 and 67uC indicating that hras-1 folds in at least two quadruplexes (Figure 2e). When the concentration of hras-1 was

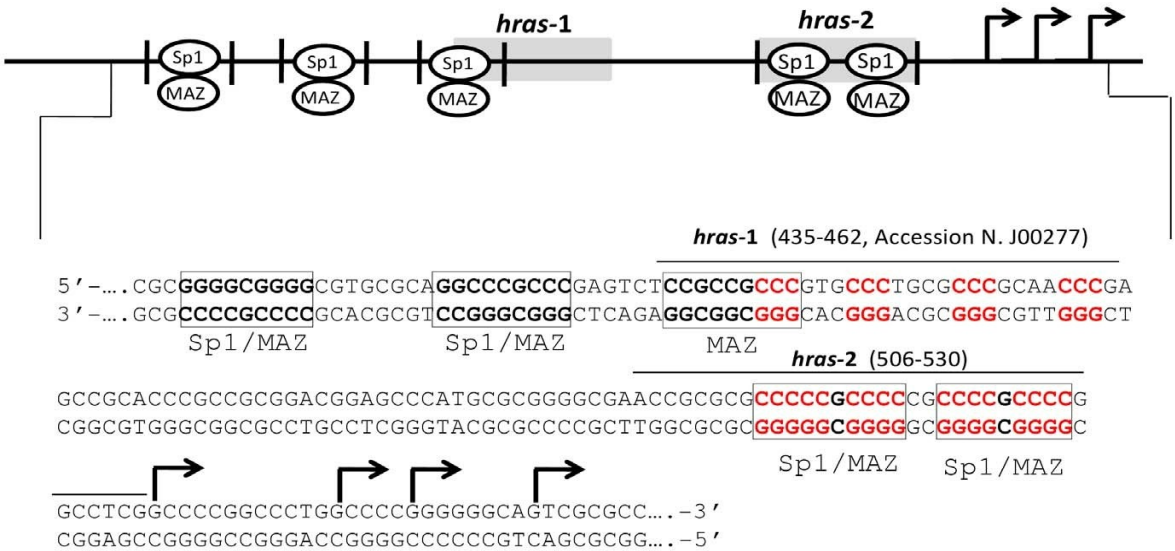


Figure 1. Sequence of the human *HRAS* promoter upstream of TSS. Two G-rich elements (hras-1 and hras-2) located upstream of transcription start site can potentially fold into G-quadruplex structures. The quadruplex-forming G-rich elements contain the binding sites for the transcription factors MAZ and Sp1.
doi:10.1371/journal.pone.0024421.g001

increased one order of magnitude from 200 nM to 2 mM, the T_M's did not change: a behavior typical of an intramolecular structure (not shown). Together, these data suggest that hras-1 adopts an antiparallel quadruplex which could assume different topologies: the one with lateral loops is shown in Figure 2f [16]. Although we know by footprinting and CD data the guanines that are involved in the formation of antiparallel qhras-1, an insight into structure of this quadruplex will only be obtained by NMR. Figure 3 shows the results obtained with sequence hras-2. The DMS-footprints of 24-mer hras-2 shows that in 10, 50, 100 and 140 mM KCl (lanes 2–5), the G-runs A–D are protected from DMS, while G10, G12, G13, G21 and G22 react with DMS. This indicates that the guanine triads forming the quadruplex should be G3-G4-G5, G7-G8-G9, G14-G15-G16, G18-G19-G20. The fact that G16 shows some reactivity to DMS suggests that the pentaguanine G12-G16 stretch can participate to the quadruplex either with G14-G15-G16 or with G13-G14-G15. The mutant h2-mut sequence and wild type hras-2 in 100 mM Li^+ or Cs^+ did not give any footprint (lanes 6–8). The footprinting in the presence of 50 nM TMPyP4 is very strong (lanes 9, 10). The CD spectrum of hras-2 shows a strong ellipticity at 260 nm indicative of a G-quadruplex with a parallel conformation (Figure 3c) [30]. This structure is so stable that the CD at 85uC still shows an intense 260 nm ellipticity. The stability at various KCl concentrations was determined with sequence hras-2 labeled with FAM and TAMRA. We performed FRET-melting experiments at 20, 40, 60, 100 and 140 mM KCl and obtained T_M values of 78 to 82, 85, .90 and .92uC, respectively (Figure 3d). In this case, too, the T_M's did not change when the concentration was increased one order of magnitude (from 200 nM to 2 mM) (not shown). Overall, our data demonstrate that hras-2 forms a parallel G-quadruplex (qhras-2) whose putative structure inferred by DMS-footprinting and CD should have either 1/1/4 or 1/2/3 loops (Figure 3e). A definitive structure assignment will only be possible on the basis of NMR data.

G4-DNA destabilizing point mutations in hras-1 and hras-2 strongly upregulate HRAS transcription

As sequences hras-1 and hras-2 are located immediately upstream of the transcription start site and form in vitro stable G-quadruplexes, we asked what happens to HRAS transcription when the capacity of quadruplex formation by sequences hras-1 and hras-2 is abolished. To address this point we constructed a plasmid, pHRAS-luc, bearing firefly luciferase driven by the HRAS promoter. From pHRAS-luc we obtained by site-directed mutagenesis two mutant plasmids: pHRAS-

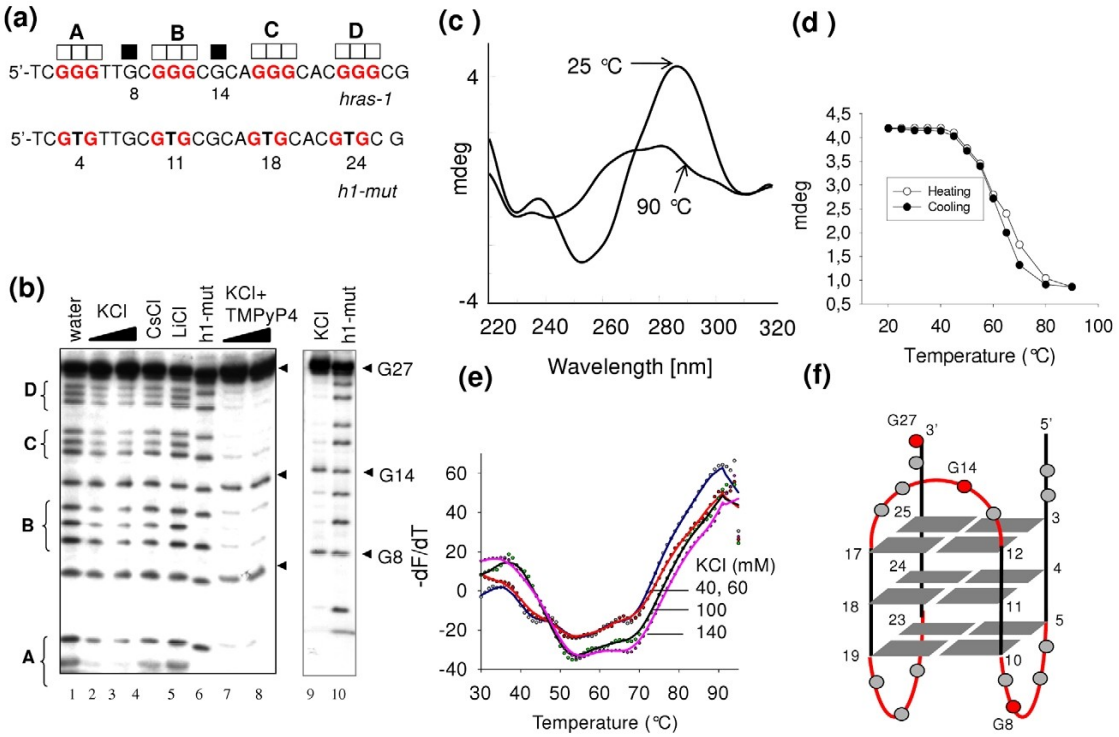


Figure 2. The G-rich element called *hras*-1 forms an antiparallel G-quadruplex structure. (a) Wild-type (hras-1) and mutated (h1-mut) sequences analyzed by DMS-footprinting. Full squares indicate DMS-reactive guanines, open squares indicate DMS-unreactive guanines; (b) DMS-footprinting of hras-1 in water (lane 1); hras-1 in 50 and 100 mM KCl (lanes 2 and 3); hras-1 in 100 mM CsCl or LiCl (lanes 4 and 5); h1-mut in 100 mM KCl (lane 6); hras-1 in 50 mM TMPyP4, 1 mM KCl (lane 7); hras-1 in 50 mM TMPyP4, 100 mM KCl (lane 8); hras-1 in 140 mM KCl (lane 9); h1-mut in 100 mM KCl (lane 10); (c) CD spectra at 25 and 90 uC show that hras-1 folds into an antiparallel G-quadruplex; (d) Melting curves for quadruplex hras-1 obtained plotting the 287-nm ellipticity against T. A T_M of about 63uC was obtained by heating and cooling curves; (e) FRET-melting curves of quadruplex hras-1 labelled with FAM and TAMRA show that it folds in two G-quadruplexes with T_M of 53 and 67uC (Ex. 475 nm, Em. 520 nm); (f) Possible conformations for hras-1, based on footprinting and CD data, are antiparallel quadruplexes with lateral and lateral/diagonal loops. The quadruplex with lateral loops is shown in the figure.
doi:10.1371/journal.pone.0024421.g002

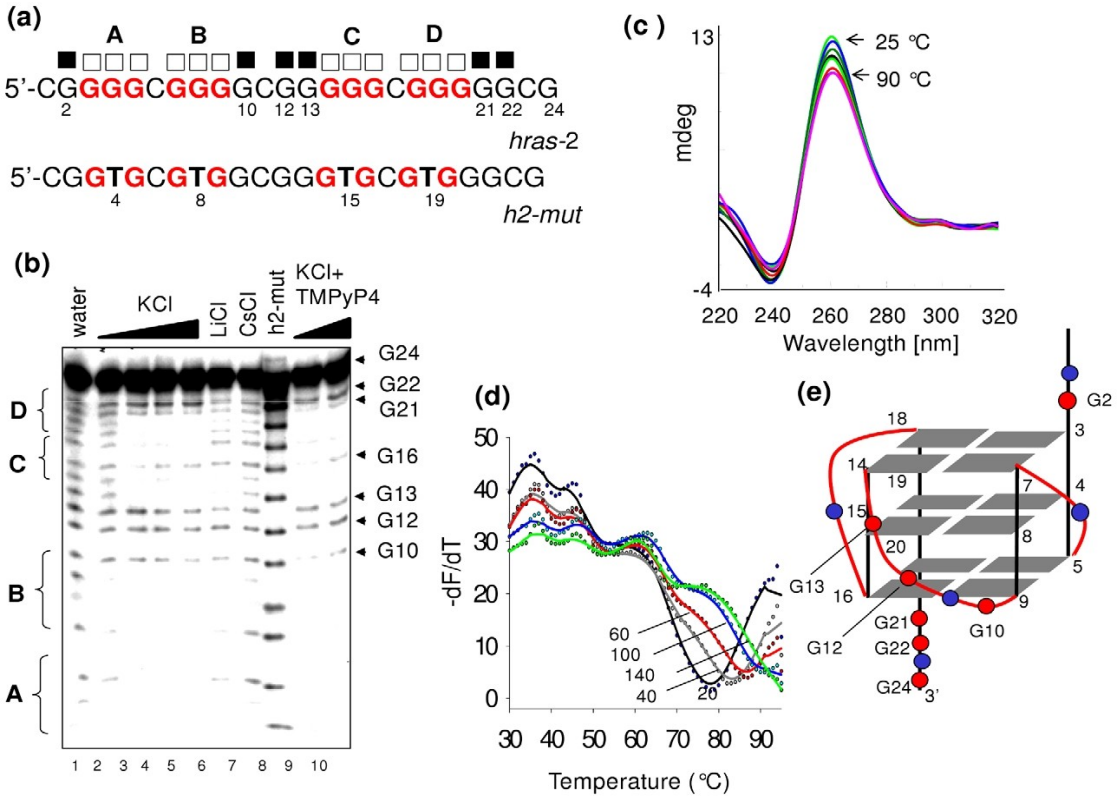


Figure 3. Sequence *hras*-2 forms a parallel G-quadruplex structure. (a) wild-type and mutated hras-2 sequences analyzed by DMS-footprinting. Full squares indicate DMS-reactive guanines, open squares indicate DMS-unreactive guanines; (b) DMS-footprinting of hras-2 in water (lane 1); hras-2 in 1, 10, 50, 100 mM KCl (lanes 2–5); hras-2 in 100 mM LiCl or CsCl (lanes 6,7); h2-mut in 100 mM KCl (lane 8); hras-2 in 50 nM TMPyP4, 1 mM KCl (lane 9); hras-2 in 50 nM TMPyP4, 100 mM KCl (lane 10). Guanines G10, G12, G13 and G21, G22 are cleaved while the other guanines are not (except G16 showing a low reactivity). Mutant h2-mut does not show any footprinting in 100 mM KCl; (c) CD spectra of quadruplex hras-2 at various temperatures between 25 and 90uC suggest the formation of an antiparallel G-quadruplex; (d) FRET-melting of quadruplex hras-2 at 20, 40, 60, 100 and 140 mM KCl shows only one transition with T_M of 78, 82, 85, .90, .92uC; (e) Possible G-quadruplex structure for hras-2, a parallel conformation with 1/1/4 or 1/2/3 loops, based on footprinting and CD data.
doi:10.1371/journal.pone.0024421.g003

mut1 and pHRAS-mut2, where two guanines in the second and third G-tetrads of the putative quadruplexes were replaced with thymine/cytosine or thymines (Figure 4a). These mutations abrogated the capacity of sequences hras-1 and hras-2 to fold into a quadruplex (Figure S2). The wild-type and mutant plasmids were co-transfected in HeLa cells with pRL-CMV, a vector encoding for Renilla luciferase. Forty-eight hours after transfection we measured firefly and Renilla luciferases activity in the lysates of untreated and treated cells. The results reported in Figure 4b show that blocking quadruplex formation causes a dramatic increase of firefly luciferase expression, up to 5-fold compared to control. This strongly indicates that the G-quadruplexes formed by sequences hras-1 and hras-2 are structural elements of the HRAS promoter that behave as repressors, as observed for the CMYC gene [21].

G4-DNA stabilizing guanidinium phthalocyanines repress HRAS transcription

Given that quadruplex DNA behaves as a repressor element for HRAS, we examined the effect on transcription of G4-DNA stabilizing ligands. Due to their specificity for G4-DNA, in our experiments we used as ligands modified phthalocyanines: tetrakis-(diisopropyl-guanidine) phthalocyanine (DIGP) and its Zn-containing derivative Zn-DIGP (Figure 5a) [31–33]. HeLa cells were first treated with 1 or 5 mM DIGP or Zn-DIGP for 24 h and then co-transfected with a mixture of pHRAS-luc and pRL-CMV. After transfection, the cells were let to grow for 48 h before firefly and Renilla luciferases activity were measured with a luminometer. As a control (i) we treated the cells with TMPyP2, a porphyrin that does not bind to G4-DNA [34]; (ii) we used as a reporter vector pHRAS-mut1 or pHRAS-mut2 bearing a mutated G-element unable to form a quadruplex structure. Figure 5b, c, d shows that DIGP and Zn-DIGP strongly inhibit luciferase from wild-type pHRAS-luc, while TMPyP2 does not. Moreover, the phthalocyanines do not inhibit luciferase in the cells treated with the mutant plasmids pHRAS-mut1 or pHRAS-mut2, in keeping with the fact that the G-quadruplexes cannot be formed by these vectors. These data provide further evidence that quadruplexes qhras-1 and qhras-2 act as transcription repressors.

Transcription factors Sp1 and MAZ bind to quadruplex forming sequences (QFS) hras-1 and hras-2

As our transcription data suggest that both sequences hras-1 and hras-2 are critical for transcription, we asked if they are recognized

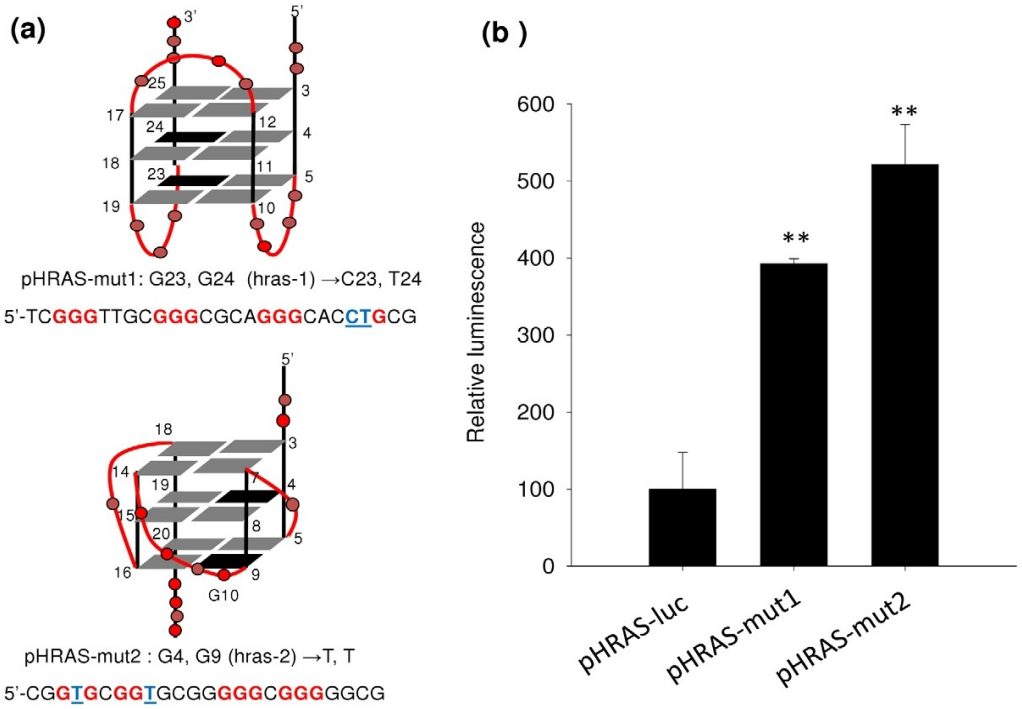


Figure 4. Effect of quadruplex DNA on *HRAS* transcription. (a) Two G4-DNA destabilizing point mutations in the hras-1 or hras-2 element have been introduced. Plasmid pHRAS-mut1 contains a mutated hras-1 element, while plasmid pHRAS-mut2 contains a mutated hras-2 element; (b) Dual luciferase assay showing that the point mutations cause up to a 5-fold increase in HRAS promoter activity (see Figure 5 legend). doi:10.1371/journal.pone.0024421.g004

by nuclear proteins. An answer to this question was obtained by mobility-shift assays with a HeLa nuclear extract. Figure 6a shows that both HRAS duplexes form with a HeLa nuclear extract distinct DNA-protein complexes, thus pointing out the relevance of these sequences for HRAS transcription. Ishii et al. showed that sequence hras-2, which contains two copies of GGGCGGG, is bound by Sp1 [13,14]. However, sequences hras-1 and hras-2 should also interact with the myc associated zinc-finger transcription factor (MAZ), its binding site being (G/C)GG(C/A)GGG [35,36]. In fact, it has been reported that MAZ and Sp1 often regulate transcription in a cooperative way [37,38].

To prove that Sp1 and MAZ bind to sequences hras-1 and hras-2 under in vivo conditions, we performed chromatin immunoprecipitation (ChIP) experiments. Living HeLa cells were treated with formaldehyde to crosslink the DNA-protein complexes, chromatin was sheared into fragments and then immunoprecipitated with anti-MAZ and anti-Sp1 antibodies (Abs). The abundance of HRAS promoter sequences in the immunoprecipitates was measured by PCR using primers specific for sequences hras-1 and hras-2. The results reported in Figure 6b show that: IgG Ab did not immunoprecipitate DNA-protein complexes containing sequence hras-1 or hras-2 (negative control); anti-MAZ Ab did immunopre-cipitate a DNA-protein complex containing hras-1 and hras-2; anti-Sp1 Ab pulled down a complex with hras-2. By measuring the intensity of the bands, we found that the HRAS sequences were more abundant in the immunoprecipitates with anti-MAZ and anti-Sp1 Abs than in the IgG Ab immunoprecipitate. We thus concluded that under in vivo conditions MAZ is associated to sequences hras-1 and hras-2, while Sp1 is associated to sequence hras-2. This was confirmed by EMSA with recombinant MAZ and Sp1 (Figure S3). As previous studies have shown that MAZ binds to G4-DNA from the murine KRAS [24] and c-myb [39] promoters, we explored whether it also binds to the HRAS quadruplexes. We performed EMSA with recombinant MAZ and Sp1, which were expressed in E. coli as GST fusion proteins (Figure S4). Figure 6c shows that at pH 7.4, 50 mM KCl, qhras-1 and qhras-2 with increasing amounts of MAZ-GST form - in the presence of 100-fold excess nonradiolabelled poly d(I-C) - two retarded bands due to the formation of two DNA-protein complexes, most likely with 1:1 and 1:2 stoichiometry. It is important to note that in 50 mM CsCl, where the HRAS sequences are unstructured (see DMS footprinting), both complexes are destabilized indicating that the DNA-protein interaction is mediated by the quadruplex structure. In a buffer at pH 9 where MAZ probably modifies its own folding, the DNA-protein interaction is also inhibited. We also tested the binding of Sp1-GST to the HRAS quadruplexes and found that Sp1-GST interacts with the antiparallel qhras-1 quadruplex, but in a weaker way compared to MAZ.

MAZ and Sp1 synergistically activate HRAS transcription

To prove that MAZ and Sp1 are involved in HRAS transcription, we co-transfected plasmid pHRAS-luc in HeLa cells either with pMAZ (encoding for MAZ) or pSp1 (encoding for Sp1) or with a mixture of both plasmids (Figure 7a). When pMAZ or pSp1 was cotransfected with the reporter vector, the level of luciferase expression increased by 50% compared to control.

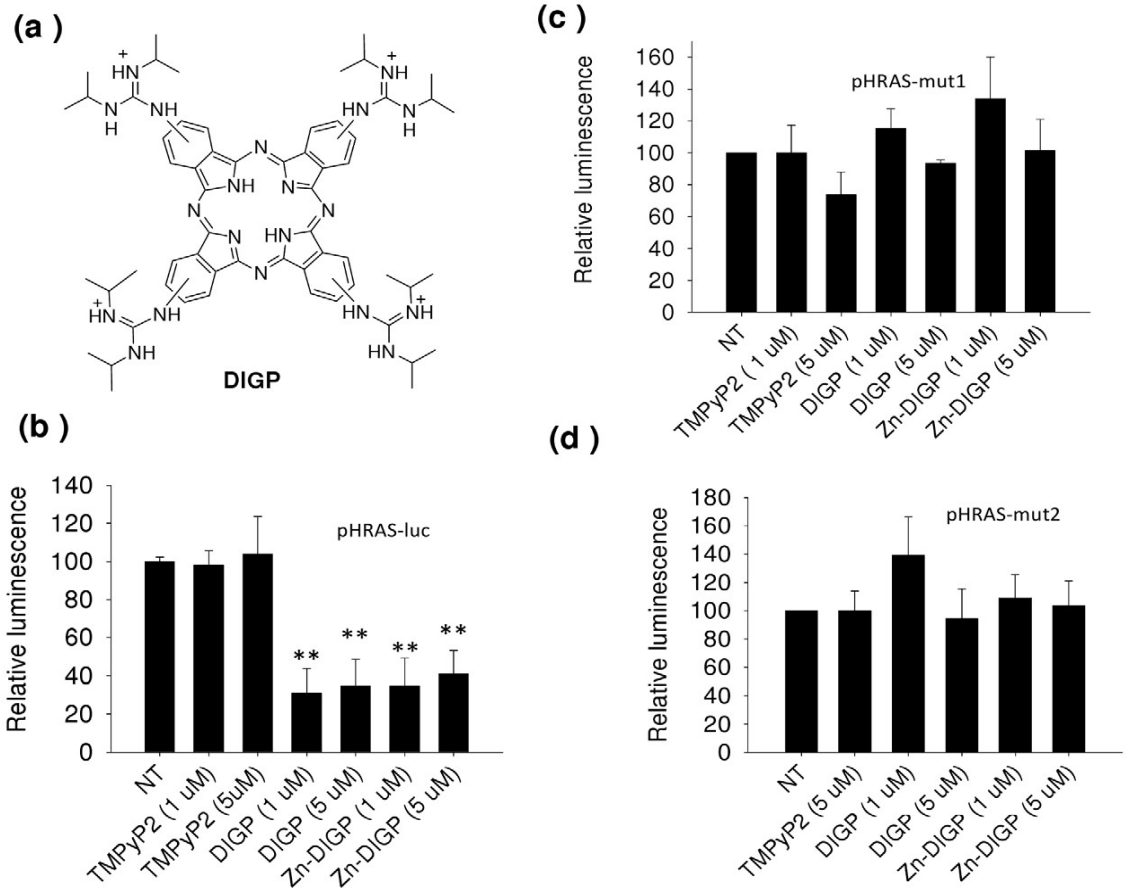


Figure 5. Effect of quadruplex-stabilizing ligands on *HRAS* transcription. (a) Structure of tetrakis-(diisopropyl-guanidine) phthalocyanine (DIGP) used, with its Zn-containing analogue Zn-DIGP, for luciferase experiments; (b) Dual luciferase assays showing that guanidinium phthalocyanines DIGP and Zn-DIGP, stabilizing hras-1 and hras-2 quadruplexes, repress HRAS promoter activity. This effect is not observed with the mutant plasmids pHRAS-mut1 and pHRAS-mut2 (control) (panels c and d). Relative luciferase is given by $R_T/R_{NT}6100$, where R_T and R_{NT} are (firefly luciferase)/(Renilla luciferase) in T24 cells treated with phthalocyanines and untreated cells. Differences from the control are supported by Student's t test, P,0.05 (one asterisk), P,0.01 (two asterisks).
doi:10.1371/journal.pone.0024421.g005

When both pMAZ and pSp1 were cotransfected with the reporter vector, transcription increased nearly 3-fold over control, indicating that both proteins synergistically activated HRAS promoter. The synergy was stronger (7-fold compared to control) when pMAZ and pSp1 were cotransfected with the mutant vector pHRAS-mut1 or pHRAS-mut2, bearing mutated hras-1 or hras-2 sequences which are unable to form quadruplex structures. This suggests that transcription is activated when sequences hras-1 and hras-2 are in the unfolded duplex conformation. Furthermore, as a proof that transcription requires both MAZ and Sp1, we knocked down with validated shRNA each of the two transcription factors separately (Figure S5). HeLa cells were treated with shRNA specific for MAZ or Sp1 and the levels of HRAS transcripts were measured by real-time PCR, at 48 and 72 h following treatment. It can be seen that HRAS transcription was reduced to 30 and 20% of control, 48 h following treatment with anti MAZ and anti Sp1 shRNA, respectively (Figure 7b). Taken together our data demonstrate that HRAS transcription is activated synergistically by MAZ and Sp1.

MAZ destabilizes the HRAS G-quadruplexes

Considering that MAZ is essential for HRAS transcription and recognizes qhras-1 and qhras-2, we asked whether these quadruplexes are unfolded by MAZ. To answer this question, we performed FRET-melting experiments with recombinant MAZ-GST. Quadruplex qhras-1 end-labelled with FAM and TAMRA was incubated in 50 mM KCl, for 30 min in the presence of increasing amounts of MAZ-GST or control protein (BSA or trypsinogen). The melting curves (2dF/dT) showed that qhras-1 alone melts with its typical biphasic profile with T_M at 53 and 67uC (Figure 8a curve 1). But when qhras-1 was incubated with MAZ-GST at (moles of protein)/(moles of DNA) ratio r = 2.5, 5, 10 (curves 4, 5 and 6), the melting profile changed significantly as the T_M's fell to ,46uC. This indicates that both qhras-1 quadruplexes are destabilized by MAZ-GST. As expected, unspecific proteins such as BSA or trypsinogen at r = 10 did not affect the melting of qhras-1 (curves 2,3). Due to its high stability in potassium (T_M = 78uC in 10 mM KCl), qhras-2 was analyzed in 100 mM NaCl where the T_M was 65uC. Incubated with MAZ-

GST at r = 2.5, 5, 10, the quadruplex melted at 42uC, indicating that also qhras-2 was destabilized by MAZ-GST, but not by BSA or trypsinogen. We also ascertained that GST had no influence on the melting of the HRAS quadruplexes (Figure S6b). An additional control that we performed to rule out that bacterial proteins did not contribute to quadruplex destabilization was to mix Glutathi-one Sepharose 4B resin with an extract obtained from non-transformed BL21 DE3 plysS bacteria. An SDS-PAGE analysis of the eluate with 10 mM reduced glutathione showed that no bacterial proteins bound non-specifically to the resin (Figure S6a).

We were surprised about the unfolding activity of MAZ, as we recently reported that MAZ stabilized the murine KRAS quad-ruplex [24]. However, it should be borne in mind that MAZ having six zinc fingers can interact with DNA in a complex way, as observed with qTBP42 and CBF-A [40,41]. These proteins disrupt the dimeric quadruplex formed by the FMR1 $d(CGG)_n$ repeat but they also stabilize the telomeric quadruplex $d(TTAGGG)_n$. CBF-4 and qTBP42 have four RNP domains but only two are involved in G4-DNA binding. It has been demonstrated that different RNP combinations are responsible for either the stabilizing or the destabilizing activity.

G4-DNA decoy ODNs mimicking HRAS quadruplexes inhibit transcription and cell growth

Considering that HRAS transcription is activated by a combined action of MAZ and Sp1 and that qhras-1 and qhras-2 bind to MAZ (qhras-1 also binds to Sp1), we designed a decoy strategy against HRAS oncogene. We reasoned that these oligonucleotides mimicking quadruplexes qhras-1 or qhras-2 (G4-decoys) should sequester MAZ and inhibit HRAS transcription as well as cell growth (Figure 8c). To increase the stability of the G4-decoys we inserted in their sequence one or two units of (R)-1-O-[4-(1-Pyrenylethynyl) phenylmethyl] glycerol (P) to cap the quadruplex ends [42]. To increase their resistance against endogenous nucleases, we introduced LNA residues at the 39 end and in one loop. We designed three G4-decoys mimicking quadruplex qhras-1 (3, 4 and 5) and three mimicking quadruplex qhras-2 (6, 7 and 8) (Figure S7) (Table 2). The latter show CD spectra similar to that of wild type qhras-2, with a strong ellipticity at 260 nm typical of a parallel quadruplex conformation, while the former show a CD slightly different from that of qhras-1, with two positive ellipticities at 260 and 287 nm, suggesting that they should form a mixed parallel-antiparallel quadruplex [43] (Figure 9a). The antiprolif-erative activity of the designed G4-decoys was tested in two types of cells: HeLa and T24 urinary bladder cancer cells that harbour a mutant HRAS [codon 12 GGC (Gly) is changed in GTC (Val)], expressing a hyperactivated HRAS protein [44]. In a first set of experiments we delivered the G4-decoys (5 mM) without any transfectant agent and did not observe any effect on cell proliferation. To see if this was due to a poor oligonucleotide uptake, we analyzed by confocal microscopy T24 bladder cells treated for 24 h with 5 mM decoy 3 covalently labelled with fluorescein (3-F). Figure 9b shows typical images: the nuclei stained with propidium iodide (red fluorescence), the intracellular

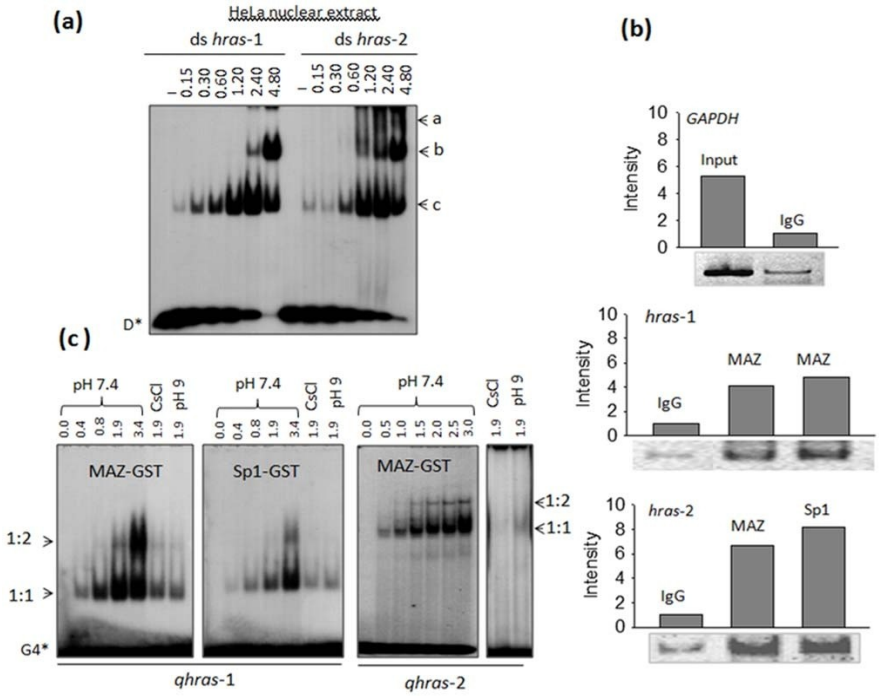


Figure 6. Proteins MAZ and Sp1 interact with sequences *hras*-1 and *hras*-2. (a) EMSA showing the formation of DNA-protein complexes between duplexes hras-1 and hras-2 and HeLa nuclear extract; amounts of extract used (mg) are indicated, radiolabelled DNA was about 4 nM. Letters a, b and c indicate the DNA-protein complexes; (b) Chromatin immunoprecipitation analysis showing the interaction of MAZ and Sp1 with sequences hras-1 and hras-2. The interaction between MAZ and Sp1 with the HRAS sequences was analyzed by a PCR amplification of 190 bp (hras-1) and 161 bp (hras-2) fragments. The two bands for complex MAZ:hras-1 have been obtained in the presence and absence of Mg^{2+} in the mixture (c) EMSA showing the binding of MAZ-GST and Sp1-GST to quadruplexes qhras-1 and qhras-2.
doi:10.1371/journal.pone.0024421.g006

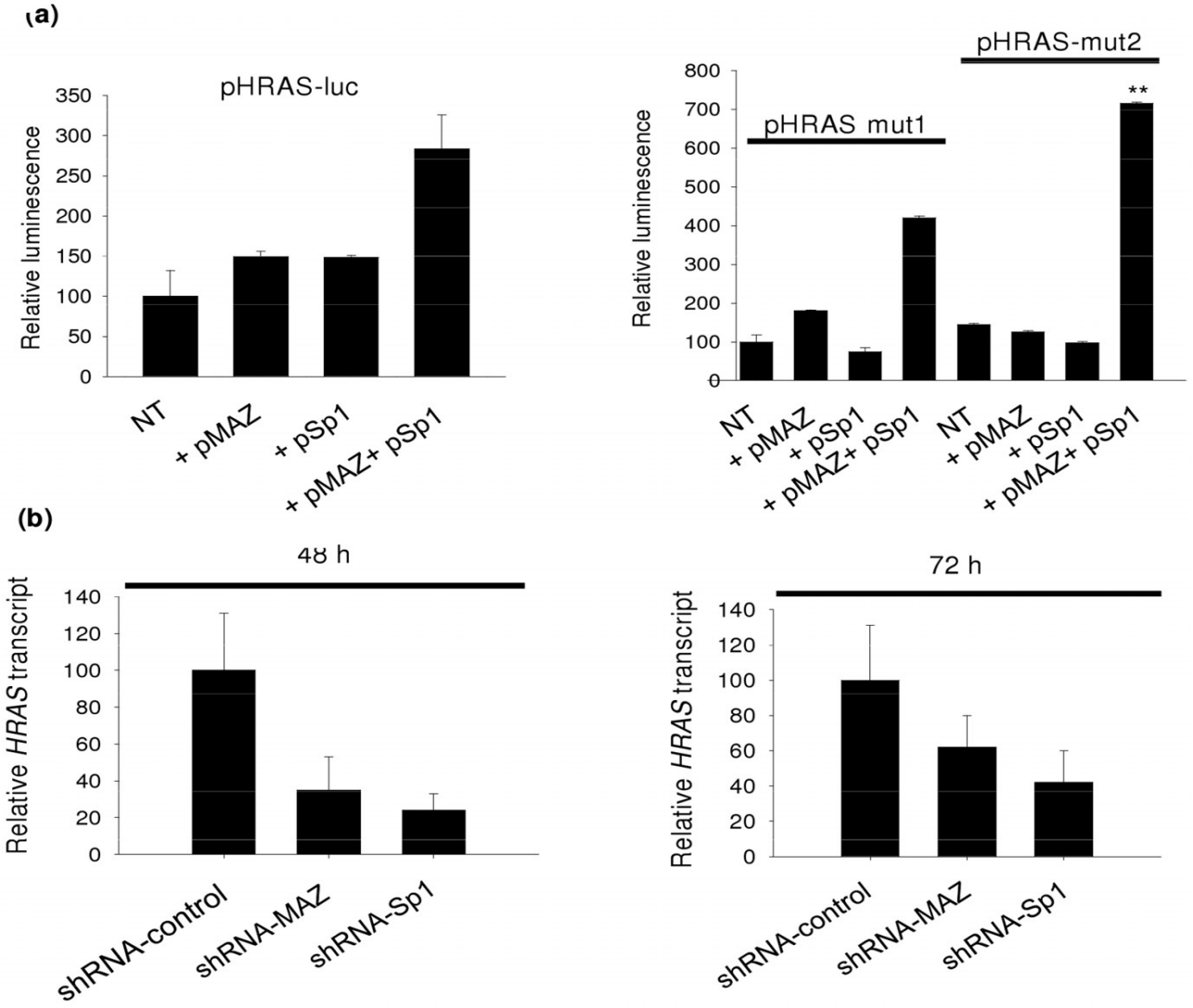


Figure 7. MAZ and Sp1 sinergistically activate *HRAS* promoter. (a) HeLa cells transfected with pHRAS-luc/pRL-CMV, pHRAS-luc/pRL-CMV/pMAZ; pHRAS-luc/pRL-CMV/pSp1; pHRAS-luc/pRL-CMV/pMAZ/pSp1. The dual luciferase assays were performed 24 h following transfection. Relative luminiscence is given by $R_T/R_{NT}6100$, where R_{NT} is (firefly luciferase)/(Renilla luciferase) in T24 cells treated with only pHRAS-luc and pRL-CMV, while R_T is (firefly luciferase)/(Renilla luciferase) in T24 cells treated with pHRAS-luc +pMAZ and/or pSp1+pRL-CMV. Differences from the control are supported by Student's t test, P,0.05 (one asterisk), P,0.01 (two asterisks); (b) Level of HRAS transcript in HeLa cells in which MAZ or Sp1 was knocked down by validated shRNAs.
doi:10.1371/journal.pone.0024421.g007

distribution of the oligonucleotide (green fluorescence) and the overlay of both emissions. It appears clear that 3-F localizes basically in the cytoplasm, and this is the reason why the decoys are not active when they are delivered as free molecules. We then treated the cells with 3-F (350 nM) mixed to polyethylenimine (PEI). Figure 9b shows that despite the low concentration, 3-F was efficiently taken up by the cells and uniformly distributed in the nucleus. We therefore decided to perform our proliferation experiments using PEI as a transfectant agent. We did a dose-response experiment by delivering the G4-decoys in two doses to T24 cells, one 48 h after the other. Three days after the first delivery, we performed a resazurin proliferation assay. It can be seen that the three G4-decoys specific for hras-1 (3, 4 and 5) promoted a dramatic inhibition of cell growth with IC_{50} of about 700 nM (Figure 9c). In contrast, only compound 6 specific for qhras-2 showed some antiproliferative effect in T24 cells. Remarkably, the oligonucleotides with the sequence of hras-1 (H1) or hras-2 (H2) and control oligonucleotide 1450 (not folding into a quadruplex) did not show any antiproliferative effect (H1 and H2 are probably degraded by nucleases). A proliferation assay performed as a function of time showed that the growth inhibition promoted by the G4-decoys did not weaken over a period of 144 h (Figure S8).

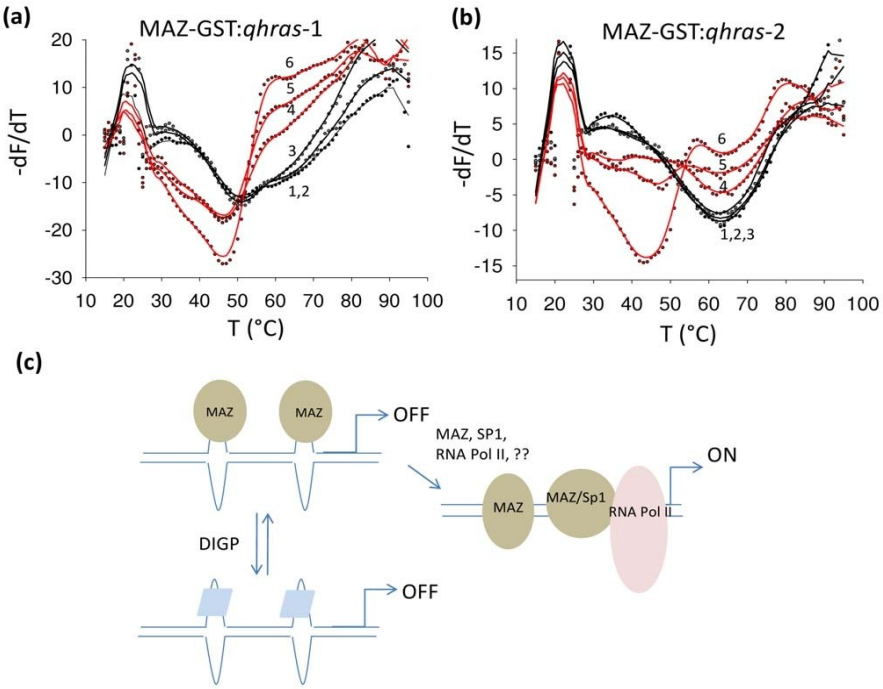


Figure 8. MAZ destabilizes quadruplexes *hras*-1 and *hras*-2. FRET-melting experiment showing that in 50 mM KCl, MAZ-GST at r = 2.5, 5 and 10 (r = [MAZ]/[G4-DNA]) (curves 4, 5 and 6) destabilizes quadruplex qhras-1 in 50 mM KCl, 50 mM Zn-acetate [curve 1 (panel a)] and quadruplex qhras-2 in 100 mM NaCl, 50 mM Zn-acetate [curve 1, (panel b)]. BSA and TR (trypsinogen) at r = 10 do not have any effect on qhras-1 and qhras-2 quadruplexes) (curves 2 and 3, panels a and b); (c) Schematic representation of HRAS transcription regulation proposed by this study. When the critical promoter elements hras-1 and hras-2 are folded, they behave as transcription repressors. This is suggested by the fact that quadruplex-destabilizing point mutations in hras-1 and hras-2 result in a significant increase of transcription, while quadruplex-stabilizing phthalocyanines are found to repress transcription. MAZ, through its capacity to recognize the quadruplex structures, should be recruited to the promoter critical region. The MAZ-DNA interaction destabilizes the G-quadruplexes, the critical hras-1 and hras-2 elements assume a duplex conformation that favors the binding of Sp1 and other proteins of the transcription machinery. These events result in the activation of transcription.
doi:10.1371/journal.pone.0024421.g008

According to the postulated mechanism of action, the G4-decoys 3, 4 and 5 should repress HRAS transcription by taking MAZ (and Sp1) away from the promoter. Actually, we found that the level of HRAS transcript, determined by real-time PCR 24 h after decoy treatment, was reduced up to 30% of control, in T24 cells treated with the active decoys 3, 4 and 5 (specific for qhras-1). In keeping with proliferation data, the decoys 7 and 8 that are not active (specific for qhras-2) do not repress HRAS transcription (Figure 10a). This result correlates with the finding that decoys 3, 4 and 5 strongly compete with the binding of MAZ to the hras-1 quadruplex, suggesting that these active decoys bind to MAZ (Figure 10b).

Finally, an insight into the killing mechanism caused by the designed G4-decoys was obtained by measuring the activity of caspases 3 and 7 in untreated and decoy-treated T24 cells. Figure 10c shows that decoys with the strongest antiproliferative activity (3, 4 and 5) considerably activate caspases 3/7 (24 h after treatment), suggesting that they promote cell death mediated by apoptosis.

Discussion

The data of this study show that HRAS transcription is activated by Sp1 and MAZ and repressed when the binding sites of these proteins closest to TSS assume a folded G-quadruplex structure.

This is a new and compelling piece of evidence pointing to a transcription mechanism which involves a simple on-off switch in a gene regulatory region where unusual G-quadruplexes behave as repressors. This was first proposed by Hurley and co-workers for the CMYC gene [21]. Following this transcription model we designed several decoy oligonucleotides in quadruplex conformation eliciting a potent antiproliferative effect in T24 urinary bladder cancer cells bearing mutant HRAS.

ChIP assays showed that the G-rich elements called hras-1 and hras-2, located upstream of TSS, are bound by the zinc-finger proteins MAZ and Sp1. MAZ interacts with these promoter sequences in a complex way, as it recognizes both duplex and quadruplex conformations of hras-1 and hras-2. The binding to the quadruplexes is catalytic in nature as qhras-1 and qhras-2 bound to MAZ go through a destabilization process that decreases the stability of the DNA-protein complexes (see EMSA). This means that PAGE gives here only an apparent affinity between MAZ and the HRAS quadruplexes, for which we estimated a K_D of about 1.5 mM. One should also consider that by using recombinant bacterially expressed proteins the binding data might be underestimated, as recombinant proteins do not undergo the post-translation modifications which occur in eukaryotic cells necessary for optimal binding. For instance, MAZ shows optimal binding to DNA when it is phosphorilated [45]. As a comparison, it has been estimated by a filter binding assay that recombinant

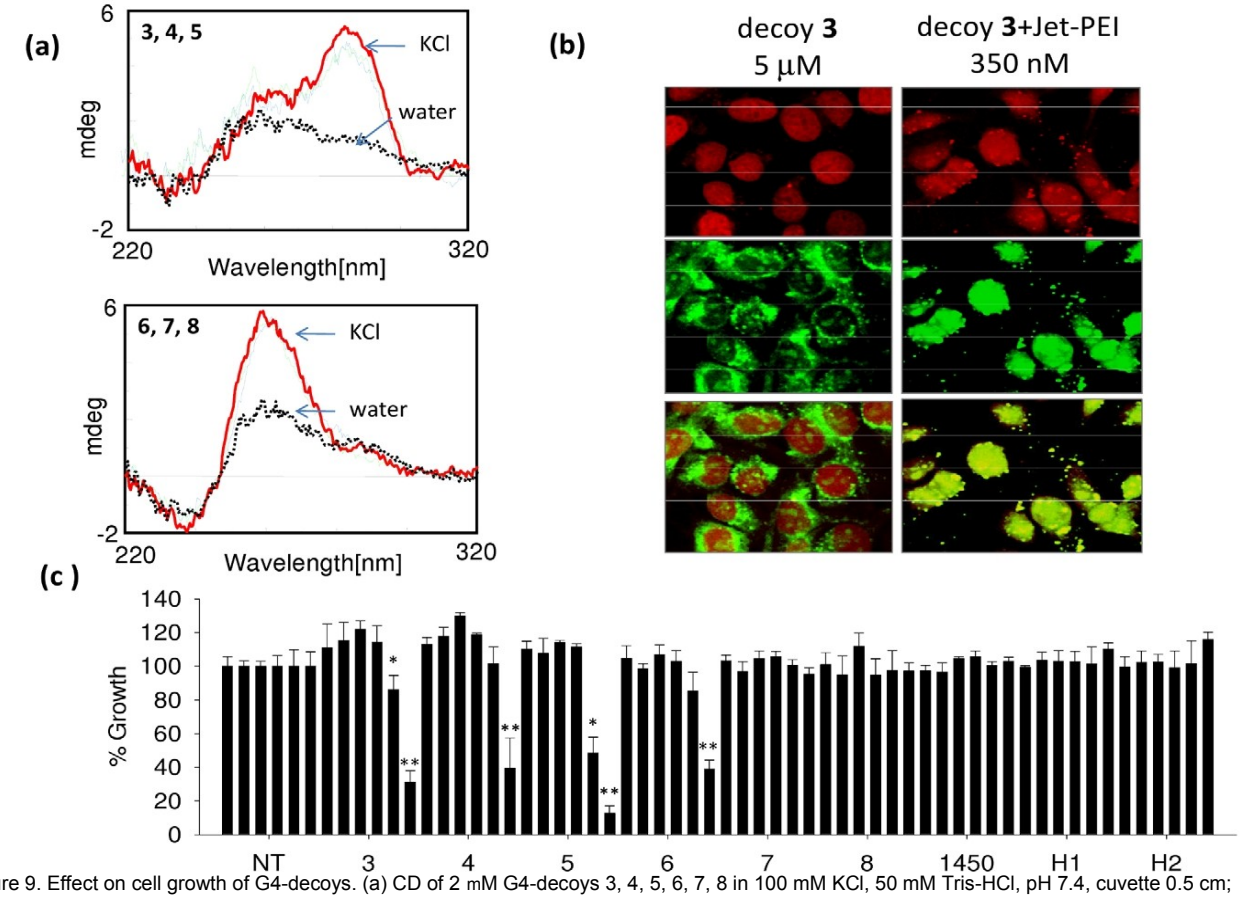


Figure 9. Effect on cell growth of G4-decoys. (a) CD of 2 mM G4-decoys 3, 4, 5, 6, 7, 8 in 100 mM KCl, 50 mM Tris-HCl, pH 7.4, cuvette 0.5 cm; (b) confocal microscopy images of T24 cells treated with decoy 3 labelled with fluorescein in the presence and absence of polyethylenimmine (jet-PEI). Top panel shows the nuclei stained with propidium iodide, mid panel shows the intracellular distribution of the decoy, bottom panel shows the overlay; (c) Dose response assays in T24 urinary bladder cancer cells treated with the G4-decoys at increasing concentrations (100, 200, 300, 400, 600 and 800 nM - double delivery); differences from the control are supported by Student's t test, P,0.05 (one asterisk), P,0.01 (two asterisks). doi:10.1371/journal.pone.0024421.g009

nucleolin binds to parallel quadruplexes with a K_D between 79 to 367 nM and it binds to mixed parallel antiparallel quadruplexes with a K_D of 0.45–2.5 mM [46]. Instead, by surface plasmon resonance assay it has been found that recombinant nucleophosmin binds to the CMYC quadruplex with a K_D of 1.9 mM [47].

MAZ was identified as a G-box binding transcription factor for CMYC [48]. Previous studies have shown that MAZ can either activate [35–37,49–51] or repress [39,52] transcription. Furthermore, in certain genes MAZ regulates transcription together with Sp1, as the two proteins can bind to the same guanine blocks: the consensus sequence for MAZ is G(g/c)GGc/a GGGG(c/a)(g/t) while that of Sp1 is (g/t)GGGCGG(g/a)(g/a)(c/t) [35]. When MAZ or Sp1 was separately knocked-down with specific shRNAs, HRAS transcription dropped respectively to 30% or 20% of control. Conversely, when MAZ and Sp1 were overexpressed, a synergistic effect was observed and HRAS transcription increased 3-fold compared to control.

To explore the role of G4-DNA on HRAS transcription we introduced two quadruplex destabilizing GRT(C) point mutations in sequences hras-1 or hras-2 and found that transcription increased 5-fold compared to control. A similar behavior was previously reported for CMYC and CMYB genes [21,22,39]. Our conclusion is that both HRAS G-quadruplexes behave as a molecular on-off switch that either provides the binding sites to MAZ and Sp1 or subtracts them. This is also supported by the fact that guanidinium phthalocyanines stabilize the HRAS G-quadruplexes and repress luciferase from pHRAS-luc to 20–30% of control, but not from mutant vectors pHRAS-mut1 and pHRAS-mut2. In addition to function as a transcription factor, MAZ is also able to remove the structural blocking of transcription by the quadruplex structures, as indicated by FRET-melting data.

In the light of all these findings, we designed G4-decoys mimicking HRAS G-quadruplexes, which show a strong antiproliferative activity in T24 urinary bladder cancer cells harboring mutant HRAS. We hypothesized that the G4-decoys should take MAZ away from the promoter and inhibit HRAS transcription. We found that in T24 bladder cells (but also in HeLa cells, not shown) the decoys specific for quadruplex qhras-1 (3, 4, and 5) displayed a dramatic inhibitory activity on cell growth, at a concentration as low as 700 nM. Instead, only decoy 6 mimicking quadruplex qhras-2 showed some activity. In keeping with previous observations, our data suggest that quadruplex formation per se is not sufficient to give rise to a bioactivity, as decoys 7 and 8 though forming a stable quadruplex, are not active.

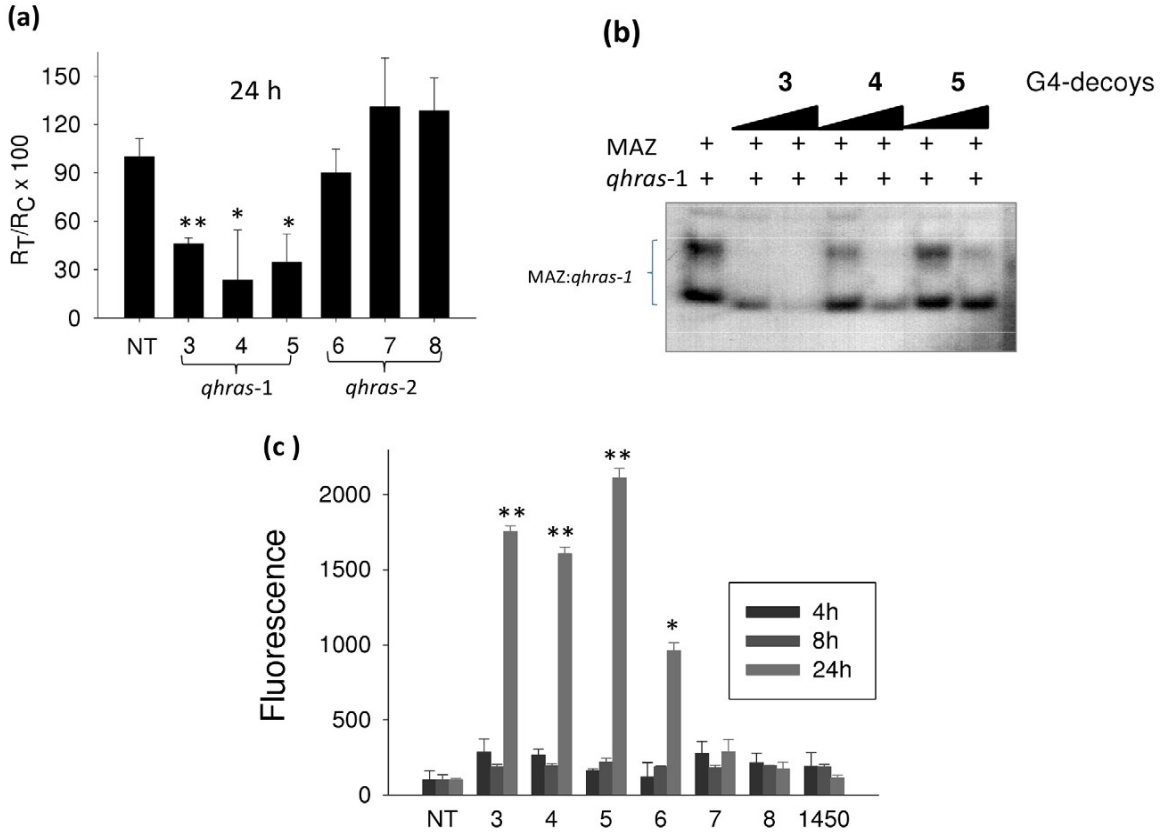


Figure 10. Effect of G4-decoys on *HRAS* transcription. (a) Real-time PCR determination of the HRAS transcripts in T24 urinary bladder cancer cells treated with 800 nM G4-decoys specific for hras-1 and hras-2 for 24 h; (b) EMSA showing that the DNA-protein complexes between 3 mg MAZ and radiolabelled qhras-1 (10 nM) are competed away by cold G4-decoys 10- and 50-fold in excess compared to radiolabelled qhras-1 quadruplex; (c) Caspase 3/7 assays showing that in T24 cells active G4-decoys 3, 4, 5 and 6 strongly activate apoptosis. Differences from the control are supported by Student's t test, P,0.05 (one asterisk), P,0.01 (two asterisks).
doi:10.1371/journal.pone.0024421.g010

It is well known that certain G-rich oligonucleotides show a clear antiproliferative effect in cancer cells which is not due to a true antisense effect, but to their propensity to fold into a G-quadruplex [26]. How these oligonucleotides precisely work, is not yet clear, but Bates and co-workers proposed that the antiprolif-erative activity of certain G-rich oligonucleotides requires: nuclease resistance; efficient cellular uptake; binding to a specific protein [53]. Our G4-decoys fulfill these requirements as: their compact structure and LNA residues make them resistant to nucleases; they efficiently penetrate cell membranes and internal-ize in the nucleus when complexed with PEI; they interact with MAZ, an essential protein for HRAS transcription.

That our G4-decoys act through their binding to a nuclear protein (MAZ) is suggested by the fact that when they are delivered without a transfectant agent, they localize in the cytoplasm and are not active. In contrast, when they are delivered with PEI, they reach the nucleus and show a strong antiprolifer-ative activity.

In accord with the proposed mechanism of action, the decoys eliciting the highest inhibition of cell growth (3, 4, 5) caused in T24 cancer cells a strong decrease of HRAS transcript and activation of caspases 3/7.

In summary, this work shows that: G4-DNA near the transcription start sites of HRAS behaves as a transcription repressor; MAZ and Sp1 bind to G-elements that can fold into quadruplex forming sequences; transcription is activated by MAZ and Sp1; MAZ destabilized the HRAS G-quadruplexes; G4-decoys mimicking the HRAS quadruplexes behave as decoy molecules against MAZ and cause a potent antiproliferative effect in T24 bladder cancer cells bearing mutant HRAS; the decoy strategy could provide a new therapeutic approach to treat bladder cancer.

Materials and Methods

Plasmids and G4-decoys synthesis

pHRAS-luc was obtained by standard cloning, while mutant plasmids pHRAS-mut1, pHRAS-mut2 were obtained by site directed mutagenesis with the gene tailor kit (Invitrogen). A 838-bp Sac I-Sac I fragment, obtained from pEJ 6.6 plasmid bearing the human HRAS promoter, was cut with Xma I restriction enzyme and the resulting 345 bp fragment was subcloned in pGL3–E1 basic plasmid in Sac I-Xma I upstream of firefly luciferase. In the resulting construct luciferase was driven by wild-type HRAS promoter (pHRAS-luc). By site directed mutagenesis we introduced in pHRAS-luc two point mutations either in hras-1 or hras-2 sequence. The primers used were 59- CGG GGC CGA GGC CGG TGC GGT GCG TGT GCG TGT GC-39 (for) and 59-CCG GCC TCG GCC CCG GCC CTG GCC C-39 (rev).

PCR was performed with 3 ng/ml DNA template, 0.1 mM each primer, 0.05 units/ml AccuPrime pfx DNA polymerase (Invitro-gen) in 16 AccuPrime pfx reaction mix for 3 min at 95uC, 30 cycles 1 min at 95uC, 30 s at 81uC, and 5 min at 68uC. Bacteria DIH 101 5a were transformed with PCR product, and plasmid DNA was extracted and sequenced (primer pGL3bpr 2 59-CTT TAT GTT TTT GGC GTC TTC CA-39).

Plasmids pCMV-MAZ (called pMAZ), pCMV-Sp1 (called pSp1), pGEX-MAZ and pGEX-Sp1 have been purchased by RIKEN (Japan).

G4-decoys were synthesized on an ExpediteTM Nucleic Acid Synthesis System Model 8909 from Applied Biosystems. Purifica-tion of oligonucleotides was accomplished using a reverse-phase semipreparative HPLC on Waters XterraTM MS C18 column. Oligonucleotide concentrations were determined by UV-absor-bance at 260 nm, 90uC and the calculated single-stranded extinction coefficients were based on a nearest neighbour model (extinction coefficient for monomers is 22400 at 260 nm).

CD and DMS footprinting experiments

CD spectra have been obtained with a JASCO J-600 spectro-polarimeter equipped with a thermostatted cell holder, 3 mM oligonucleotides in 50 mM Tris–HCl, pH 7.4, 100 mM KCl.

DMS footprinting experiments were performed using PAGE purified oligonucleotides, 27-mer hras-1 and h1-mut, 24-mer hras-2 and h2-mut (20 nM), end-labelled with [c-^{32}P]ATP (Perkin Elmer) and polynucleotide kinase (New England Biolabs, MA). Before the reactions they were incubated overnight at 37uC, in 50 mM sodium cacodylate, pH 8, 1 mg sonicated salmon sperm DNA, 1 mM EDTA, 100 mM KCl or LiCl or CsCl or KCl+TMPyP4 as specified in figure legends. Dimethylsulphate (DMS) dissolved in ethanol (DMS:ethanol, 4/1, vol/vol) was added to the DNA solution (1 ml to a total volume of 70 ml) and left to react for 5 min at room temperature. The reactions were stopped by the addition of 1:4 volumes of stop solution (1.5 M sodium acetate, pH 7, 1 M b-mercaptoethanol and 1 mg/ml tRNA). DNA was precipitated with 4 volumes of ethanol and resuspended in piperidine 1 M. After cleavage at 90uC for 30 min, reactions were stopped with chilling in ice followed by precipitation with 0.3 M sodium acetate, pH 5.2 and 3 volumes of ethanol. The DNA samples were resuspended in 90% formamide and 50 mM EDTA, denatured at 90uC and run for 2 h on a denaturing 20% polyacrylamide gel, prepared in TBE

Table 1. RT-PCR primers.

Oligonucleotides	59-39 sequence	T, 6C
hHRAS for	GGG GCA GTC GCG CCT GTG AA	60
hHRAS rev	CCG GCG CCC ACC ACC ACC AG	60
hGAPDHfor	CCC TTC ATT GAC CTC AAC TAC ATG	60
hGAPDHrev	TGG GAT TTC CAT TGA TGA CAA GC	60
hHPRTfor	AGA CTT TGC TTT CCT TGG TCA GG	60
hHPRTrev	GTC TGG CTT ATA TCC AAC ACT TCG	60
hb2mglobulinrev	CAT TCC TGA AGC TGA CAG CAT TC	60
hb2mglobulinfor	TGC TGG ATG ACG TGA GTA AAC C	60
hmazfor	CTC CAG TCC CGC TTC T	55
hmazrev	GGG AGC AAG TCC ACC T	55
hSp1for	CCC TTG AGC TTG TCC CT	50
hSp1rev	CCT GTG AAA AGG CAC CA	50

doi:10.1371/journal.pone.0024421.t001

Table 2. G4-decoys used for HRAS gene.

Oligonucleotides	59-39 sequence
hras-2 biotin	B-CGAGGCCGGGGCGGGGCGGGGGCGGGGGCGCGCGGT
hras-1 biotin	B-CGGCTCGGGTTGCGGGCGCAGGGCACGGGCGGC
hras-1 (H1)	TCGGGTTGCGGGCGCAGGGCACGGGCG
hras-2 (H2)	CGAGGCCGGGGCGGGGCGGGGGCGGGGGCGCGCGGT
3	TCPGGGTTGCGGGCGCAGGGCA CGGGCGG
3-F	F- TCPGGGTTGCGGGCGCAGGGCA CGGGCGG
4	TCPGGGTTGCGGGCGCAGGGPCA CGGGCGG
5	TCPGGGTTGCGGGCGCAGGGPCA CGGGCGG
6	CGPGGGCGGGGCGGGGGCGGGGGCG
7	CGPGGGCGGGGCGGGGGCGGGGGCG
8	CGPGGGCGGGPGCGGGGGCGGGGGCG
1450	GCGGTGTCGCPAAGACGCAAGACGCGGAGGCAG

B = biotin; P = TINA (S_4); Underlined bases indicate LNA residue. doi:10.1371/journal.pone.0024421.t002

and 8 M urea, pre-equilibrated at 55uC in a Sequi-Gen GT Nucleic Acids Electrophoresis Apparatus (Bio-Rad, CA), which was equipped with a thermocouple that allows a precise temperature control. After running, the gel was fixed in a solution containing 10% acetic acid and 10% methanol, dried at 80uC and exposed to Hyperfilm MP (GE Healthcare) for autoradiography.

Cell culture and proliferation assay

HeLa and T24 urinary bladder cancer cells were maintained in exponential growth in Dulbecco's Modified Eagle's Medium (DMEM) containing 100 U/ml penicillin, 100 mg/ml strepto-mycin, 20 mM L-glutamine and 10% foetal bovine serum (Euroclone, Milan, Italy). For proliferative assays T24 cells were seeded (1250 cells/well) the day before decoy treatment in a 96-well plate. The G4-decoys, mixed to jetPEI (Polyplus transfec-tion), were delivered to the cells at increasing concentrations up to 800 nM. 48 h after the first delivery a second dose of oligonucleotides was given to the cells. 72 h after the second delivery, the cell viability was measured by resazurin assays following standard procedures.

Dual luciferase assays

Firefly luciferase activity in cell lysates was measured and normalized for Renilla luciferase activity using the Dual-Glo Luciferase Assay System (Promega) following vendor's instruc-tions. Transfection was performed by mixing each vector 250 ng/ well with control plasmid pRL-CMV expressing Renilla luciferase under control of CMV promoter 10 ng/well using Metafectene transfection reagent (Biontex Laboratories, GmbH) following manufacturer instruction. For cotransfection with pMAZ and/or pSp1 100 ng of pHRAS-luc or mutant, with 10 ng of pRL-CMV, were transfected with either 100 ng of pMAZ/pSp1 or pcDNA3 plasmid (empty vector) as mass for control transfections. Each transfection was performed in triplicate. Luciferase assays were performed 48 h after transfection following instructions. Samples were read with Turner Luminometer and expressed as Relative luminescence (see figure captions).

Recombinant MAZ and Sp1: purification and EMSA

Recombinant MAZ and Sp1 tagged to GST were expressed in Escherichia coli BL21 DE_3 plys using plasmid pGEX-MAZ and

pGEX-Sp1. The bacteria were grown for 1–2 h at 37uC to an optical density at 600 nm between 0.5–2.0 prior to induction with IPTG (0.2 mM final concentration). Cells were allowed to grow overnight at 30uC before harvesting. The cells were centrifuged at 5000 rpm, 4uC, the supernatant removed and the cells washed twice with PBS. The pellet was resuspended in a solution of PBS with 1 mM PMSF, 1 mM DTT and Protease Inhibitor Cocktails (only for GST-tagged proteins) (Sigma-Aldrich). The bacteria were lysed by sonication and centrifuged for 30 min at 4uC, 16000 rpm. Glutathione Sepharose 4B (GE Healthcare) (50% slurry in PBS) was added to the supernatant from the previous step and incubated for 30 min at 4uC while shaking. The mix was centrifuged for 5 min at 500 g and the resin was washed 3 times with PBS. Proteins were eluted from the resin with a buffer composed by 50 mM Tris-HCl pH 8 and 10 mM reduced glutathione.

Chromatin immunoprecipitation

HeLa cells were plated in 2615 cm diameter plates, grown to 80% confluency (about 1.5610^7 cells) and fixed in formaldehyde 1% in PBS for 2 or 5 minutes. Chromatin immunoprecipitation assays were performed using the ChIP-IT™ Express kit (Active Motif, Rixensart, Belgium), as previously reported [24]. The antibodies used are MAZ H-50 (sc-28745, Santa Cruz, Biotech-nology, Inc), Sp1 (PEP 2) (sc-59, Santa Cruz, Biotechnology, Inc) used at 20 ng/ml. The primers used are: (i) for control reaction, specific for GAPDH (provided by the kit) that give a 160 bp product; (ii) 59-GGCTCCTGACAGACGGG (hras-1for) and 59-GCATGGGCTCCGTCC (hras-1rev) giving a 190 bp product; (iii) 59-GGACGGAGCCCATGC (hras-2for) and 59-CGTATTGCTGCCGCCT (hras-2rev) giving a 161 bp product. Amplification products were separated by a 10% acrylamide gel in TBE and visualized with a Gel–DOC apparatus (Bio-Rad Laboratories, CA, USA).

shRNA transfections, RNA extraction and real-time PCR

Cells were seeded 200–400000/well in a 24 well plate. shRNA plasmids were transfected the day after plating 0.5 mg/well. Plasmids used are shRNA-control, shRNA-MAZ and shRNA-Sp1 (Santa Cruz Biotechnology, CA, USA). Cells were collected 48 and 72 h after transfection.

RNA extraction: RNA was extracted using TRIzol reagent (Invitrogen, Carlsbad, CA). cDNA synthesis: 5 ml of RNA in DEPC-water was heated at 55uC and placed in ice. The solution was added with 7.5 ml of a mix containing (final concentrations) 16buffer; 0.01 M DTT (Invitrogen); 1.6 mM primer dT [MWG Biotech, Ebersberg, Germany; $d(T)_{16}$]; and 1.6 mM random primers (Promega); 0.4 mM dNTPs solution containing equi-molar amounts of dATP, dCTP, dGTP and dTTP (Euroclone, Pavia, Italy); 0.8 U/ml RNAse OUT; 8 U/ml of M-MLV reverse transcriptase (Invitrogen). The reactions were incubated for 1 h at 37uC and stopped with heating at 95uC for 5 min. As a negative control the reverse transcription reaction was per-formed with a sample containing DEPC water. Real-time PCR reactions were performed with 16 iQ™ SYBR Green Super-mix (Bio-Rad Laboratories, CA, USA), 300 nM of each primer, 1 ml of RT reaction. The sequences of the primers used for HRAS, GAPDH, b-2-microglobulin, Hypoxanthine Ribosyl Transferase, MAZ and Sp1 amplifications are reported in Table 1. The PCR cycle was: 3 min at 95uC, 40 cycles 10 s at 95uC, 30 s at 60uC with an iQ_5 real-time PCR controlled by an Optical System software version 2.0 (Bio-Rad Laboratories, CA, USA).

Electrophoresis mobility-shift assays

Oligonucleotides were end-labelled with [c-33P]ATP and T4 polynucleotide kinase. Duplexes hras-1 and hras-2 were prepared annealing (10 min at 95uC, overnight at room temperature) a mixture containing 1:1,2 ratio of radiolabelled hras-1 or hras-2 with the complementary strand in 50 mM Tris–HCl, pH 7.4, 100 mM NaCl. Radiolabelled hras-1 and hras-2 were allowed to assume a G4-DNA structure in 50 mM Tris–HCl, pH 7.4, 100 mM KCl, by heating at 95uC and overnight incubation at 37uC. Before EMSA, the radiolabelled oligonucleotides were treated for 30 min at room temperature with different amounts of MAZ or Sp1 (or extract), in 20 mM Tris–HCl, pH 8, 30 mM KCl, 1.5 mM $MgCl_2$, 1 mM DTT, 8% glycerol, 1% Phosphatase Inhibitor Cocktail I (Sigma, Milan, Italy), 5 mM NaF, 1 mM Na_3VO_4, 2.5 ng/ml poly [dI-dC]. After incubation, the reaction mixtures were loaded in 5% TBE (1X) polyacrylamide gel, thermostatted at 20uC. After running, the gel was dried and exposed to autoradiography (GE Healthcare, Milan) for 24–36 h at 280uC.

FRET-melting experiments

FRET melting experiments were performed on a real-time PCR apparatus (CFX 96, BioRad, Hercules, CA), using a 96-well plate filled with 50 ml solutions of dual-labelled qhras-1 and qhras-2, in 50 mM Tris–HCl, pH 7.4, 50 mM KCl (qhras-1) or 100 mM NaCl (qhras-2). The protocol, used for the melting experiments, is the following: (i) equilibration step of 5 min at low temperature (20uC); (ii) stepwise increase of the temperature of 1uC/min for 76 cycles to reach 95uC. All the samples in the wells were melted in 76 min.

Caspase assays

We performed Caspase activity assays using Apo-ONE™ Homogeneous Caspase-3/7 Assay (Promega), according to the manufacturer's protocol using a Microplate Spectrofluorometer System (Molecular Devices, Concorde, Canada).

Supporting Information

Figure S1 Polymerase stop assays with a wild-type DNA template containing hras-1 or hras-2 and a mutant template in which four GRT point mutations were introduced in sequence hras-2 to abrogate quadruplex formation. Taq polymerase is arrested at the 39 end of hras-2, before the first run of guanines, in keeping with the formation of a quadruplex structure by hras-2 (experimental conditions: 37uC, 140 mM KCl, 50 mM Tris-HCl pH 7.4) When the DNA template is incubated with G4-DNA ligands that stabilize quadruplex DNA, Taq polymerase is completely arrested and only the truncated product is produced. This is observed with porphyrin $TMPyP_4$ and guanidine phthalocyanines DIGP and Zn-DIGP at r = 4 (r = [ligand]/[template]). Instead, $TMPyP_2$, which does not bind to G4-DNA does not affect the processivity of Taq polymerase. When the experiment is performed with the mutated template, Taq polymerase does not stop at the G-element and full product is observed. A longer truncated product is observed with the mutated template, probably due to a hairpin structure stabilized by CG and GT base pairs. Polymerase stop assays with a DNA template containing hras-1show that Taq polymerase is arrested in the presence of phthalocyanines, indicating that the G-quadruplex formed by hras-1 is less stable than that formed by hras-2.(TIF)

Figure S2 CD spectra of the hras-1 and hras-2 mutants. CD spectra in 50 mM Tris-HCl pH 7.4, 100 mM KCl of hras-1, hras-1

mut (59-TCGGGTTGCGGGCGCAGGGCACCTGCG), hras-2 and hras-2 mut (59-CGAGGCCGGTGCGGTGCGGGGGC-GGGGGCGCGCGGT). Cuvette 0.5 cm, DNA concentration 6 mM.(TIF)

Figure S3 EMSA showing the binding between recombinant MAZ-GST and Sp1-GST with duplexes hras-1 and hras-2. (a) radiolabelled hras-1 duplex, 15 nM incubated for 30 min with 0, 0.5, 1, 1.5, 2 and 2.5 mg MAZ-GST; (b) radiolabelled hras-2 duplex, 15 nM incubated for 30 min with 0, 0.5, 1, 1.5, 2 and 2.5 mg MAZ-GST; (c) radiolabelled hras-2 duplex, 15 nM incubated for 30 min with 0, 0.5, 1 and 2.5 mg Sp1-GST.(TIF)

Figure S4 SDS-PAGE of eluate from Glutathione Sepharose 4B column loaded with protein extract of BL21 bacteria transformed with a plasmid encoding for MAZ-GST or Sp1-GST (a) lane 1: bacterial extract; lane 2: proteins that did not bind the resin (flow through); lane 3, fraction eluted with 10 mM glutathione; (b) lane 1: molecular weights; lane 2: bacterial extract; lane 3: proteins that did not bind the resin (flow through); lane 4, column wash; lane 5: 1st elution with 10 mM glutathione; lane 6: 2nd elution with 10 mM glutathione.(TIF)

Figure S5 Silencing in HeLa cells of MAZ and Sp1 by commercial shRNAs. HeLa cells have been treated with MAZ shRNA or Sp1 shRNA complexed with Metafectene. As control we treated HeLa cells with control shRNA. MAZ and Sp1 specific shRNA and control shRNA have been purchased from Santa Cruz Biotechnology (USA). After 48 h, total RNA was extracted, transformed in cDNA and used for real-time experiments. The levels of MAZ and Sp1 transcripts have been measured and reported in graph relatively to the expression of three housekeep-ing genes: GAPDH, b2- microglobulin, hypoxanthine ribosyl transferase.(TIF)

Figure S6 (a) SDS-PAGE of fractions eluted from a Glutathione Sepharose 4B column loaded with protein extract obtained from non-transformed BL21 DE3 plysS bacteria (lanes 5 and 6). Fractions eluted with 10 mM reduced glutathione do not contain bacterial proteins bound non-specifically to the resin; (b) (top) FRET melting of 200 nM quadruplex hras-1 in 50 mM Tris-HCl, pH 7.4, 50 mM KCl, 50 mM Zn-acetate in the presence of FPLC purified GST at DNA:protein ratios of 1:0, 1:1 and 1:5; (bottom) FRET melting of 200 nM quadruplex hras-2 in 50 mM Tris-HCl, pH 7.4, 100 mM NaCl, 50 mM Zn-acetate in the presence of FPLC purified GST at DNA:protein ratios of 1:0, 1:1 and 1:5. (TIF)

Figure S7 (Top) putative structures of the designed G4-decoys. The yellow rectangles represent the TINA unit (P); (bottom) Structure of the TINA unit covalently inserted in the decoy oligonucleotides.(TIF)

Figure S8 Proliferation assay with T24 cells untreated and treated with 800 nM G4-decoys 3,4,5 (mimicking qhras-1) and G4-decoys 6,7,8 (mimicking qhras-2). Decoy 637 is a random sequence containing one P unit. Two doses, one 48 h after the other, of 800 nM G4-decoys mixed with polyethylenimine have been delivered to T24 cells. Viable cells, measured by a resazurin assay, have been performed at increasing times from 1st treatment. (TIF)

Acknowledgments

Plasmids pGEX-MAZ, pGEX-Sp1, pCMV-MAZ and pCMV-Sp1 have been provided by RIKEN BRC which is participating in the National Bio-Resources Project of the MEXT. We thank Prof. Nathan Luedtke (Zurich University) for the guanidinium phthalocyanines and Dr. Manikandan Paramasivam for DMS-footprinting experiment assistance.

Author Contributions

Conceived and designed the experiments: LX AM SC EP. Performed the experiments: AM SC. Analyzed the data: LX. Contributed reagents/ materials/analysis tools: LX EP. Wrote the paper: LX.

References

14. Lowy DR, Willumsen BM (1993) Function and regulation of RAS. Annu Rev Biochem 62: 851–891.

15. Clarke S (1992) Protein isoprenylation and methylation at carboxyl terminal cysteine residues. Annu Rev Biochem 61: 355–386.

16. Rebollo A, Martinez CA (1999) Ras proteins: recent advances and new functions. Blood 94: 2971–2980.

17. Downword J (1998) Ras signalling and apoptosis. Curr Opin Genet Dev 8: 49–54.

18. Schlessinger J (1993) How receptor tyrosin kinases activate Ras. Trends Biol Sci 18: 273–275.

19. Porter AC, Vaillancourt RR (1998) Tyrosine kinase receptor-activated signal transduction pathways which lead to oncogeneses. Oncogene 17: 13434–13452.

20. Fujita J, Yoshida O, Yuasa Y, Rhim JS, Hatanak M, et al. (1984) Ha-ras oncogenes are activated by somatic alterations in human urinary tract tumors. Nature 309: 464–466.

21. Bos JL (1989) Ras oncogenes in human cancer: a review. Cancer Res 49: 4682–4689.

22. Schubbert S, Shannon K, Bollag G (2007) Hyperactive Ras in developmental disorders and cancer. Nat Rev Cancer 7: 295–308.

23. Dinney CP, McConkey DJ, Millikan RE, Wu X, Bar-Eli M, Adam L, et al. (2004) Focus on bladder cancer. Cancer Cell 6: 111–116.

24. Vageli D, Kiaris H, Delakas D, Anezinis P, Cranidis A, et al. (1996) Transcriptional activation of H-ras, K-ras and N-ras proto-oncogenes in human bladder tumors. Cancer Lett 107: 241–247.

25. Mo L, Zheng X, Huang HY, Shapiro E, Lepor H, et al. (2007) Hyperactivation of Ha-ras oncogene, but not Ink4a/Arf deficiency, triggers bladder tumorigen-esis. J Clin Invest 117: 314–325.

26. Ishii S, Merlino GT, Pastan I (1985) Promoter region of the human Harvey ras proto-oncogene: similarity to the EGF receptor proto-oncogene promoter. Science 230: 1378–1381.

33. Ishii S, Kadonaga JT, Tjian R, Brady JN, Merlino GT, et al. (1986) Binding of the Sp1 transcription factor by the human Harvey ras1 proto-oncogene promoter. Science 232: 1410–1413.

34. Sen D, Gilbert W (1988) Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. Nature 334: 364–366.

35. Burge S, Parkinson GN, Hazel P, Todd AK, Neidle S (2006) Quadruplex DNA: sequence, topology and structure. Nucleic Acids Res 34: 5402–5415.

36. Eddy J, Maizels N (2006) Gene function correlates with potential for G4 DNA formation in the human genome. Nucleic Acids Res 34: 3887–3896.

37. Huppert JL, Balasubramanian S (2007) G-quadruplexes in promoters through-out the human genome. Nucleic Acids Res 35: 406–413.

38. Saxonov S, Berg P, Brutlag DL (2006) A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc Natl Acad Sci USA 103: 1412–1417.

39. Zhao Y, Du Z, Li N (2007) Extensive selection for the enrichment of G4 DNA motifs in transcriptional regulatory regions of warm blooded animals. FEBS Lett 581: 1951–1956.

40. Siddiqui-Jain A, Grand CL, Bears DJ, Hurley LH (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. Proc Natl Acad Sci U S A 99: 11593–11598.

41. Brooks TA, Hurley LH (2009) The role of supercoiling in transcription control of CMYC and its importance in molecular therapeutics. Nature Reviews Cancer 9: 849–861.

42. Cogoi S, Xodo LE (2006) G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription. Nucleic Acids Res 34: 2536–2549.

43. Cogoi S, Paramasivam M, Membrino A, Yokoyama KK, Xodo LE (2010) The KRAS promoter responds to MYC-associated zinc finger and poly[ADP-ribose]polymerase 1 proteins which recognize a critical quadruplex-forming GA-element. J Biol Chem 285: 22003–22016.

25. Bejugam M, Sewitz S, Shirude PS, Rodriguez R, Shahid R, et al. (2007) Trisubstituted isoalloxazines as a new class of G-quadruplex binding ligands: small molecule regulation of c-kit oncogene expression. J Am Chem Soc 129: 12926–12927.

26. Choi EW, Nayak LV, Bates PJ (2010) Cancer-selective antiproliferative activity is a general property of some G-rich oligodeoxynucleotides. Nucleic Acids Res
38: 1623–35.

27. Bates PJ, Kahlon JB, Thomas SD, Trent JO, Miller DM (1999) Antiproliferative activity of G-rich oligonucleotides correlates with protein binding. J Biol Chem
274: 26369–26377.

28. Todd AK, Neidle S (2008) The relationship of potential G-quadruplex sequences in cis-upstream regions of the human genome to SP1-binding elements. Nucleic Acids Res 36: 2700–4.

29. Han H, Langley DR, Rangan A, Hurley LH (2001) Selective interactions of cationic porphyrins with G-quadruplex structures. J Am Chem Soc 123: 8902–8913.

30. Paramasivan S, Rujan I, Bolton PH (2007) Circular dichroism of quadruplex DNAs: applications to structure, cation effects and ligand binding. Methods 43: 324–31.

31. Membrino A, Paramasivam M, Cogoi S, Alzeer J, Luedtke NW, et al. (2010) Cellular uptake and binding of guanidine-modified phthalocyanines to KRAS/ HRAS G-quadruplexes. Chem Commun (Camb) 46: 625–627.

32. Alzeer J, Roth PJ, Luedtke NW (2009) An efficient two-step synthesis of metal-free phthalocyanines using a Zn(ii) template. Chem Commun (Camb) 21: 1970–1971.

33. Alzeer J, Vummidi BR, Roth PJ, Luedtke NW (2009) Guanidinium-modified phthalocyanines as high-affinity G-quadruplex fluorescent probes and transcrip-tional regulators. Angew Chem Int Ed Engl 48: 9362–9365.

34. Han H, Langley DR, Rangan A, Hurley LH (2001) Selective interactions of cationic porphyrins with G-quadruplex structures. J Am Chem Soc 123: 8902–8913.

35. Parks CL, Shenk T (1996) The serotonin 1a receptor gene contains a TATA-less promoter that responds to MAZ and Sp1. J Biol Chem 271: 4417–4430.

36. Parks CL, Shenk T (1997) Activation of the adenovirus major late promoter by transcription factors MAZ and Sp1. J Virol 71: 9600–9607.

37. Leroy C, Manen D, Rizzoli R, Lombe`s M, Silve C (2004) Functional importance of Myc-associated zinc finger protein for the human parathyroid hormone (PTH)/PTH-related peptide receptor-1 P2 promoter constitutive activity. J Mol Endocrinol 32: 99–113.

38. Song J, Ugai H, Kanazawa I, Sun K, Yokoyama KK (2001) Independent repression of a GC-rich housekeeping gene by Sp1 and MAZ involves the same cis-elements. J Biol Chem 276: 19897–19904.

39. Palumbo SL, Memmott RM, Uribe DJ, Krotova-Khan Y, Hurley LH, et al. (2008) A novel G-quadruplex-forming GGA repeat region in the c-myb promoter is a critical regulator of promoter activity. Nucleic Acids Res 36: 1755–1769.

40. Weisman-Shomer P, Naot Y, Fry M (2000) Tetrahelical forms of the fragile X syndrome expanded sequence d(CGG)(n) are destabilized by two heterogeneous nuclear ribonucleoprotein-related telomeric DNA-binding proteins. J Biol Chem 275: 2231–2238.

41. Weisman-Shomer P, Cohen E, Fry M (2002) Distinct domains in the CArG-box binding factor A destabilize tetraplex forms of the fragile X expanded sequence d(CGG)n. Nucleic Acids Res 30: 3672–3681.

42. Cogoi S, Paramasivam M, Filichev V, Ge´ci I, Pedersen EB, et al. (2009) Identification of a new G-quadruplex motif in the KRAS promoter and design of pyrene-modified G4-decoys with antiproliferative activity in pancreatic cancer cells. J Med Chem 52: 564–8.

43. Dai J, Dexheimer TS, Chen D, Carver M, Ambrus A, et al. (2006) An intramolecular G-quadruplex structure with mixed parallel/antiparallel G-strands formed in the human BCL-2 promoter region in solution. J Am Chem Soc 128: 1096–1098.

44. Reddy EP, Reynolds RK, Santos E, Barbacid M (1982) A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene. Nature 300: 149–152.

45. Komatsu M, Li HO, Tsutsui H, Itakura K, Matsumura M, et al. (1997) MAZ, a Myc-associated zinc finger protein, is essential for the ME1a1-mediated expression of the c-myc gene during neuroectodermal differentiation of P19 cells. Oncogene 15: 1123–1131.

46. Gonzales V, Guo K, Hurley L, Sun D (2009) Identification and characterization of nucleolin as a c-myc G-quadruplex-binding protein. J Biol Chem 284: 23622–23635.

47. Federici L, Arcovito A, Scaglione GL, Scaloni F, LoSterzo C, et al. (2010) Nucleophosmin C-terminal leukemia-associated domain interacts with G-rich quadruplex forming DNA. J Biol Chem 285: 37138–37148.

48. Bossone SA, Asselin C, Patel AJ, Marcu KB (1992) MAZ, a zinc finger protein, binds to c-MYC and C2 gene sequences regulating transcriptional initiation and termination. Proc Natl Acad Sci U S A 89: 7452–7466.

49. Her S, Claycomb R, Tai TC, Wong DL (2003) Regulation of the rat phenylethanolamine N-methyltransferase gene by transcription factors Sp1 and MAZ. Molecular Pharmacology 64: 1180–1188.

50. Lew A, Rutter WJ, Kennedy GC (2000) Unusual DNA structure of the diabetes susceptibility locus IDDM2 and its effect on transcription by the insulin promoter factor Pur-1/MAZ. Proc Natl Acad Sci U S A 97: 12508–12512.

51. Himeda CL, Ranish JA, Hauschka SD (2008) Quantitative proteomic identification of MAZ as a transcriptional regulator of muscle-specific genes in skeletal and cardiac myocytes. Mol Cell Biol 28: 6521–3655.

52. Karantzoulis-Fegaras F, Antoniou H, Lai SL, Kulkarni G, D'Abreo C, et al. (1999) Characterization of the human endothelial nitric-oxide synthase promoter. J Biol Chem 274: 3076–3093.

53. Bates PJ, Laber DA, Miller DM, Thomas SD, Trent JO (2009) Discovery and development of the G-rich oligonucleotide AS1411 as a novel treatment for cancer. Exp Mol Pathol 86: 151–164

Supporting Information

G4-DNA formation in *HRAS* promoter and rationale design of decoy oligonucleotides for cancer therapy

Alexandro Membrino[1], Susanna Cogoi[1], Erik B. Pedersen[2] and Luigi E. Xodo[1]*

[1]Department of Medical and Biological Science, P.le Kolbe 4, School of Medicine, 33100 Udine, Italy; Tel (+39) 0432.494395; Fax (+39) 0432.494301; [2]Nucleic Acid Center, Institute of Physics and Chemistry, University of Southern Denmark, DK-5230 Odense M, Denmark.

Corresponding author, E-mail: luigi.xodo@uniud.it

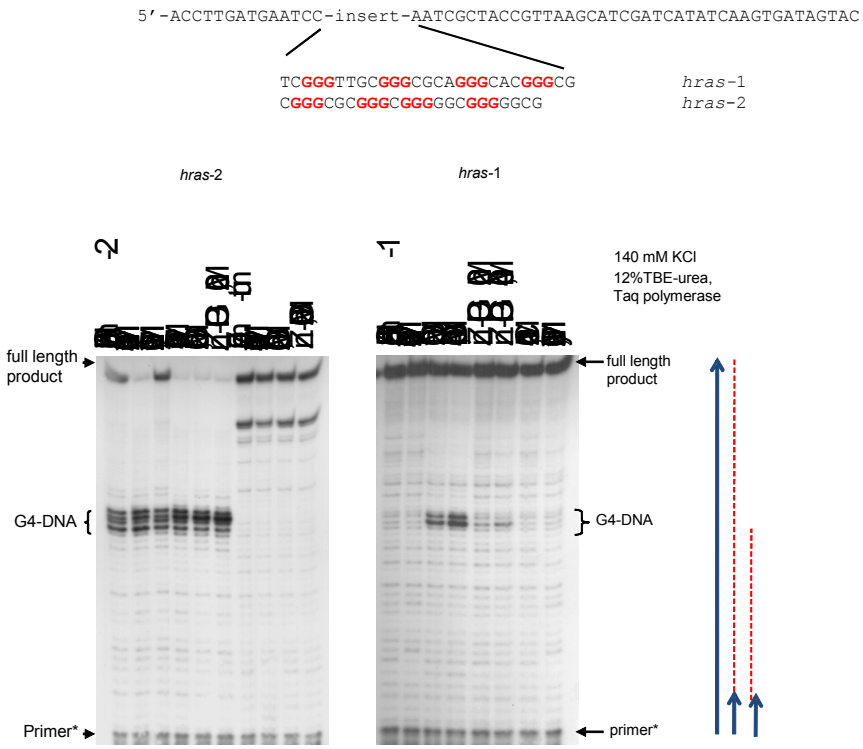


Fig. S$_1$: Polymerase stop assay. The formation of G-quadruplexes by *hras*-1 and *hras*-2 was assessed by polymerase stop assays. We designed two wild-type DNA templates containing

hras-1 or *hras*-2 and one mutant template in which four G→T point mutations were introduced in *hras*-2, to abrogate quadruplex formation. Primer-extension reactions showed that Taq polymerase was partly arrested at the 3' end of *hras*-2, just before the first run of guanines, in keeping with the formation of a quadruplex structure by *hras*-2. Under the experimental condition adopted, 37°C, 140 mM KCl, the polymerase was not completely arrested, as the full-length product was present in lane 1. However, when the DNA template was incubated with G4-DNA ligands that stabilize quadruplex DNA, Taq polymerase was completely arrested and only the truncated product was produced. This was observed with porphyrin $TMPyP_4$ and guanidine phthalocyanines DIGP and Zn-DIGP at r=4 (r=[ligand]/[template]). Instead, $TMPyP_2$, which does not bind to G4-DNA did not affect the processivity of Taq polymerase. When the experiment was performed with the mutated template whose *hras*-2 element cannot fold into a G-quadruplex (the second guanine in each block was replaced by T), Taq polymerase did not stop at the G-element and full product is observed. A longer truncated product is observed with the mutated template, due to a hairpin structure stabilized by CG and GT base pairs. Polymerase stop assays were performed also with a DNA template containing *hras*-1. In this case we observed arrest of Taq polymerase in the presence of phthalocyanines, indicating that the G-quadruplex formed by *hras*-1 was less stable than that formed by *hras*-2.

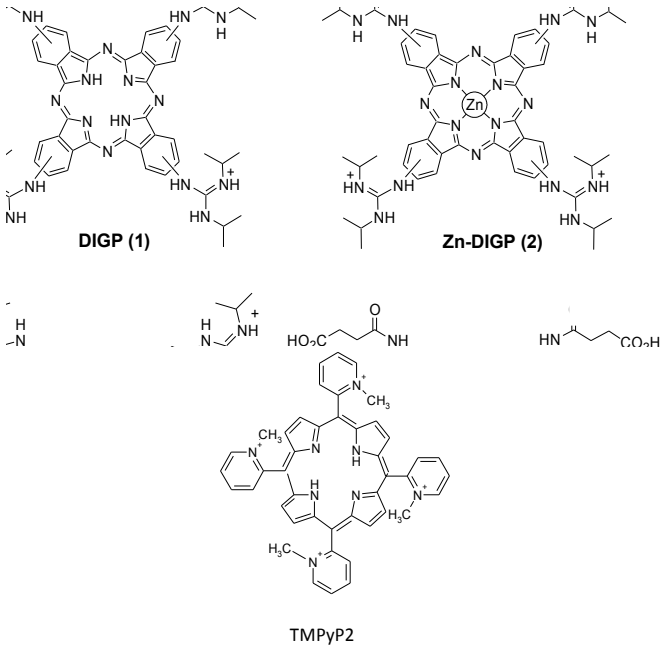


Fig. S_2: Structure of guanidine phthalocyanines DIGP, Zn-DIGP and porphyrin TMPyP2 used in dual luciferase assays

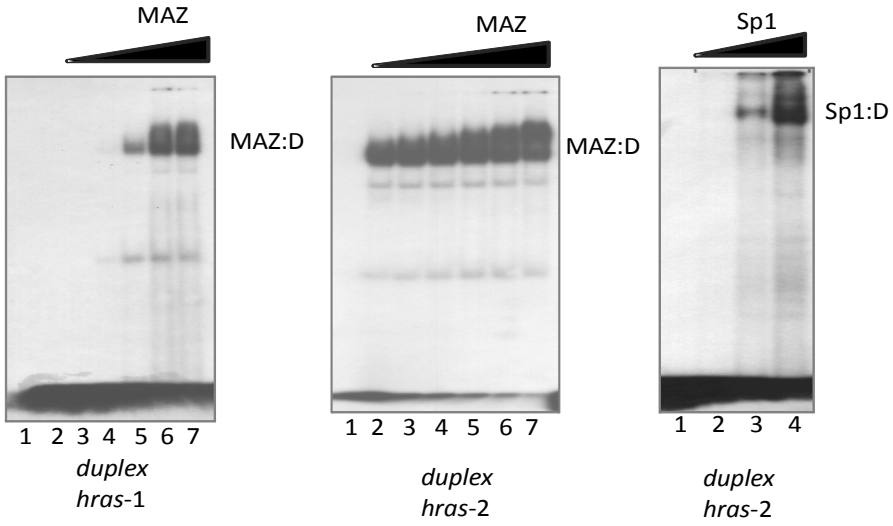


Fig. S_3: EMSA showing the binding between recombinant MAZ-GST and Sp1-GST with duplexes *hras*-1 and *hras*-2.

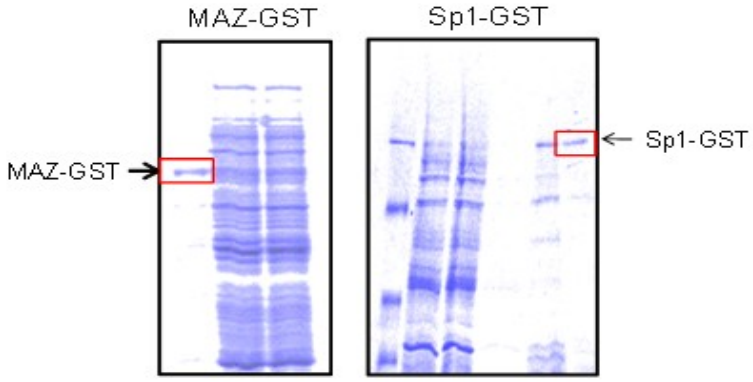


Fig. S_4: Purification of MAZ-GST and Sp1-GST by glutathione sepharose 4B affinity chromatography

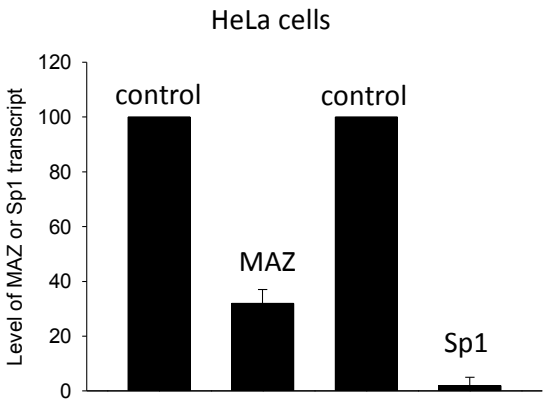


Fig. S_5: Silencing in HeLa cells of MAZ and Sp1 by commercial shRNAs. HeLa cells have been treated with MAZ shRNA or Sp1 shRNA complexed with Metafectene. After 48 h, total RNA was extracted, transformed in cDNA and used for real-time experiments. The levels of

MAZ and Sp1 transcripts have been measured and reported in graph relatively to the expression of three housekeeping genes: *GAPDH*, beta-2 microglobulin, hypoxanthine ribosyl-transferase.

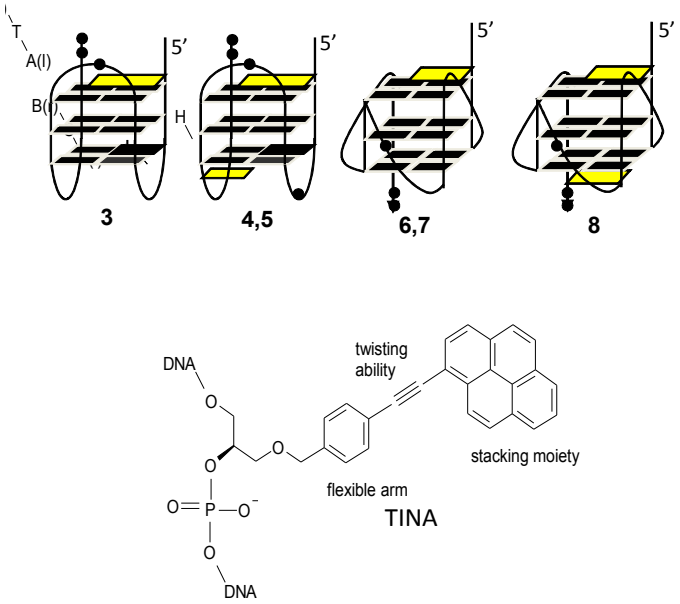


Fig. S$_6$: putative structures of the designed G4-decoys. The yellow rectangles represent the TINA unit (P).

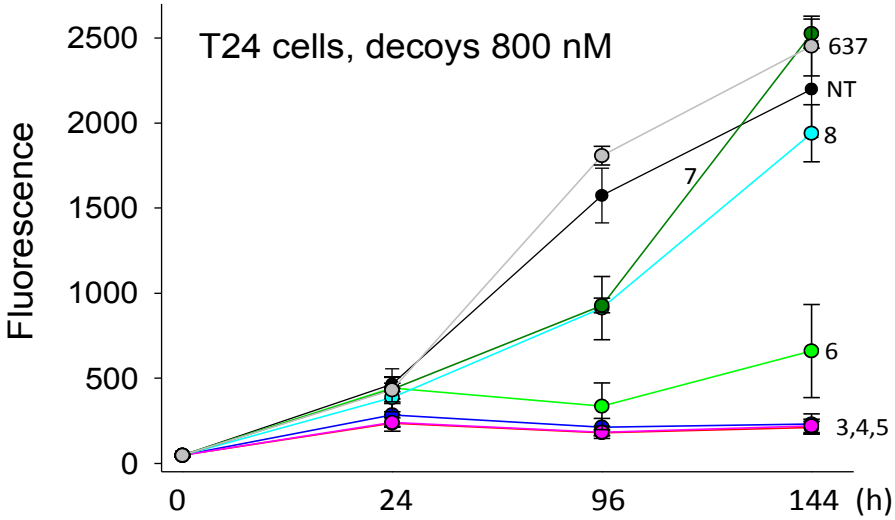


Fig. S$_7$: Proliferation assay with T24 cells untreated and treated with 800 nM G4-decoys **3**,**4**,**5** (mimicking *qhras*-1) and 6,7,8 (mimicking *qhras*-2). Decoy 637 is a random sequence containing one P unit.

5. Conclusions

The main target of my PhD was to demonstrate that quadruplex DNA has a function on transcription. Most laboratories have addressed their efforts to understand why the 5' end of mammalian genes are rich in guanine, while the 3' end is not. The asymmetric distribution of the guanines in the genome has raised the hypothesis that quadruplex DNA should be involved in gene expression regulation either at the promoter or at the transcript level. This PhD work has addressed the query whether quadruplex DNA is an integral part of the mechanism controlling transcription. To provide an answer we took into consideration the *HRAS* gene, as it possesses upstream of TSS two G-elements composed by runs of guanines each folding into a quadruplex structure. Furthermore, *HRAS* was a good study model because the quadruplex-forming sequences overlap the binding site of transcription factors MAZ and Sp1. The mechanism for transcription regulation that we have proposed is based on the distinctive observation that MAZ recognizes the *HRAS* quadruplexes and unfolds them. So MAZ has a double function within the *HRAS* promoter: a. unfolds the quadruplexes and remove the block on transcription and b. activates transcription together with Sp1. Likewise, we can say that quadruplex DNA also has two functions: a. inhibits transcription and b. recruits MAZ to the *HRAS* promoter when transcription has to be activated.

The molecular strategies that have been proposed to down-regulate transcription in cancer cells may have a potential in cancer therapy. We developed oligonucleotides decoys engineered to be stable in serum. These molecules showed strong antiproliferative effect in cancer at a very low concentration. Contrarily to other G-rich oligonucleotides whose antiproliferative effect has been correlated with their capacity to bind to nucleolin, the quadruplex decoys proposed in this PhD work behave with a mechanism based on their binding to the transcription factors MAZ and Sp1. The good performance of these molecules in cancer cell lines allowed us to continue this project by testing the activity of the decoys *in vivo*. A problem that should be addressed in order to enhance the potency of the decoy approach is to enhance their cellular uptake using nanotechnologies.

The second approach consisted in small molecules that easily penetrate the cell membrane and bind to the quadruplex. The promoter will be locked in the folded, non active form, inhibiting gene expression. Guanidinio phthalocyanines are quadruplex ligands that nicely stabilize quadruplex DNA but have a poor access to the nucleus. This strategy will be improved when ligands that efficiently penetrate the nucleus will be available.

6. Bibliography

- ✓ Alzeer J, Roth PJC, et al. **An efficient two-step synthesis of metal-free phtalo-cyanines using a Zn(II) template.** ChemComm 2009, 1970-1971
- ✓ Bar-Sagi D **A Ras by any other name.** Molecular and Cellular Biology 2001, 21, 1441-1443
- ✓ Bates PJ, Kahlon JB, et al. **Antiproliferative activity of G-rich oligonucleotides correlates with protein binding.** The Journal of Biological chemistry 1999, 274, 26369-26377
- ✓ Bates PJ, Mergny JL, et al. **Quartets in G-major.** EMBO reports 2007, 8, 1003-1009
- ✓ Bates PJ, Laber DA, et al **Discovery and development of the G-rich oligonucleotides AS1411 as a novel treatment for cancer.** Exp. Mol. Pathol. 2009, 86, 151-164
- ✓ Bejugam M, Sewitz S, et al. **Trisubstituted isoalloxazines as a new class of G-quadruplex binding ligands: small molecule regulation of c-kit oncogene expression.** J. Am. Chem. Soc. 2007, 129, 12926-12927
- ✓ Bos JL ***ras* oncogenes in human cancer: a review.** Cancer research 1989, 49, 4682-4689
- ✓ Bos JL **All in the family? New insights and questions regarding interconnectivity of Ras, Rap1 and Ral.** The EMBO Journal 1998, 17, 6776-6782
- ✓ Brooks TA, Hurley LH **The role of supercoiling in transcription control of CMYC and its importance in molecular therapeutics.** Nature Reviews Cancer 2009, 9, 849-861
- ✓ Burge S, Parkinson GN, et al. **Quadruplex DNA: sequence, topology and structure.** Nucleic Acids Research 2006, 34, 5402-5415
- ✓ Burmer GC, Loeb LA **Mutations in the *KRAS2* oncogene during progressive stages of human colon carcinoma.** Proc. Natl. Acad. Sci. USA 1989, 86, 2403-2407
- ✓ Campbell SL, Khosravi-Far R, et al. **Increasing complexity of Ras signaling.** Oncogene 1998, 17, 1395-1413
- ✓ Choi EW, Nayak LV, et al **Cancer-selective antiproliferative activity is a general property of some G-rich oligodeoxynucleotides.** Nucleic Acids Research 2010, 38, 1623-1635
- ✓ Cogoi S, Quadrifoglio F, et al **G-rich oligonucleotide inhibits the binding of a nuclear protein to the Ki-ras promoter and strongly reduces cell growth in human carcinoma pancreatic cells.** Biochemistry 2004, 43, 2512-2523
- ✓ Cogoi S, Xodo LE **G-quadruplex formation within the promoter of the *KRAS* proto-oncogene and its effect on transcription.** Nucleic Acids Research 2006, 34, 2536-2549
- ✓ Cogoi S, Paramasivam M, et al. **Structural polymorphism within a regulatory element of the human *KRAS* promoter: formation of G4-DNA recognized by nuclear proteins.** Nucleic Acids Research 2008, 36, 3765-3780
- ✓ Cogoi S, Paramasivam M, et al. **The KRAS promoter responds to MYC-associated zinc finger and poly[ADP-ribose]polymerase 1 proteins which recognize a critical quadruplex-forming GA-element.** J Biol Chem. 2010, 285, 22003-22016

- ✓ De A, Michor F **DNA secondary structures and epigenetic determinants of cancer genome evolution.** Nature structural & Molecular Biology 2011, 18, 950-956
- ✓ Dinney CP, McConkey DJ, et al. **Focus on bladder cancer.** Cancer Cell 2004, 6, 111-116
- ✓ Downward J **Ras signalling and apoptosis.** Current Opinion in Genetics & Development 1998, 8, 49-54
- ✓ Dunn KL, Espino PS, et al. **The Ras-MAPK signal transduction pathway, cancer and chromatin remodeling.** Biochem. Cell. Biol. 2005, 83, 1-14
- ✓ Eddy J, Maizels N **Gene function correlates with potential for G4 DNA formation in the human genome.** Nucleic Acid Research 2006, 34, 3887-3896
- ✓ Enokizono Y, Konishi Y, et al. **Structure of hnRNP D complexed with single-stranded telomere DNA and unfolding of the Quadruplex by heterogeneous nuclear ribonucleoprotein D.** The Journal of Biological Chemistry 2005, 280, 18862-18870
- ✓ Federici L, Arcovito A, et al. **Nucleophosmin C-terminal leukemia-associated domain interacts with G-rich quadruplex forming DNA.** The Journal of Biological Chemistry 2010, 285, 37138-37149
- ✓ Gellert M, Lipsett MN, et al. **Helix formation by guanylic acid.** Proc. N.A.S. 1962, 48, 2013-2018
- ✓ Guo K, Gokhale V, et al. **Intramolecularly folded G-quadruplex and i-motif structures in the proximal promoter of the vascular endothelial growth factor gene.** Nucleic Acids Research 2008, 36, 4598-4608
- ✓ Han H, Hurley LH **G-quadruplex DNA: a potential target for anti-cancer drug design.** TiPS 2000, 21, 136-142
- ✓ Han H, Langley DR, et al. **Selective interactions of cationic porphyrins with G-quadruplex structures.** J Am Chem Soc. 2001, 123, 8902-13
- ✓ Herrick G, Alberts B **Purification and physical characterization of nucleic acid helix-unwinding proteins from calf thymus.** The Journal of Biological Chemistry 1976, 251, 2124-2132
- ✓ Hershman SG, Chen Q, et al. **Genomic distribution and functional analyses of potential G-quadruplex-forming sequences in *Saccharomyces cerevisiae*.** Nucleic Acids Research 2008, 36, 144-156
- ✓ Hoffman EK, Trusko SP, et al. **An S1 nuclease-sensitive homopurine/homo-pyrimidine domain in the c-Ki-ras promoter interacts with a nuclear factor.** Biochemistry 1990, 87, 2705-2709
- ✓ Honkawa H, Masahashi W, et al. **Identification of the principal promoter sequence of the c-H-ras transforming oncogene: deletion analysis of the 5'-flanking region by focus formation assay.** Molecular and Cellular Biology, 1987, 7, 2933-2940
- ✓ Huppert JL, Balasubramanian S **Prevalence of quadruplexes in the human genome.** Nucleic Acids Research 2005, 33, 2908-2916
- ✓ Huppert JL, Balasubramanian S **G-quadruplexes in promoters throughout the human genome.** Nucleic Acids Research 2007, 35,406-413
- ✓ Hurley LH **Secondary DNA structures as molecular targets for cancer therapeutics.** Biochemical Society 2001, 29, 692-696

- ✓ Ishii S, Merlino GT, et al. **Promoter region of the human Harvey ras proto-oncogene: similarity to the EGF receptor proto-oncogene promoter.** Science 1985, 230, 1378-1381
- ✓ Ishii S, Kadonaga JT, et al. **Binding of the Sp1 transcription factor by the human Harvey ras1 proto-oncogene promoter.** Science 1986, 232, 1410-1413
- ✓ Jantos K, Rodriguez R, et al. **Oxazole-based peptide macrocycles: a new class of G-quadruplex binding ligands.** J. Am. Chem. Soc. 2006, 128, 13662-13663
- ✓ Jing N, Li Y, et al. **G-quartet oligonucleotides: a new class of signal transducer and activator of transcription 3 inhibitors that suppresses growth of prostate and breast tumors through induction of apoptosis.** Cancer Research 2004, 64, 6603-6609
- ✓ Jordano J, Perucho M **Chromatin structure of the promoter region of the human c-K-ras gene.** Nucleic Acids Research 1986, 14, 7361-7378
- ✓ Kankia BI, Barany G, et al. **Unfolding of DNA quadruplexes induced by HIV-1 nucleocapsid protein.** Nucleic Acids Research 2005, 33, 4395-4403
- ✓ Keniry MA **Quadruplex structures in nucleic acids.** Biopolymers 2001, 56, 123-146
- ✓ Kumari S, Bugaut A, et al. **A RNA G-quadruplex in the 5'-UTR of the *NRAS* proto-oncogene modulates translation.** Nat. Chem. Biol. 2007, 3, 218-221
- ✓ Lee JT, McCubrey JA **The Raf/MEK/ERK signal transduction cascade as a target for chemotherapeutic intervention in leukemia.** Leukemia 2002, 16, 486-507
- ✓ Leroy C, Manen D, et al. **Functional importance of Myc-associated zinc finger protein for the human parathyroid hormone (PTH)/PTH-related peptide receptor-1 P2 promoter constitutive activity.** J Mol Endocrinol. 2004, 32, 99-113
- ✓ Lipps HJ, Rhodes D **G-quadruplex structures: *in vivo* evidence and function.** Cell 2009, 19, 414-422
- ✓ Lory, D.R. and Willumsen, B.M. **Function and regulation of RAS.** Annu Rev Biochem 1993, 62, 851-891.
- ✓ Lowndes NF, Paul J, et al. **c-Ha-ras gene bidirectional promoter expressed in vitro : and regulation.** Molecular and Cellular Biology, 1989, 9, 3758-3770
- ✓ Lowndes NF, Bushel P, et al. **A short, highly repetitive element in intron –1 of the human c-Ha-ras gene acts as a block to transcriptional readthrough by a viral promoter.** Molecular and Cellular Biology, 1990, 10, 4990-4995
- ✓ Maizels N **Dynamic roles for G4 DNA in the biology of eukaryotic cells.** Nature Structural & Molecular Biology 2006, 13, 1055-1059
- ✓ Marshall CJ **Ras effectors.** Current opinion in Cell Biology 1996, 8, 197-204
- ✓ Membrino A, Paramasivam M, et al. **Cellular uptake and binding of guanine-modified phthalocyanines to *KRAS*/*HRAS* G-quadruplexes.** ChemComm, 2010, 46, 625-627
- ✓ Membrino A, Cogoi S, et al. **G4-DNA formation in the HRAS promoter and rational design of decoy oligonucleotides for cancer therapy.** PLoS ONE 2011, 6, e24421
- ✓ Mitin N, Rossman KL, et al. **Signaling Interplay in Ras superfamily function.** Current Biology 2005, 15, 563-574

- ✓ Mo L, Zheng X, et al. **Hyperactivation of Ha-ras oncogene, but not Ink4a/Arf deficiency, triggers bladder tumorigenesis.** The Journal of Clinical Investigation 2007, 117, 314-325
- ✓ Neidle S, Read MA **G-quadruplexes as therapeutic targets.** Biopolymers 2001, 56, 195-208
- ✓ Neidle S **The structures of quadruplex nucleic acids and their drug complexes.** Current opinion in Structural Biology 2009, 19, 239-250
- ✓ Ou T, Lu Y, et al. **G-quadruplexes: targets in anticancer drug design.** ChemMedChem 2008, 3, 690-713
- ✓ Paeschke K, Capra JA, et al. **DNA replication through G-quadruplex motifs is promoted by the Saccharomyces cerevisiae Pif DNA helicase.** Cell, 2011, 145, 678-691
- ✓ Palumbo SL, Memmott RM, et al. **A novel G-quadruplex –forming GGA repeat region in the c-myb promoter is a critical regulator of promoter activity.** Nucleic Acids Research 2008, 36, 1755-1769
- ✓ Paramasivan S, Rujan I, et al. **Circular dichroism of quadruplex DNAs: applications to structure, cation effects and ligand binding.** Methods 2007, 43, 324-331
- ✓ Paramasivam M, Membrino A, et al. **Protein hn RNP A1 and its derivative Up1 unfold quadruplex DNA n the human *KRAS* promoter: implications for transcription.** Nucleic Acids Research, 2009, 37, 2841
- ✓ Parks CL, Shenk T **The serotonin 1a receptor gene contains a TATA-less promoter that responds to MAZ and Sp1.** J Biol Chem. 1996, 271, 4417-4430
- ✓ Parks CL, Shenk T. **Activation of the adenovirus major late promoter by transcription factors MAZ and Sp1.** J Virol. 1997, 71, 9600-9607
- ✓ Pestov DG, Dayn A, et al. **H-DNA and Z-DNA in the mouse c-Ki-ras promoter.** Nucleic Acids Research 1991, 19, 6527-6532
- ✓ Porter AC, Vaillancourt RR **Tyrosine kinase receptor-activated signal transduction pathways which lead to oncogenesis.** Oncogene 1998, 16, 1343-1352
- ✓ Qin Y, Rezier EM, et al. **Characterization of the G-quadruplexes in the duplex nuclease hypersensitive element of the *PDGF-A* promoter and modulation of *PDGF-A* promoter activity by TMPyP4.** Nucleic Acids Research 2007, 35, 7698-7713
- ✓ Qin Y, Hurley LH **Structures, folding patterns, and functions of intramolecular DNA G-quadruplexes found in eukaryotic promoters regions.** Biochimie. 2008, 90, 1149-1171
- ✓ Ragazzon P, Chaires JB **Use of competition dialysis in the discovery of G-quadruplex selective ligands.** Methods 2007, 43, 313-323
- ✓ Rankin S, Reszka AP, et al. **Putative DNA quadruplex formation within the human *c-kit* oncogene.** J. Am. Chem. Soc. 2005, 127, 10584-10589
- ✓ Rawal P, Kummarasetti VBR, et al. **Genome-wide prediction of G4 DNA as regulatory motifs: role in *Escherichia coli* global regulation.** Genome Research 2006, 16, 644-655
- ✓ Rebollo A, Martinez-A C **Ras proteins: recent advances and new functions.** Blood 1999, 94, 2971-2980
- ✓ Reddy EP, Reynolds RK et al. **A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene.** Nature, 1982, 300, 149-152

- ✓ Reuther GW, Der CJ **The Ras branch of small GTPases: Ras family members don't fall far from the tree.** Current Opinion in Cell Biology 2000, 12, 157-165
- ✓ Salas TR, Petruseva I, et al. **Human replication protein A unfolds telomeric G-quadruplexes.** Nucleic Acids Research 2006, 34, 4857-4865
- ✓ Saxonov S, Berg P, et al. **A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters.** PNAS 2006, 103, 1412-1417
- ✓ Sen D, Gilbert W **Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis.** Nature 1988, 334, 364-366
- ✓ Shafer RH, Smirnov I **Biological aspects of DNA/RNA quadruplexes.** Biopolymers 2001, 56, 209-227
- ✓ Shklover J, Etzioni S, et al. **MyoD uses overlapping but distinct elements to bind E-box and tetraplex structures of regulatory sequences of muscle-specific genes.** Nucleic Acids Research 2007, 35, 7087-7095
- ✓ Schubbert S, Shannon K, Bollag G. **Hyperactive Ras in developmental disorders and cancer.** Nat Rev Cancer 2007, 7, 295-308
- ✓ Siddiqui-Jain A, Grand CL, et al. **Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-*MYC* transcription.** PNAS 2002, 99, 11593-11598
- ✓ Smith JS, Chen Q, et al. **Rudimentary G-quadruplex-based telomere capping in *Saccharomices cerevisiae*.** Nature Structural & Molecular Biology 2011, 18, 478-486
- ✓ Song J, Ugai H, et al. **Independent repression of a GC-rich housekeeping gene by Sp1 and MAZ involves the same cis-elements.** J Biol Chem. 2001, 276, 19897-19904
- ✓ Spiller DG, Giles RV, et al. **Improving the intracellular delivery and molecular efficacy of antisense oligonucleotides in chronic myeloid leukemia cells: a comparison of streptolysin-O permeabilization, electroporation, and lipophilic conjugation.** Blood 1998, 12, 4738-4746
- ✓ Sun D, Liu WJ, et al. **The proximal promoter region of the human vascular endothelial griwth factor gene has a G-quadruplex structure that can be targeted by G-quadruplex-interactive agents.** Mol. Cancer Ther. 2008, 7, 880-889
- ✓ Teng Y, Girvan AC, et al. **AS1411 alters the localization of a complex containing protein arginine methyltransferase 5 and nucleolin.** Cancer Res 2007, 67, 10491-10500
- ✓ Thakur RK, Kumar P, et al. **Metastases suppressor NM23-H2 interaction with G-quadruplex DNA within c-*MYC* promoter nuclease hypersensitive et induces c-*MYC* expression.** Ucleic Acids Research 2009, 37, 172-183
- ✓ Tomonaga T, Levens D **Activating transcription from single stranded DNA.** Proc. Natl. Acad. Sci. USA. 1996, 93, 5830-5835
- ✓ Todd AK **Bioinformatics approaches to quadruplex sequence location.** Methods 2007, 43, 246-251
- ✓ Todd AK, Neidle S **The relationship of potential G-quadruplex sequences in cis-upstream regions of the human genome to SP1-binding elements.** Nucleic Acids Res. 2008, 36, 2700-4
- ✓ Vageli D, Kiaris H, et al. **Transcriptional activation of H-ras, K-ras and N-ras proto-oncogene in human bladder tumors.** Cancer Letters 1996, 107, 241-247

- ✓ Waller ZAE, Sewitz SA, et al. **A small molecule that disrupts G-quadruplex DNA structure and enhances gene expression.** J. AM. Chem. Soc. 2009, 131, 12628-12633
- ✓ Wang F, Podell ER, et al. **The POT1-TPP1 telomere complex is a telomerase processivity factor.** Nature 2007, 445, 506-510
- ✓ Weisman-Shomer P, Naot Y, et al. **Tetrahelical forms of the fragile X syndrome expanded sequence d(CGG)n are destabilized by two heterogeneous nuclear ribonucleoprotein-related telomeric DNA-binding proteins.** The Journal of Biological Chemistry 2000, 275, 2231-2238
- ✓ Weisman-Shomer P, Cohen E, et al. **Distinct domains in the CarG-box binding factor A destabilize tetraplex forms of the fragile X expanded sequence d(CGG)n.** Nucleic Acids research 2002, 30, 3672-3681
- ✓ Wennerberg K, Rossman KL, et al. **The Ras superfamily at a glance.** Journal of Cell Science 2005, 118, 843-846
- ✓ Wierstra I. **Sp1: Emerging roles—Beyond constitutive activation of TATA-less housekeeping genes.** Bioc. Biophys. Research Comm., 2008, 372, 1-13
- ✓ Zaug AJ, Podell ER, et al. **Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension *in vitro*.** PNAS 2005, 102, 10864-10869
- ✓ Zhang QS, Manche L, et al. **hnRNP A1 associates with telomere ends and stimulates telomerase activity.** RNA 2006, 12, 1116-1128
- ✓ Zhao Y, Du Z, et al. **Extensive selection for the enrichment of G4 DNA motifs in transcriptional regulatory regions.** of warm blooded animals. FEBS Letters 2007, 581, 1951-1956

6.1. Figures sources

Figure 1: Huppert JL, Balasubramanian S **G-quadruplexes in promoters throughout the human genome** Nucleic Acids Research 2007, 35, 406-413

Figure 2: Maizels N **Dynamic roles for G4 DNA in the biology of eukaryotic cells.** Nature Structural & Molecular Biology 2006, 13, 1055-1059

Figure 3: Keniry MA **Quadruplex structures in nucleic acids.** Biopolymers 2001, 56, 123-146

Figure 4: Patel DJ, Phan AT, et al. **Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics** Nucleic Acids Research, 2007, 35, 7429-7455

Figure 5: Maizels N **Dynamic roles for G4 DNA in the biology of eukaryotic cells.** Nature Structural & Molecular Biology 2006, 13, 1055-1059

Figure 6: Francis SH, Blount MA, et al. **Mammalian Cyclic Nucleotide Phosphodiestera-ses: Molecular Mechanisms and Physiological Functions** Physiol. Rev. 2011, 91, 651–690

Figure 7: Keniry MA **Quadruplex structures in nucleic acids.** Biopolymers 2001, 56, 123-146

Figure 8: Perentesis JP, Bathia S, et al. **RAS oncogene mutations and outcome of therapy for childhood acute lymphoblastic leukemia** Leukemia 2004, 18, 685-692