



HAL
open science

hnRNPA1/UP1 Unfolds KRAS G-Quadruplexes and Feeds a Regulatory Axis Controlling Gene Expression

Annalisa Ferino, Julien Marquevielle, Himanshi Choudhary, Giorgio Cinque, Coralie Robert, Anne Bourdoncle, Raffaella Picco, Jean-Louis Mergny, Gilmar Salgado, Luigi Xodo

► **To cite this version:**

Annalisa Ferino, Julien Marquevielle, Himanshi Choudhary, Giorgio Cinque, Coralie Robert, et al.. hnRNPA1/UP1 Unfolds KRAS G-Quadruplexes and Feeds a Regulatory Axis Controlling Gene Expression. ACS Omega, 2021, 6 (49), pp.34092-34106. 10.1021/acsomega.1c05538 . hal-03504886

HAL Id: hal-03504886

<https://hal.science/hal-03504886>

Submitted on 30 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

hnRNPA1/UP1 Unfolds *KRAS* G-Quadruplexes and Feeds a Regulatory Axis Controlling Gene Expression

Annalisa Ferino, Julien Marquevielle, Himanshi Choudhary, Giorgio Cinque, Coralie Robert, Anne Bourdoncle, Raffaella Picco, Jean-Louis Mergny, Gilmar F. Salgado,* and Luigi E. Xodo*



Cite This: *ACS Omega* 2021, 6, 34092–34106



Read Online

ACCESS |



Metrics & More

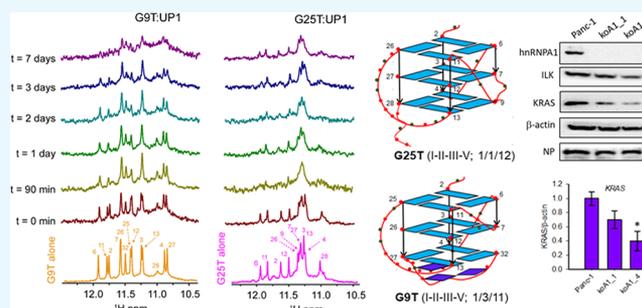


Article Recommendations



Supporting Information

ABSTRACT: Recent studies have proven that the genetic landscape of pancreatic cancer is dominated by the *KRAS* oncogene. Its transcription is controlled by a G-rich motif (called 32R) located immediately upstream of the TSS. 32R may fold into a G-quadruplex (G4) in equilibrium between two G4 conformers: G9T ($T_M = 61.2$ °C) and G25T ($T_M = 54.7$ °C). We found that both G4s bind to hnRNPA1 and its proteolytic fragment UP1, promoting several contacts with the RRM protein domains. 1D NMR analysis of DNA imino protons shows that, upon binding to UP1, G25T is readily unfolded at both 5' and 3' tetrads, while G9T is only partially unfolded. The impact of hnRNPA1 on *KRAS* expression was determined by comparing Panc-1 cells with two Panc-1 knockout cell lines in which hnRNPA1 was deleted by the CRISPR/Cas9 technology. The results showed that the expression of *KRAS* is inhibited in the knockout cell lines, indicating that hnRNPA1 is essential for the transcription of *KRAS*. In addition, the knockout cell lines, compared to normal Panc-1 cells, show a dramatic decrease in cell growth and capacity of colony formation. Pull-down and Western blot experiments indicate that conformer G25T is a better platform than conformer G9T for the assembly of the transcription preinitiation complex with PARP1, Ku70, MAZ, and hnRNPA1. Together, our data prove that hnRNPA1, being a key transcription factor for the activation of *KRAS*, can be a new therapeutic target for the rational design of anticancer strategies.



INTRODUCTION

The transcription of human *Kirsten ras* gene (*KRAS*) is regulated by a G-rich element (called 32R) located between –144 and –112 from the transcription start site (TSS).^{1,2} Sequence 32R forms a stable G-quadruplex (G4) structure recognized by nuclear proteins including PARP-1, Ku70, and hnRNPA1. These proteins have been identified by biotin–streptavidin pull-down assays coupled to mass spectrometry.³ In addition, a DNA-binding protein tool (Matinspector, Genomatix) predicted that the Myc-associated zinc-finger protein (MAZ) should also recognize 32R. This was indeed confirmed by EMSA and chromatin immunoprecipitation.^{4,5} Further studies suggested that the 32R G4 should act as a platform for the recruitment of TFs to the promoter to form the transcription preinitiation complex.⁶ Indeed, by silencing MAZ or PARP-1 with specific siRNA, we observed a downregulation of *KRAS* transcription.^{5,6} Within this framework, a question still remains unanswered: what is the role of hnRNPA1 in the *KRAS* promoter?

hnRNPA1 is a multifunctional protein regulating several aspects of mRNA metabolism, nuclear export,^{7–12} translation,^{13,14} and telomerase activity.¹⁵ Protein hnRNPA1 is composed of 322 amino acids, and its N-terminal contains two RNA recognition motif (RRM) domains followed by a highly

flexible glycine-rich (Gly-rich) C-terminal region, which acts as an RNA-binding domain and as a nuclear targeting sequence.¹⁶ Its N-terminal portion of 195 amino acids containing the RRM domains, called UP1, has been extensively studied by X-ray crystallography and NMR spectroscopy.¹⁷ Some high-resolution crystal structures of the two tandem RRMs of hnRNPA1 have been obtained with the free protein or with the protein bound to telomeric DNA repeats at a resolution of 1 Å.^{18–20} In addition to RNA, hnRNPA1 has been found to be associated with promoter sequences and to participate in the regulation of transcriptional events.^{7–9} The association of hnRNPA1 with the promoters of thymidine kinase (TK)⁷ and gamma-fibrinogen⁸ was found to repress transcription, while hnRNPA1 acts as an activator in the promoters of the ApoE⁹ and interferon-inducible RNA-dependent protein kinase genes.²¹

Received: October 5, 2021

Accepted: November 12, 2021

Published: November 30, 2021



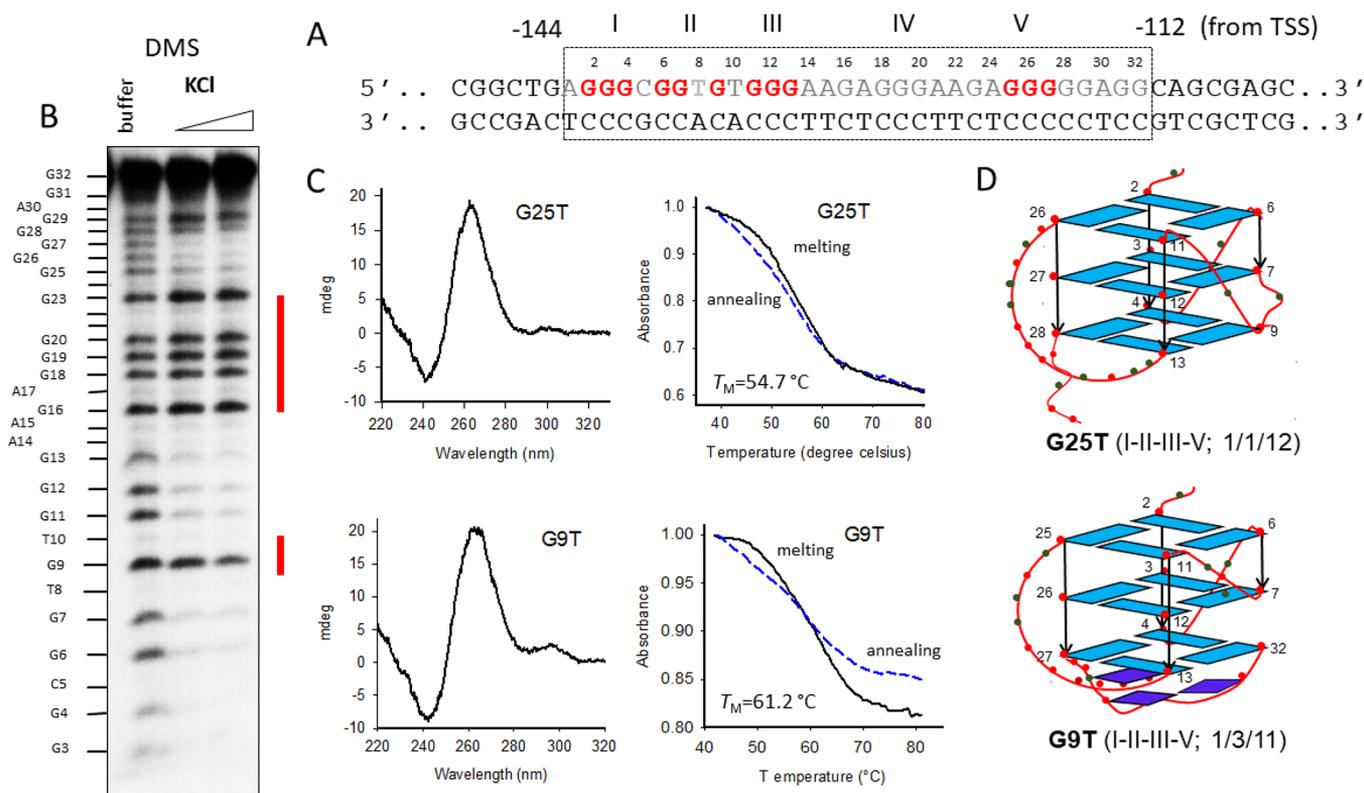


Figure 1. (A, B) 32R sequence and its typical DMS footprinting in 0, 50, and 100 mM KCl, (C) CD and UV-melting profiles in 50 mM N-acacodylate (pH 7.4) and 100 mM KCl of 3 μM G25T and G9T. Guanines in red form the G-tetrads according to DMS-footprinting; (D) NMR structure of the two G4 conformers formed by 32R adapted from ref 29 (*Nucleic Acids Res.* 2020, 48, 9336–9345), Oxford University Press.

Some authors have reported that hnRNPA1 is able to recognize and unfold DNA and RNA G4 structures. The first paper reporting this feature was published in 2002 by Fukuda *et al.*,²² who demonstrated that the G4 structures from the minisatellite repeat and telomeric DNA are unfolded by UP1. Some years later, our laboratory discovered that the G4 structure formed by 32R is recognized and unfolded by hnRNPA1 and UP1.²³ Moreover, single molecule FRET experiments showed that the telomeric DNA overhang is partially unfolded by hnRNPA1.²⁴ It has been reported that the RGG-box of hnRNPA1 recognizes telomeric G4 DNA and enhances the G4 unfolding of UP1.²⁵ The same authors also reported that the glycine–arginine-rich domain (RGG-box) of hnRNPA1 specifically recognizes TERRA G4 RNA but not single-stranded RNA.²⁶ All these studies suggest that hnRNPA1 is a nuclear protein associated with unusual DNA and RNA G4 structures, for which the association model and role are still unknown.

A recent NMR study from our laboratories showed that 32R folds into two co-existing conformers, called G25T and G9T, characterized by a different structure. Here, by EMSA and NMR, we explored the interaction between UP1 and the two KRAS G4 conformers. In addition, we tried to define the role of hnRNPA1 in the KRAS promoter. Previous work showed that hnRNPA1 is able to interact with the G4 formed by 32R^{23,27} and that KRAS is controlled by the KRAS-ILK-hnRNPA1 regulatory loop.²⁸ To further address this issue, we employed CRISPR/Cas9 technology to obtain hnRNPA1 knockouts of Panc-1 cells (koA1). We compared the expression of the ras genes in the normal and knockout cells and the capacity of these cells to survive and grow. We

concluded that hnRNPA1 is an essential TF for the transcription of KRAS. Our study opens a new therapeutic window for designing anticancer drugs to treat pancreatic ductal adenocarcinomas (PDACs).

RESULTS AND DISCUSSION

The KRAS 32R Sequence Forms Two G4 Structures.

Although KRAS holds two G4 motifs (32R and G4-mid), the vast majority of the studies reported in the literature focused on 32R, as this G-rich motif overlaps with a nuclease hypersensitive site and is an important platform for the recruitment of TFs.³ The first evidence that 32R spontaneously folds into a G4 structure was observed by running primer-extension experiments using two plasmids as DNA templates: one bearing the human 32R sequence and the other bearing its murine homolog.^{2,3} The finding that DNA polymerase I paused at the 3'-end of both G-rich motifs suggested the formation of a folded G4 structure by both templates. To determine the guanines of the G4 motif involved in the formation of the G-tetrad core, DMS footprinting experiments were carried out.^{2,3} In Figure 1A,B, we report a typical cleavage pattern of the human 32R motif. The expected folding involving G-runs I, III, IV, and V was not observed. The footprinting showed that G-run IV (G18–G19–G20) is strongly reactive to DMS, while guanines G6 and G7 are instead protected and G9 partially protected. This indicates that G-run II (GGTG) takes part in the formation of the G-tetrad core, while G-run IV does not. Combining footprinting and CD data for the critical 32R motif of the human KRAS promoter, we proposed a tri-stacked G-tetrad parallel G4 structure with two

1-nt and one 11-nt loops and a T-bulge in one strand (1/1/11 topology)³ (Figure S1).

Recently, we carried out an NMR study to gain insight into the folding of 32R.²⁹ The results indicated that 32R assumes two major G4 conformations, which are reported in Figure 1C,D (Table 1). The one called G25T has a structure similar

Table 1. Oligonucleotides Used in this Study^a

Name	Sequence 5' → 3'
32R	AGGGCGGTGTGGGAAGAGGGAAGAGGGGGAGG
92 ⁴⁰	AGGGC ^{ox} GGTGTGGGAAGAGGGAAGAGGGGGAGG
96 ⁴⁰	AGGGCGGTGTGGGAA ^{ox} GA ^{ox} GGAAGAGGGGGAGG
G9T	AGGGCGGT ^T TGGGAAGAGGGAAGAGGGGGAGG
G25T	AGGGCGGTGTGGGAAGAGGGAAGA ^T GGGGAGG
32Rmut	AGTCCGTTTTGTGAAGAGGGAAGAGTGGGAGG
b-32R	Biotin-TTTTTAGGGCGGTGTGGGAAGAGGGAAGAGGGGGAGG
b-G9T	Biotin-TTTTTAGGGCGGT ^T TGGGAAGAGGGAAGAGGGGGAGG
b-G25T	Biotin-TTTTTAGGGCGGTGTGGGAAGAGGGAAGA ^T GGGGAGG
Cy5-32R	Cy5.5-TTTTTAGGGCGGTGTGGGAAGAGGGAAGAGGGGGAGG
Cy5-G9T	Cy5.5-TTTTTAGGGCGGT ^T TGGGAAGAGGGAAGAGGGGGAGG
Cy5-G25T	Cy5.5-TTTTTAGGGCGGTGTGGGAAGAGGGAAGA ^T GGGGAGG

^aoxG = 8-oxoguanine; b = biotin; red T = G/T substitution in G9T and G25T compared to 32R.

to that proposed for 32R on the basis of CD and DMS footprinting: a tri-stacked G-tetrad G4 with a T-bulge in one strand, two 1-nt, a 12-nt loop, and all guanines in *anti*-conformation. G25T ($T_M = 54.7$ °C) is in equilibrium with G9T ($T_M = 61.2$ °C), which exhibits a structure characterized by a fold-back guanine in *syn* conformation (G32) and a triad (G29, A30, and G31) capping the 3'-end. In addition, in previous studies,^{30,31} we observed by primer extension experiments that G-runs I, II, III, and IV may fold into an alternative G4 with a 1/1/4 topology in the presence of a G4-stabilizing phthalocyanine (DIGP). However, the fact that, in the absence of DIGP, Taq polymerase paused only at G32 clearly suggests that 32R folds spontaneously into the G9T/G25T G4 conformers, which can be considered the major G4 structures of sequence 32R.

The Two G4 Structures of KRAS Interact with hnRNPA1 and UP1. In 2008, we carried out pull-down and mass spectrometry experiments and found that the critical 32R G4 motif is recognized by several nuclear proteins including PARP-1, Ku70, and hnRNPA1.³ Later on, we discovered that MAZ also binds to 32R.^{4,5,32} As stated above, in this study, we focused on the role played by hnRNPA1 in the KRAS promoter. First, we investigated by EMSA if both 32R G4 conformers are recognized by hnRNPA1/UP1. Figure 2A shows that G9T and G25T with hnRNPA1 form two DNA-protein complexes, c1 and c2, of different electrophoretic mobility. The wild-type 32R sequence forms, in addition to c2, another complex of very low mobility. In contrast, the 32R duplex shows little affinity for hnRNPA1. We also examined the proteolytic fragment of 196 amino acids of hnRNPA1 called unfolding protein 1 (UP1), which maintains both the binding and G4-unfolding capacity of the entire protein.²² It can be seen that UP1 also forms with the two G4 conformers DNA-protein complexes. The fact that these complexes do not run with sharp bands may be due to the complexity of the interaction involving the disruption of the G4 structures. Considering that the protein upon binding to G4 unfolds the structure, a 32R mutant unable to form a G4 (32Rmut) is also bound by UP1. The structure of the DNA-protein complexes

observed by EMSA can be predicted from the crystal of UP1 bound to the telomeric d(TTAGGG)₂ oligonucleotide (TR2).²⁰ TR2 and UP1 form a dimeric complex consisting of two oligonucleotides and two protein molecules. The two TR2 strands are antiparallel to one another and completely unfolded. The complex is stabilized by multiple interactions occurring between the TTAGGG hexamers and the two RRM protein domains. In keeping with the TR2-UP1 crystal,²⁰ a structural model for the complexes formed by the G4 conformers G25T/G9T and hnRNPA1/UP1 is proposed: a U-shaped complex with (1:1, c1) and (1:2, c2) stoichiometry.^{23,24}

Subsequently, we determined the K_D 's of the interaction between the G4 structures and UP1 by isothermal titration calorimetry. Owing to low yields in expressing UP1 and hnRNPA1, titrations were conducted with UP1 in the sample cell and G4 in the injection syringe (reverse titration). Figure 2B shows the binding curves obtained by plotting the area of the peak versus the G4/protein molar ratios. The binding curve analysis gave dissociation constants (K_D 's) between 0.49 and 1.1 μ M and ΔG of complex formation between -8.5 and -8.9 kcal/mol (Table 2). We also obtained a 1:2 stoichiometry for complex G9T-UP1, in keeping with EMSA. Instead, the binding curves of G25T-UP1 and 32R-UP1 suggested a more complex stoichiometry, >1:2, probably owing to the apparent extra degree of flexibility that these sequences seem to have from NMR spectra.

Interaction between hnRNPA1/UP1 and KRAS G4s by NMR. The interaction between G25T/G9T and UP1 was investigated by NMR. We performed titrations with uniformly ¹³C and ¹⁵N isotopically labeled UP1 followed by the evolution of each protein residue upon the addition of either G9T or G25T by 2D ¹H-¹⁵N HSQC NMR experiments. The analysis of the chemical shift deviations ($\Delta\delta$ /ppm) of the amide group for the most affected amino acids is proportional to the change in the chemical environment caused by the interaction with G4. To better assess the chemical shift differences, we superimposed the spectra before and after each successive G4 addition and depicted the most important $\Delta\delta$ as a function of the residues (Figure 3A-C).

The spectrum of UP1 alone showed peaks that were separated and well resolved as described in the literature.¹⁷ Upon the addition of G4, the UP1 spectrum became more complex and some peaks, especially in a central region around 8.5 ppm for the ¹H dimension and around 122.5 ppm for the ¹⁵N dimension, faded in intensity, which are typical of peaks undergoing chemical exchange from local unfolding events and dynamics. Although these perturbations are complex to interpret, they were accurately examined to determine the binding with the G4s. Nevertheless, the vast majority of peaks were identifiable up to molar ratios of 1:1. Among the remaining peaks that could be analyzed, some peaks either did not shift or disappeared. While the former are probably not involved in the interaction, the latter ones may play a specific, yet undefined, role. We compared the global chemical shift peak pattern in the HSQC, and we identified unambiguously some residues, such as R7, K45, and R75 (purple arrows), that have relatively important shifts. The same residues are also involved through hydrogen bonding in the binding of UP1 with a telomeric repeat sequence.³³ These residues belong to a nucleic-acid binding region β -sheet platform and a short α -helical turn interdomain that connects both RRM domains. The global shifts, calculated from eq 1 described in the

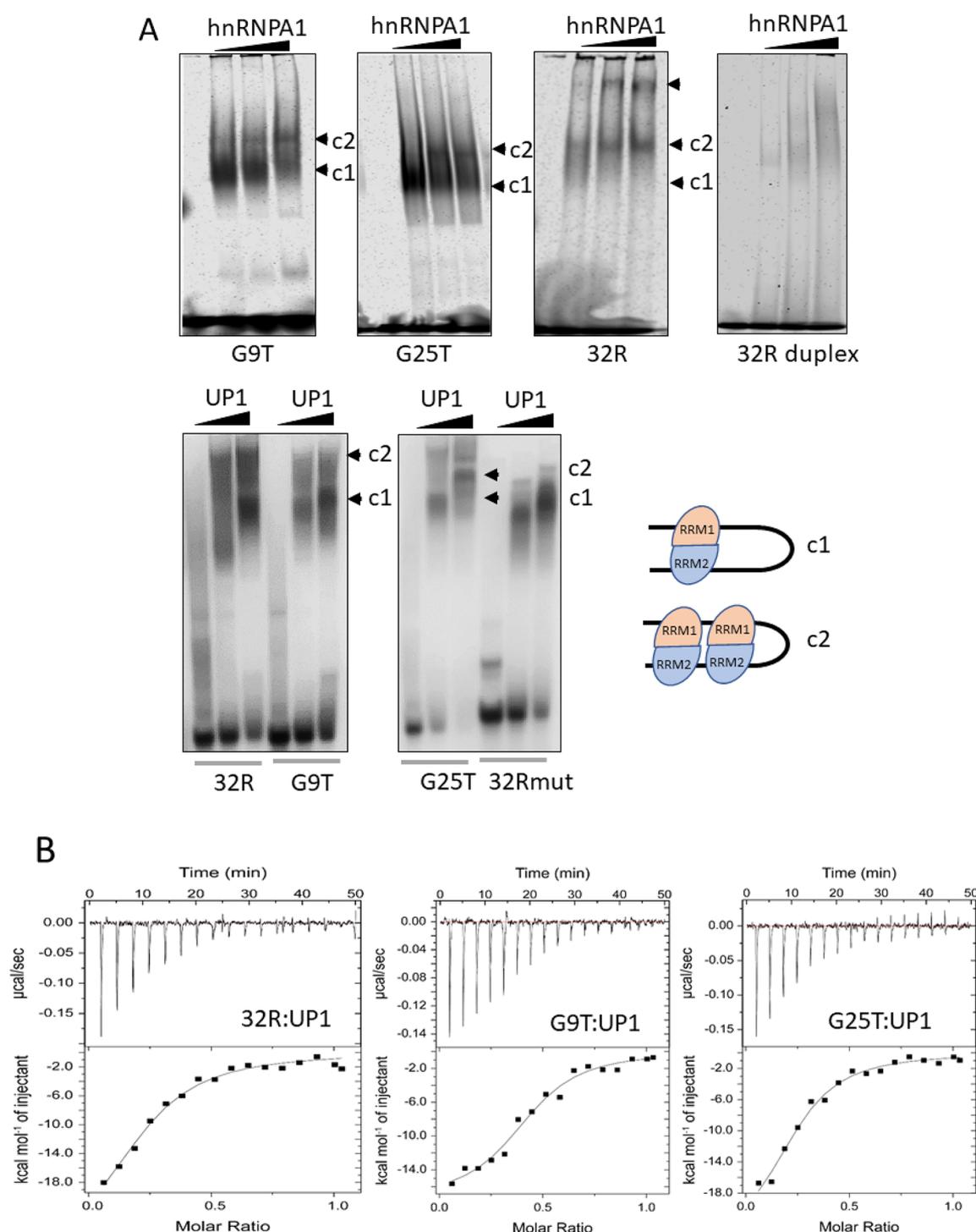


Figure 2. (A) EMSA showing the binding of hnRNPA1 and UP1 to G9T, G25T, and 32R G4 structures. Samples with hnRNPA1, containing 50 nM G4 labeled with Cy5.5 and 0, 2.5, 5, or 10 µg of protein, were incubated in 50 mM Tris–HCl (pH 7.4), 50 mM KCl, 2.5 ng/mL poly[dI-dC], 1 mM EDTA, 50 mM Zn-acetate, 1 mM NaV, 5 mM NaF, 0.01% phosphatase inhibitor, 1 mM DTT, and 8% glycerol for 30 min at 25 °C and then run in 5% PAGE in TBE. EMSA with UP1 were run in a 10% PAGE. 32Rmut is unable to assume a G4 structure (Table 1). Proposed models for the complexes c1 and c2 between hnRNPA1/UP1 and G9T/G25T are shown. (B) Isothermal calorimetry titrations relative to the binding of UP1 to 32R, G9T, and G25T G4 structures at 37 °C in a phosphate buffer and 50 mM KCl (pH 6.6). Binding isotherms from which the thermodynamic parameters of the interaction were obtained (Table 2).

experimental procedures, are plotted in Figure 3C. The plot shows that interactions with the G9T and G25T G4 conformers involve residues in both RRM domains of UP1, with some preference for domain 1. This observation is in agreement with the crystal structure between UP1 and TR2²⁰

and the results supported by other NMR and ITC experiments available in the literature.²⁵ The residues with the most intense $\Delta\delta$ have been plotted in red and orange within the UP1 structure (Figure 3C). To probe the effect of UP1 on the folding of both 32R G4 conformers, we performed 1D ¹H

Table 2. Thermodynamic Parameter Concerning the Interaction between UP1 and KRAS G4s

sequence	K_D (μM)	ΔG (kcal/mol)	ΔH (kcal/mol)	$T\Delta S$ (kcal/mol)
32R	1.1 ± 0.25	-8.5 ± 1.9	-29 ± 4.0	-20.5 ± 5.9
G9T	0.49 ± 0.12	-8.9 ± 2.2	-17 ± 1.0	-8.1 ± 3.2
G25T	0.79 ± 0.18	-8.6 ± 2.0	-25 ± 2.6	-16.4 ± 4.6

NMR experiments as a function of time, looking at the G-quadruplex imino signatures at a 1:1 molar ratio. The NMR spectra were acquired at different time periods, and the samples were kept at 37 °C for a week (Figure 4). Ninety minutes after the addition of UP1, we observed a significant decrease in the intensity of all imino peaks of G25T, suggesting that the sequence bound to UP1 was unfolded, although probably not completely. Instead, the G9T showed a far lower drop in the intensity of the imino peaks. It can be seen that the G9T most affected imino protons, showing broadening and decline in intensity after 90 min incubation with UP1, are those of the guanines corresponding to the 5'-end tetrad (G2, G6, G11, and G25) (Figure 1D). The imino protons of the guanines of the 3'-end tetrad (G4, G32, G13, and G27) instead show an initial broadening, but their intensity slowly decreases within a time scale of hours/days (after 1 day, the G4 should be partially unfolded in rapid refolding equilibrium when not bound to UP1). As for G25T, all imino peaks except those of the central G-tetrad (G3, G7, G12, and G27) disappeared after 90 min exposure to UP1, showing that the end-tetrads are disrupted. The data suggest that, under the experimental conditions of the experiments, UP1 binds to the G4-ends of the structures (Figure S2). It is noteworthy that, gradually over time, some peaks in both G4 conformers reappeared. This process is mostly due to the unfolding of UP1 itself, especially after more than 2 days at 37 °C, confirmed by an HSQC experiment (not shown). Taken together, both the 2D and 1D NMR experiments indicate that G25T in the presence of UP1 is practically unfolded, while conformer G9T is only partially unfolded. In agreement with the data of the TR2-UP1 complex,²⁰ chemical shift plots reported in Figure 3C show that there are many contacts between G25/G9T and the two RRM domains of the proteins. As the unfolding of both G4s appears to be not complete, we can hypothesize that UP1 without the glycine C-terminus domain (RGG-box) is not as efficient in unfolding as the entire hnRNPA1 protein, as observed in the case of human telomeric G-quadruplex Tel22,²⁵ or that not all the G4 molecules are bound to UP1 at a 1:1 ratio. In fact, at G4/protein ratios of 1:5 and 1:10, FRET experiments suggest that G9T bound to hnRNPA1 or UP1 is unfolded (Figure S3).

hnRNPA1 Is Upregulated in PDAC Cells and Plays a Key Role in the KRAS Promoter. Previous studies suggested that hnRNPA1 should be involved in the mechanism regulating KRAS transcription.^{23,27,28} We then asked what the real impact of hnRNPA1 in KRAS expression is. To address this issue, we compared KRAS expression in normal and knockout Panc-1 cells, in which hnRNPA1 was deleted by the CRISPR/Cas9 technology (Panc-1 cells are human PDAC cells bearing the KRAS mutation G12D). The genome editing of Panc-1 cells was carried out by Synthego (CA), which provided us with a pool of Panc-1 edited cells from which we managed to isolate three clones: koA1_1, koA1_4, and koA1_8, which were fully knocked out for hnRNPA1.

Figure 5A,B reports the guide and target sequences used to obtain the hnRNPA1 knockout cell lines as well as a typical Western blot showing that koA1_1, koA1_4, and koA1_8 do not express hnRNPA1, while they do express β -actin. To confirm the specificity of hnRNPA1 knockout, we detected the level of hnRNPA1 isoforms such as hnRNP M, hnRNP F/H, and hnRNP A2/B1 (Figure 5C). We observed that these isoforms are equally expressed in normal and knockout cells, as expected. Subsequently, we reasoned that if hnRNPA1 is a TF essential for KRAS, the knockout cell lines should express a lower level of KRAS compared to normal Panc-1 cells. To test this, we measured the levels of KRAS in koA1_1 and koA1_4 knockout cells (Figure 6A,B).

Compared to β -actin and nucleoporin, both koA1_1 and koA1_4 express a lower level of KRAS protein: residual KRAS is ~ 70 and $\sim 40\%$ in koA1_1 and koA1_4, respectively, compared to normal Panc-1 cells. As a control, we silenced hnRNPA1 in normal Panc-1 cells (residual hnRNPA1 $\sim 50\%$) by using a specific siRNA and observed that a transient suppression of hnRNPA1 resulted in the downregulation of KRAS by $\sim 50\%$, in agreement with the results obtained with koA1_1 and koA1_4 knockout cell lines (Figure 6C,D). Taken together, the data obtained with the knockout and normal Panc-1 cells treated with siRNA clearly suggest that hnRNPA1 is important for KRAS expression. We then asked ourselves whether in the knockout cell lines the downregulation of KRAS is compensated by an overexpression of HRAS and NRAS.

The levels of the HRAS and NRAS proteins in koA1_1 and normal Panc-1 cells were measured by specific monoclonal antibodies (Figure 6E). We found that the knockout and normal cells show roughly similar levels of the HRAS and NRAS proteins. In conjunction with literature data, the role played by hnRNPA1 in the mechanism controlling KRAS transcription in pancreatic cancer can be represented as in Figure 6F. When KRAS transcription is stimulated by oxidative stress, *i.e.*, by treating the cells with H_2O_2 or with ROS-generating porphyrins,³⁴ we observed by CHIP that (i) hnRNPA1, MAZ, and PARP-1 are recruited to the KRAS promoter in the region containing the 32R motif³² and (ii) the level of 8OG increases in the 32R region more than in other genomic G-rich regions lacking G4 motifs.³² Therefore, we hypothesized that 8OG-modified G4 in the KRAS promoter acts as a platform for the recruitment of PARP-1, MAZ, and hnRNPA1 and the assembly of the transcription preinitiation complex. This and previous studies provide evidence that the function of hnRNPA1 is to unfold the G4 and facilitate the reconstitution of the duplex before the formation of the preinitiation complex with the recruited proteins. Chu *et al.*²⁸ showed that KRAS expression depends not only on hnRNPA1 but also on ILK, which forms an axis, ROS-KRAS-ILK-hnRNPA1, that maintains the expression of KRAS in PDAC high, as illustrated in Figure 6F. The high metabolic rate of PDAC enhances the level of ROS that stimulate TF recruitment and KRAS expression via ILK and hnRNPA1. If hnRNPA1 is suppressed, the axis and thus KRAS activities fall, together with the KRAS-induced metabolic rewiring necessary to produce biomass for cell growth.³⁵

KRAS promotes a complex downstream signaling involving the RAF/MEK/ERK and PI3K/PDK1/AKT pathways.³⁶ Recent findings have shown that the initiation, progression, and maintenance of PDAC heavily depend on the KRAS/PI3K/PDK1/AKT signaling, which stimulates cell growth and survival.³⁷ By Western blots, we investigated the activity of the

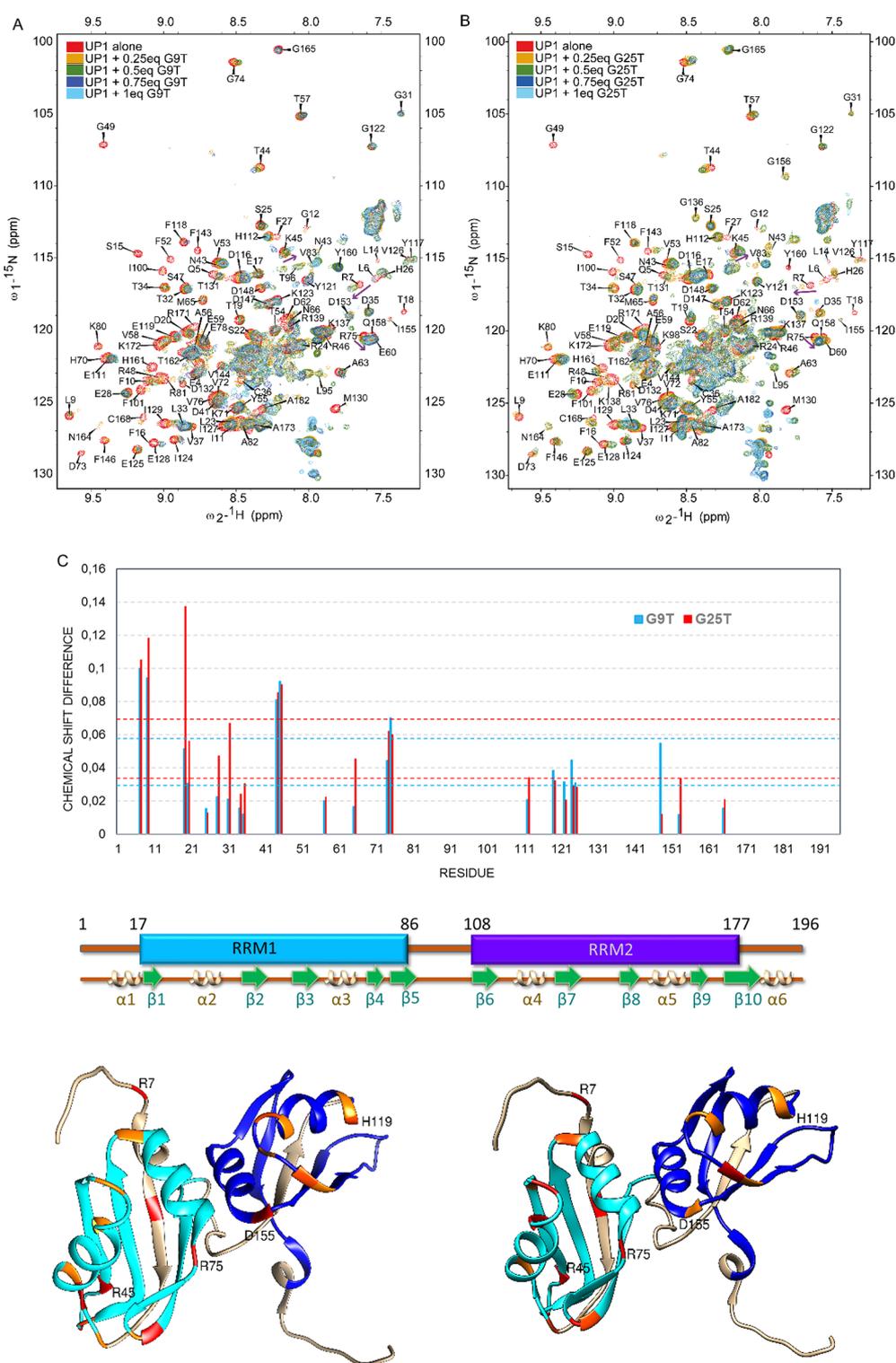


Figure 3. (A, B) Superimposition of ^{15}N - ^1H NMR HSQC spectra of UP1 showing each residue with NH bond of the backbone, measured alone and with an increasing amount of KRAS G9T (*top*) and G25T (*bottom*) G4 structures. (C) Plotted chemical shifts calculated using eq 1. Dotted lines indicate values of one and two sigma above standard deviation (SD). The corresponding residue with a schematic view of the UP1 structure to identify regions implicated in the interaction with the G9T and G25T G4 conformers. Structures of the two UP1 RRM domains with the most shifted residues in the presence of G9T (*left*) or G25T (*right*) are colored in red for strong shifts and in orange for medium shifts, e.g., 2 and 1 SD, respectively.

two pathways in Panc-1 and koA1_1 cell lines (Figure 7A). It can be seen that in koA1_1, which is characterized by a lower expression of KRAS (*vide infra*), the MEK/ERK pathway is substantially active as in normal cells, while the more critical

PI3K/PDK1/AKT pathway appeared inhibited, as indicated by the low level of phosphorylated AKT. As hnRNPA1 is a critical TF for KRAS, the knockout cell lines should exhibit a lower metabolic activity, proliferation, and colony formation

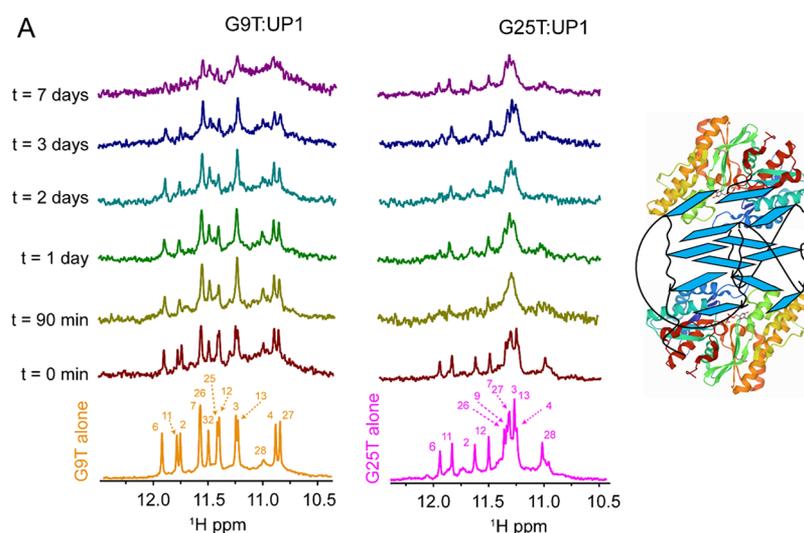


Figure 4. (A) G9T and G25T G4 imino proton region after the addition of 1 equiv of UP1 at different time periods. At a G4/UP1 ratio of 1:1, UP1 binds to the end-tetrads of the G4s and unfolds completely (G25T) or partially (G9T) the structures.

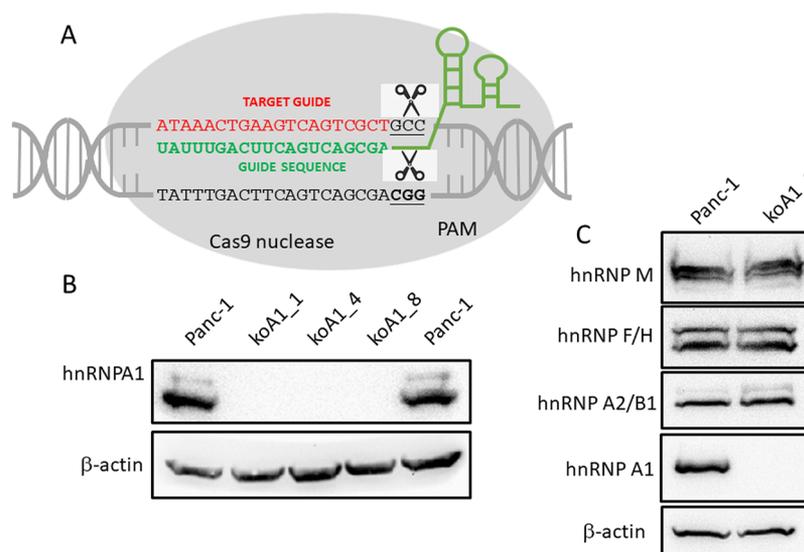


Figure 5. (A) Target and guide sequences used to suppress hnRNPA1 in Panc-1 cells. Two knockout clones for hnRNPA1 were isolated, koA1_1 and koA1_4, which show no expression of hnRNPA1. (B) Western blots showing that koA1_1 and koA1_4, but not the wild-type Panc-1 cells, do not express hnRNPA1, while they do express β -actin. (C) The DNA editing by the CRISPR Cas9 technology does not affect the expression of the hnRNP M, hnRNP F/H, and hnRNP A2/B1 isoforms of hnRNPA1.

compared to normal Panc-1 cells. To test this, we first carried out a resazurin assay that evaluates metabolic activity (resazurin in viable cells is enzymatically reduced to highly fluorescent resorufin). Figure 7B–D shows that three knockout cell lines (including koA1_8) have a lower metabolic activity than normal Panc-1 cells and a significantly lower proliferation over a period of 6 days from cell seeding. Moreover, a clonogenic assay showed that the suppression of hnRNPA1 in Panc-1 cells results in $\sim 60\%$ drop in colony formation. Together, these data provide strong evidence that hnRNPA1 plays a vital role in PDAC, as it stimulates the expression of KRAS, the oncogene to which pancreatic cancer cells are addicted.

The 32R G4 Motif Is a Platform for the Formation of the Preinitiation Complex. Previous studies support the notion that the KRAS G4 structures may function as a platform for the recruitment of TFs.^{6,32} Recently, we reported that upon

binding to the 32R G4, PARP-1 undergoes auto-PARylation, becomes negatively charged, and stimulates the recruitment of cationic TFs such as hnRNPA1 and MAZ ($pI > 7.4$). We therefore asked ourselves whether both G4 conformers of 32R are able to form a multiprotein complex when they are incubated with a nuclear extract from Panc-1 cells. To address this issue, we used a streptavidin–biotin pull-down approach. We synthesized G25T, G9T, and 32R linked to biotin and let them fold into G4 in a buffer containing 100 mM KCl. The biotinylated oligonucleotides in G4 conformation were used as G4 baits in the pull-down experiments (Figure 8A). Each biotinylated G4 was incubated with 80 μ g of nuclear extract in the presence of poly[dI-dC] to suppress unspecific binding for 30 min, and the proteins bound to G4 were pulled down with streptavidin-coated magnetic beads. The captured proteins (bound to the beads) were eluted with Laemmli buffer and analyzed by immunoblotting with antibodies specific for MAZ,

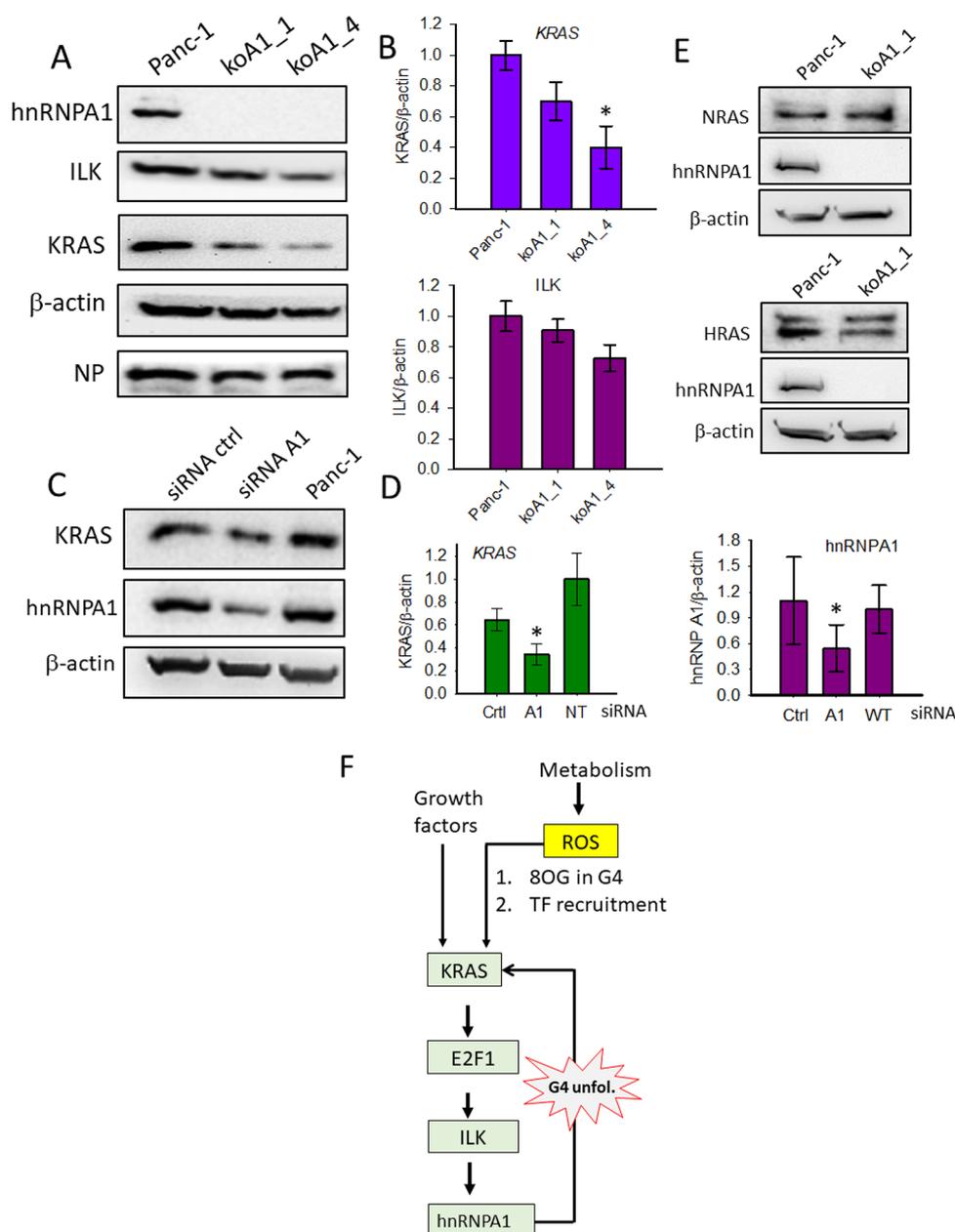


Figure 6. (A, B) Western blot showing the levels of hnRNPA1, ILK, KRAS, β -actin, and nucleoporin (NP) in the koA1_1 and koA1_4 knockouts and in wild-type Panc-1 cells. (C, D) Western blot showing the level of KRAS, hnRNPA1, and β -actin in Panc-1 cells untreated and treated with a specific siRNA for hnRNPA1 and control siRNA. (E) Western blots showing the levels of HRAS and NRAS in normal Panc-1 cells and in koA1_1 knockout cells. (F) The KRAS-ILK-hnRNPA1 axis controlling the expression of KRAS in PDAC cells. (*) = $P < 0.05$.

hnRNPA1, Ku70, and PARP-1. It was observed that the G4s pulled down all four TFs, suggesting that they can indeed act as a platform for the formation upstream of the TSS of the transcription preinitiation complex (Figure 8B).

The eluates from the streptavidin-coated beads, incubated with the nuclear extract in the absence of the G4 bait, contained a small amount of proteins owing to unspecific interactions between the magnetic beads coated with streptavidin and the nuclear proteins (lane "beads"). The result obtained with G25T is quite similar to that observed with 32R, while conformer G9T appears less efficient in pulling down the proteins. The fact that G9T seems to be a less efficient platform than 32R and G25 correlates with its higher resistance to modifying its structure upon interacting with the hnRNPA1/UP1.

Another point that we considered is the following: as the distribution of the TFs in the promoter is dynamic and their recruitment is expected to be the result of the balance between protein–protein and DNA–protein interactions, we asked ourselves whether the proteins recruited to the KRAS promoter act independently or interact with one another. To investigate this point, we carried out an immunoprecipitation assay (Figure 8C). The nuclear extract from Panc-1 cells was incubated one by one with the monoclonal antibodies (Abs) specific for the TFs. The proteins bound directly or indirectly to the antibodies were pulled down by magnetic beads coated with protein A and analyzed by Western blots. It can be seen that anti-PARP1 Ab pulled down in addition to PARP-1 also Ku70, suggesting that these two proteins are associated with each other. Anti-Ku70 Ab gave a similar result: it pulled down

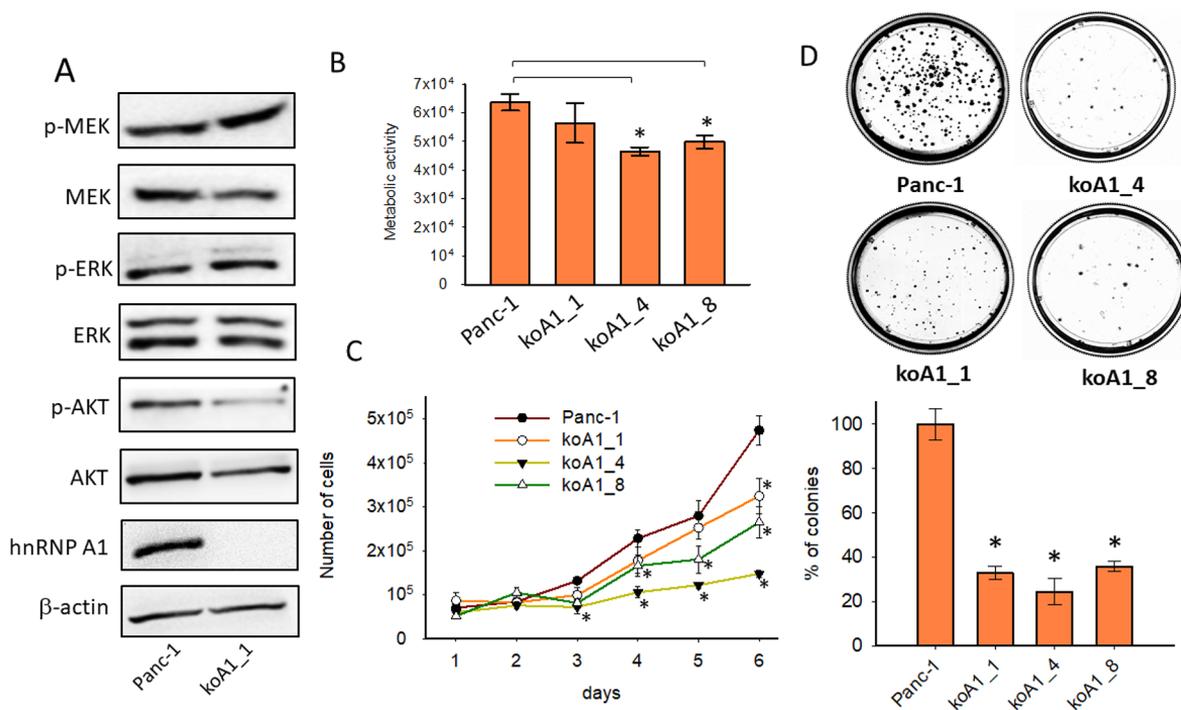


Figure 7. (A) Level of phosphorylation of KRAS downstream effector proteins. (B) Metabolic activity of normal and hnRNPA1-knockout Panc-1 cells measured 72 h after cell plating. (C) Cell growth assay reporting the number of normal and hnRNPA1-knockout Panc-1 cells up to 6 days from plating. (D) Clonogenic assay showing that normal Panc-1 cells form more colonies than the hnRNPA1-knockout clones. The bar plot shows a reduction of colony formation by the knockout clone of >60% compared to the wild-type cells. (*) = $P < 0.05$.

both Ku70 and PARP-1, confirming the contact between the two proteins. Anti-MAZ Ab only pulled down MAZ, while anti-hnRNPA1 Ab pulled down hnRNPA1 and MAZ in large amounts. The result suggests that MAZ and hnRNPA1 are strongly associated with each other. The fact that anti-MAZ Ab does not pull down hnRNPA1 indicates that the association between the two proteins overlaps the epitope recognized by anti-MAZ. In a second set of experiments, we used the 32R G4 containing 8-oxoguanine (8OG) as a bait to mimic a ROS-oxidized G4. We designed two oxidized G4 structures, called 92 and 96, the former bearing 8OG in G-run I and the latter in the major groove (Figure S4 and Table 1).³² When we incubated the wild-type and oxidized 32R sequences with a Panc-1 extract, we observed that the oxidized G4s pulled down the TFs as efficiently as wild-type 32R, indicating that the oxidized G4 acts as a platform for the recruitment of the TFs (Figure 8D). Interestingly, when we carried out a pull-down experiment with an extract obtained from the koA1_1 knockout cell line, we found that not only hnRNPA1 but also MAZ was not pulled down, in agreement with the fact that MAZ in the multiprotein complex is associated with hnRNPA1 (Figure 8E). In Figure 8F, we propose a mechanism for KRAS transcription activation. Under enhanced oxidative stress, typical of cancer cells, PARP-1 and its associated Ku70 protein are recruited to the KRAS promoter in the region containing the 32R G4 motif, most likely with 8OG modification. Upon binding to G4, PARP-1 undergoes autoparalation and becomes negatively charged.^{6,38} Ku70, which is associated with PARP-1, having a $pI = 6.23$, is also anionic under physiological conditions. The resulting G4-PARP1-Ku70 complex forms a strongly anionic platform capable of recruiting cationic TFs such as hnRNPA1 ($pI = 9.2$). The electrostatic attraction of

hnRNPA1 to the promoter should also recruit MAZ as it is associated with hnRNPA1.

The enrichment of the TFs in the neighboring G4 creates the conditions for the formation of the transcription preinitiation complex. Owing to the G4 unfolding property of hnRNPA1 and MAZ,^{6,22,23} the G4 structures are unfolded and the transcription preinitiation complex is assembled on double-stranded DNA.

Finally, we compared the morphology of the knockout cell line koA1_1 with normal Panc-1 cells by performing confocal microscopy experiments (Figure S5). We obtained images of Panc-1 cells stained with phalloidin, syto-14, and Hoechst. Phalloidin binds to actin filaments and stains the cytoskeleton of the cells, syto-14 binds to cellular RNA, and Hoechst stains the nucleus. Compared to wild-type Panc-1 cells, the koA1_1 knockout appears more aggregated in keeping with the fact that the downregulation of KRAS affects cell adhesion.³⁹

Correlation between the KRAS-ILK-hnRNPA1 Axis and PDAC Survival Probability. As the development, growth, and maintenance of PDAC heavily depend on KRAS,^{40,41} we asked ourselves whether the oncogene and the TFs recognizing the KRAS G4 structures are overexpressed in PDAC patients. We consulted a publicly available microarray data set (GSE15471) to examine the differential expression of these genes between normal and tumor tissue samples. GSE15471 reports the global gene expression of 36 pairs of normal and PDAC samples obtained from resected pancreas of cancer patients. The results are reported in Figure 9A in the form of box plots. It can be seen that KRAS is almost twofold upregulated in PDAC compared to normal tissues ($P < 10^{-7}$). Remarkably, the genes encoding for PARP-1, hnRNPA1, and Ku70 that recognize the 32R G4 are also upregulated in PDAC tissues ($P < 0.007$). Only MAZ seems to be slightly

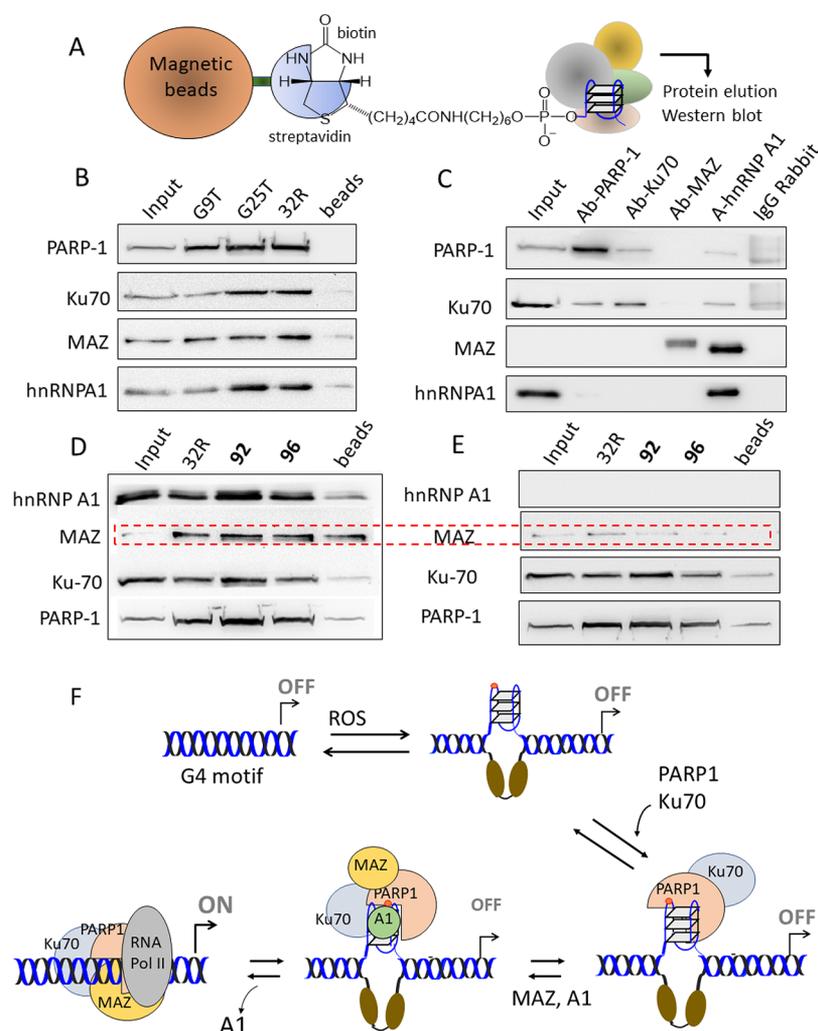


Figure 8. (A) Scheme of the pull-down experiments is illustrated. (B) Pull-down with biotinylated G9T, G25T, and 32R G4s. The biotinylated G4 (80 nM) was incubated with 80 μ g of nuclear extract for 30 min at RT. The DNA bait–protein complexes formed were pulled down with streptavidin magnetic beads. The pull-down proteins were recovered and analyzed by Western blot with anti MAZ, anti PARP-1, and anti hnRNP A1 primary antibodies and a secondary antibody conjugated to horseradish peroxidase. (C) Pull-down with antibodies specific for PARP-1, Ku70, MAZ, and hnRNP A1. The recovered proteins from the pull-down were analyzed by Western blots. (D, E) Pull-down assays with biotinylated 32R and oxidized analogues 92 and 96, in the G4 structure, used as bait with the extract from normal Panc1 cells (left panel) and knockout koA1_1 cells (right panel). (F) Proposed mechanism for the activation of KRAS transcription. First, PARP-1 binds to the KRAS promoter at the G4 motif. After binding the protein, it undergoes auto-PARYlation, becoming anionic. The G4–PARP-1 complex acts as a platform for the recruitment of the TFs. Protein hnRNP A1 should unfold the G4, thus promoting the G4 to duplex transformation at the promoter near TSS and the initiation of transcription.

downregulated ($P < 0.0012$). However, this finding is not in agreement with a recent study of Zhu *et al.*,⁴² who reported that MAZ is also upregulated in PDAC (the discrepancy may be due to a different method of analysis).

As it is now established that KRAS is controlled by an axis involving hnRNP A1 and ILK,²⁸ we also focused on ILK and found that its expression in PDAC is higher than in normal tissues ($P = 0.0012$). So, the crucial KRAS-ILK-hnRNP A1 axis controlling KRAS expression is composed by effector proteins that are overexpressed in PDAC. To provide further support of the clinical relevance of the KRAS-ILK-hnRNP A1 axis, we investigated if its expression level correlates with the overall clinic outcomes of different tumors. We obtained Kaplan–Meier plots and found that PDAC patients with a highly expressed KRAS-ILK-hnRNP A1 axis showed a lower survival probability than patients with a lowly expressed axis (Figure 9B). We divided the data of 178 PDAC patients into two

groups: one of 147 patients characterized by a high expression of KRAS-ILK-hnRNP A1 (group 1) and one of 31 patients with a low expression of the same genes (group 2). We then calculated the survival probability and found that group 2 had a survival probability significantly higher than that of group 1, $P = 0.038$. These data confirm the central role of the KRAS-ILK-hnRNP A1 axis in the maintenance of PDAC and suggest that hnRNP A1 is an interesting target for the rational design of anticancer drugs to treat PDAC.

CONCLUSIONS

The G4-motif located upstream of the transcription start site folds into a G-quadruplex in equilibrium between two G4 conformers: G9T ($T_M = 61.2$ °C) and G25T ($T_M = 54.7$ °C).²⁹ Here we have demonstrated that both G4s interact with hnRNP A1 and its proteolytic fragment UP1. 1D NMR analysis of G4 imino protons shows that, upon binding to UP1, G25T

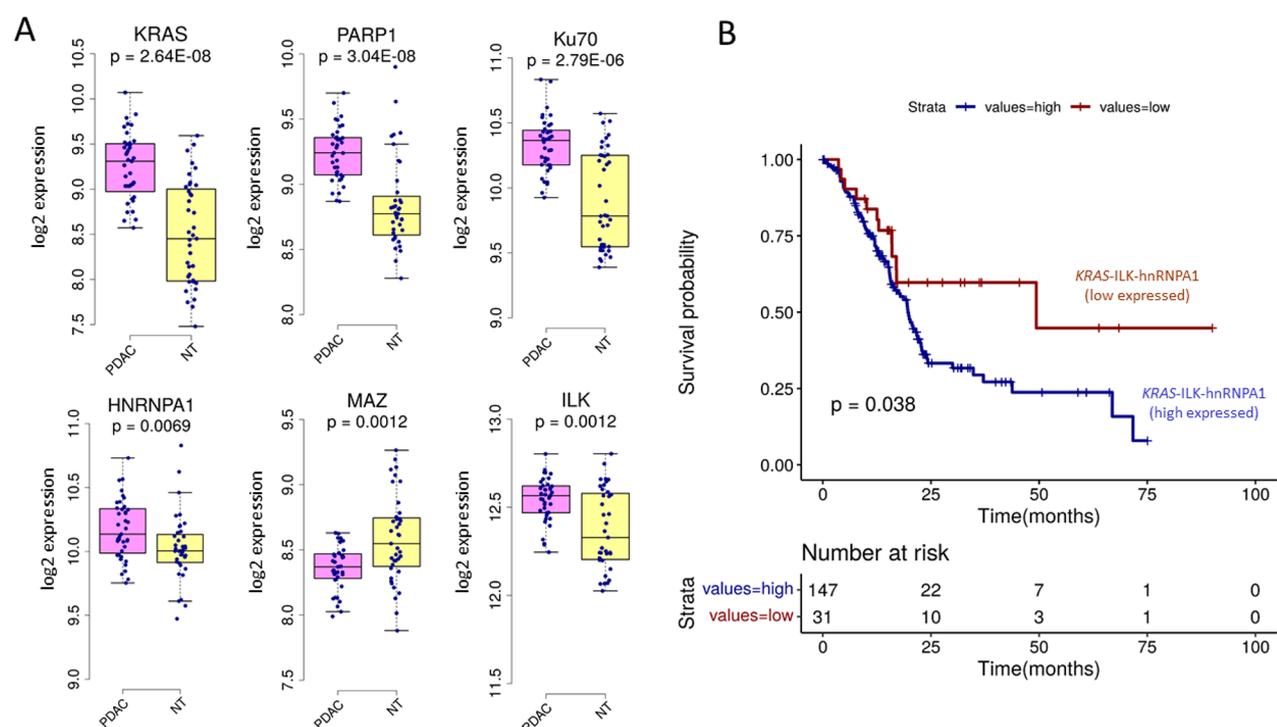


Figure 9. (A) Box plots showing the expression of genes related with KRAS in normal and PDAC pancreatic tissues (yellow and pink, respectively) obtained from the GSE15471 data set. (B) Survival probability of PDAC patients with the KRAS-ILK-hnRNPA1 axis upregulated or downregulated.

is practically unfolded, whereas G9T is only partly unfolded. As observed for the interaction between UP1 and telomeric G4, the UP1 residues showing important shifts upon binding to the KRAS G4 conformers are located in the two RRM domains.

The ability of hnRNPA1/UP1 to unfold G4 DNA suggests that this protein should play an important role in transcription regulation.^{9,21,23,28,43} By using a Panc-1 knockout cell line in which hnRNPA1 was deleted by the CRISPR/Cas9 technology, we found that hnRNPA1 is essential for the transcription of KRAS and for cell growth. Pull-down/Western blot experiments indicate that conformer G25T is a better platform than conformer G9T for the assembly of the transcription preinitiation complex with PARP1, Ku70, MAZ, and hnRNPA1. A growing body of evidence indicates that PDAC cells are addicted to KRAS, which is regulated by the KRAS-ILK-hnRNPA1 axis.^{28,44} Its expression correlates with the clinical outcome of PDAC patients. Kaplan–Meier plots show that the survival probability of PDAC patients with a high expression of the KRAS-ILK-hnRNPA1 axis is significantly lower than that of PDAC patients with a low expression of the axis. Together, the data confirm the central role of KRAS-ILK-hnRNPA1 in the maintenance of PDAC and suggest that hnRNPA1 can be an attractive target for the design of new anticancer drugs against PDAC.

MATERIALS AND METHODS

Oligonucleotides. The oligonucleotides have been purchased from Microsynth-AG, Balgach, Switzerland, or alternatively from Integrated DNA Technologies, Leuven, Belgium. Their sequences are reported in Table 1. DNA concentration was determined from the absorbance at 260 nm of the oligonucleotides diluted in milli Q water using as extinction coefficients 7500, 8500, 15,000, and 12,500 M⁻¹

cm⁻¹ for C, T, A, and G, respectively. The oligonucleotides, including those labeled to Cy-3, were HPLC purified.

Cell Culture, Metabolic Activity, and Proliferation Assay. Normal and hnRNPA1-deleted Panc-1 cells were maintained in exponential growth in Dulbecco's modified Eagle's medium (DMEM) containing 100 U/mL penicillin, 100 mg/mL streptomycin, 20 mM L-glutamine, and 10% fetal bovine serum (Euroclone, Italy). The metabolic activity assay was performed on a 96-well plate by seeding 9×10^3 cells per well. The cells were then treated with resazurin following a standard procedure. Cell growth assay was performed by seeding the cells in a 24-well plate and counting the cells on a cell counter every day for 6 days. Clonogenic assays were carried out with normal and hnRNPA1-deleted Panc-1 cells seeded in DMEM at a very low density and left for a period of 15 days. The colonies of at least 50 cells were counted, and the results were plotted in a histogram.

UV, CD, Fluorescence, and DMS Footprinting Experiments. UV melting was performed by using a Jasco V-750 UV–visible spectrophotometer equipped with a Peltier temperature control system (ETCS-761) (Jasco Europe, Cremella, Italy). The spectra were analyzed with Spectra Manager (Jasco Europe, Cremella, Italy). The oligonucleotides (3 μM) were annealed in 50 mM Na-cacodylate (pH 7.4) and 100 mM KCl (5 min at 95 °C, overnight at RT). The melting curves were recorded at 295 nm in a 0.5 cm path length quartz cuvette, heating (25–95 °C) at a rate of 0.5 °C/min.

CD spectra have been obtained with a JASCO J-600 spectropolarimeter equipped with a thermostated cell holder. CD experiments were carried out with 3 μM oligonucleotides in 50 mM Na-cacodylate (pH 7.4) and 100 mM KCl. Spectra were recorded in 0.5 cm quartz cuvettes. The spectra were calculated with the J-700 Standard Analysis software (Japan Spectroscopic Co., Ltd) and are reported as ellipticity (mdeg)

versus wavelength (nm). Each spectrum was recorded three times and subtracted to the baseline.

DMS footprinting was carried out as previously described.³

Production of Recombinant UP1. The recombinant protein comprising the RRM domains of UP1 (residues 17 to 196) was inserted into a modified pGEX vector containing a GST marker and then transformed on a Petri dish. The expression of UP1 in *E. coli* BL21 (DE3) bacteria was carried out in an LB medium (5 g/L yeast extract, 10 g/L peptone, and 10 g/L NaCl) at 37 °C overnight with ampicillin at 100 µg/mL. Bacteria were then transferred to a TB medium (24 g/L yeast extract, 12 g/L tryptone, 5 g glycerol, and 100 mM phosphate buffer (KH₂PO₄/K₂HPO₄)) supplemented with 100 µg/mL of Amp. For ¹⁵N, ¹³C labeled production, a minimal M9T medium (300 µM CaCl₂, 1 mM MgSO₄, 6 g Na₂HPO₄, 3 g KH₂PO₄, 0.5 g NaCl, 1 mg vitamin B1, 1 g NH₄Cl ¹⁵N, and 2 g glucose ¹³C) was used, always supplemented with 100 µg/mL of Amp. Expression was induced at an OD 600 nm between 1.5 and 2.0 with IPTG at 1 mM, overnight at 17 °C. The bacterial pellets were then recovered by centrifugation at 6500 rpm for 20 min at 4 °C, resuspended with PBS, and incubated under agitation for 30 min with 100 mM PMSF, lysozyme, and 1 M DTT. A lysis by sonication (40%: 45 s on, 45 s off for 4 min and 30 s) was then carried out, the lysate was then ultracentrifuged for 1 h at 4 °C at 42,000 RPM, and the supernatant was collected. Glutathione Sepharose 4B (GE Healthcare) (50% slurry in PBS) was added to the supernatant and incubated for 2 h at 4 °C with a slow shaking. The mix was centrifuged at 2000 rpm for 5 min, and the pellet was washed five times in PBS and eluted with an elution buffer containing 20 mM NaCl, 20 mM reduced glutathione, and 200 mM Tris–HCl (pH 7.5). GST tag was then cleaved from purified UP1 using PreScission protease (1.5 mg/mL) after exchange with a cleavage buffer (200 mM Tris HCl, 20 mM NaCl, 0.5 mM EDTA, and 1 mM DTT) overnight at 4 °C. GST and purified UP1 were then separated by size exclusion chromatography using GF S75 after equilibration in a buffer (50 mM KCl, 10 mM KPi (pH 6.66), and 0.5 mM DTT) overnight. Finally, purification was checked by SDS-PAGE and concentration was determined by measuring absorbance at 280 nm.

NMR Experiments. NMR spectra were recorded on a Bruker Advance III 700 MHz spectrometer equipped with a liquid TXI 1H/13C/15N/2H probe. All samples were prepared in 1X buffer (10 mM K₂HPO₄/KH₂PO₄; 50 mM KCl; pH 6.6) with the addition of 10% D₂O for lock purposes and all spectra were acquired in 3 mm NMR tubes. For kinetics experiments containing UP1 and DNA, the G4 concentration (G9T and G25T) were 184 µM followed with the addition of one molar-equivalent of non-labelled UP1 in 1X buffer in presence of 1 mM DTT. Spectra of G4 mixed with UP1 were recorded at different time periods after addition of UP1 (0, 90 minutes, 1, 2, 3 and 7 days). In the 1D (1H) NMR experiments, the water signal was suppressed using excitation sculpting with gradients (zgesgpppe; d1=2sec; 512 scans; time domain=64k). Samples were maintained at 37 °C between each NMR experiment. Identification of UP1 residues implicated in the interaction with both KRAS32R conformers (G9T and G25T) have been done by using 2D NMR acquisitions with ¹⁵N, ¹³C isotopically enriched samples of UP1 in 1X buffer in presence of 1 mM DTT. We used SOFAST (Band-Selective Optimized-Flip-Angle Short-Transient) HMQC based on 2D H-1/X correlation via Heteronuclear zero and double quantum

with decoupling during acquisition (sfhmqcf3gpqh; d1=0.3sec; 256 scans, F2 (1H) time domain=2k; F1 (15N) time domain=160). Each residue has been identified by the –NH from its backbone connection and assigned using the deposited data from PDB structure 1L3K. Increasing amounts molar fractions of G4 (G9T or G25T) have been successively included (0.25, 0.5, 0.75, 1) and we followed UP1 chemical shift peak shifting after each oligo addition. Shifts have been determined for several peaks using the equation:

$$\delta = \sqrt{(\delta H)^2 + \left(\left(\frac{\delta N}{5}\right)^2\right)} \quad (1)$$

with δH and δN being the chemical shifts in 1H and 15N dimensions respectively. Deviations of the chemical shifts were then plotted in function of the corresponding residue in Origin 8.6.

Isothermal Titration Calorimetry (ITC). ITC experiments were performed using a Microcal ITC200 instrument (Malvern). All experiments were performed at 37 °C. All samples were dialyzed in 1X buffer (10 mM K₂HPO₄/KH₂PO₄ and 50 mM KCl (pH 6.6)) with the addition of 1 mM DTT overnight and thoroughly degassed prior to use. Titrations were conducted with wild-type 32R and G4 conformers G9T and G25T. For the ITC titrations, the sample cell was filled to capacity with a dilute solution of UP1 at 10 µM and titrated with DNA at 50 µM in the same buffer. Titration has been done with 16 injections of 2.5 µL aliquots of the titrant with titrant injections made at 300 s intervals, with 600 rpm for stirring. The integrated heat data were corrected considering the heat of the dilution and blank effects. The corrected data were fit with a binding model by nonlinear regression. The binding isotherms were sigmoidal and well fit with the standard one-site binding model incorporated into the Microcal Origin ITC software.

CRISPR-Cas9 Suppression of hnRNPA1. HnRNPA1-deleted Panc-1 clones were generated by genome editing with the CRISPR/Cas9 system. The genome editing of Panc-1 cells has been carried out by Synthego (CA), which provided us a pool of Panc-1 edited cells. Individual clones were tested by Western blot to verify the deletion of the hnRNPA1 protein. Clones with extremely affected morphologic phenotype were excluded from further experiments.

Nuclear Extract and Biotin–Streptavidin Pull-Down Assay. To obtain nuclear extracts, six plates of 15 cm diameter of Panc-1 cells at a given confluence were washed with PBS and treated with 0.1 mM H₂O₂ in serum-free DMEM-high glucose for 30 min. The cells were collected in a PBS buffer and centrifuged at 800g for 10 min at 4 °C. Then, the cells were resuspended in a hypotonic buffer (10 mM HEPES-KOH (pH 7.9), 1.5 mM MgCl₂, 10 mM KCl, 0.2 mM PMSF, 0.5 mM DTT, 5 mM NaF, and 1 mM Na₃VO₄) and kept in ice for 10 min. Swollen cells were homogenized with a Dounce homogenizer and the nuclei, pelleted by centrifugation, and resuspended in a low-salt buffer (20 mM HEPES-KOH (pH 7.9), 25% glycerol, 1.5 mM MgCl₂, 20 mM KCl, 0.2 mM EDTA, 0.2 mM PMSF, and 0.5 mM DTT). The nuclear proteins were obtained by the addition of a high-salt buffer (low-salt buffer containing 1.2 M KCl). Protein concentration was determined according to the Bradford method. Biotinylated 32R, G9T, and G25T were folded in 50 mM Tris–HCl (pH 7.4) and 100 mM KCl by heating the solutions at 95 °C for 5 min and successive incubation overnight at RT. The

nuclear extract (80 μ g) was incubated for 30 min at RT with 80 nM biotinylated 32R, G9T, or G25T in 20 mM Tris–HCl (pH 7.4), 150 mM KCl, 8% glycerol, 1 mM DTT, 0.1 mM ZnAc, 5 mM NaF, 1 mM Na₃VO₄, and 2.5 ng/ μ L poly[di-dC]. Then Streptavidin MagneSphere Paramagnetic Particles (Promega Italia, Milano, Italy) were added and left to incubate for 30 min at RT. The beads were captured with a magnet and washed three times. The proteins were eluted with Laemmli buffer (4% SDS, 20% glycerol, 10% 2-mercaptoethanol, 0.004% bromophenol blue, and 0.125 M Tris–HCl).

Electrophoresis Mobility Shift Assays (EMSA). Cy5.5-end labeled oligonucleotides 32R, G9T, and G25T were allowed to adopt their structure in 50 mM Tris–HCl (pH 7.4) and 100 mM KCl (heated at 95 °C for 5 min and annealed overnight at RT). Cy5.5-oligonucleotides (50 nM) were treated for 15 min at 25 °C with increasing amounts of hnRNPA1 in 50 mM Tris–HCl (pH 7.4), 50 mM KCl, 2.5 ng/mL poly[di-dC], 1 mM EDTA, 1 mM Na₃VO₄, 5 mM NaF, 0.01% Phosphatase Inhibitor Cocktail I (Merck Life Science, Milano, Italy), 1 mM DTT, and 8% glycerol. The reaction mixtures were incubated for 10 min in ice, loaded in 5% TB (1 \times) polyacrylamide gel, and then run at 300 V, 50 mA, and 30 W for 3 h at 20 °C. After running, the gel was analyzed with the Odyssey CLx Imaging System (Li-COR Biosciences, Lincoln, NE, USA).

Immunoprecipitation Assay. Panc-1 cells were seeded onto 15 cm diameter plates. At 80% confluence, the cells were treated with 0.1 mM H₂O₂ in serum-free DMEM high-glucose medium for 30 min. Then the nuclear proteins were extracted and quantified as described in the **Nuclear Extract and Biotin–Streptavidin Pull-Down Assay** section. For immunoprecipitation, 1.5 mg of Protein A-Dynabeads (ThermoFisher Scientific-Invitrogen, Waltham, MA, USA) was incubated with 3 μ g of anti-PAR (Poly/Mono-ADP Ribose (E6F6A) Rabbit mAb #83732, Cell Signaling Technology, Leiden, The Netherlands), anti-PARP-1 (46D11, Cell Signaling Technology, Leiden, The Netherlands), anti-Ku70 (D10A7, Cell Signaling Technology, Leiden, The Netherlands), anti-MAZ (clone 133.7, IgG mouse, Santa Cruz Biotechnology, Dallas, TX, USA), anti-hnRNPA1 (clone 9H10, IgG mouse, Merck Life Science, Milano, Italy), and IgG Rabbit (ThermoFisher Scientific-Invitrogen, Waltham, MA, USA) as negative control in 20 mM Tris–HCl (pH 7.4), 20 mM KCl, 8% glycerol, 1 mM DTT, and 0.1 mM ZnAc for 15 min at RT. After one wash with the same buffer, 80 μ g was allowed to react with anti-PAR- and IgG rabbit-derivatized Dynabeads for 30 min at RT. The beads were captured with a magnet and washed three times with the same buffer. The proteins were denatured and eluted with Laemmli buffer (4% SDS, 20% glycerol, 10% 2-mercaptoethanol, 0.004% bromophenol blue, and 0.125 M Tris–HCl).

Western Blot Assays. Protein samples were separated in 10% SDS-PAGE and blotted onto the nitrocellulose membrane at 70 V for 2 h. The nitrocellulose membrane was blocked for 1 h with 5% nonfat dried milk in PBS and 0.1% Tween (Merck Life Science, Milano, Italy) at room temperature.

The primary antibodies used were as follows: anti-MAZ (clone 133.7, monoclonal antibody, IgG mouse, Santa Cruz Biotechnology, Dallas, TX, USA), anti-hnRNP A1 (clone 9H10, monoclonal antibody, IgG mouse, Merck Life Science, Milano, Italy), anti-PARP-1 (clone H-300, polyclonal antibody, IgG rabbit, Cell Signaling Technology, Leiden, The Netherlands), anti-PAR (Poly/Mono-ADP Ribose, clone E6F6A,

monoclonal antibody, IgG Rabbit, Cell Signaling Technology, Leiden, The Netherlands), anti-Ku70 (clone 3C3.11, monoclonal antibody, IgG mouse, Cell Signaling Technology, Leiden, The Netherlands), anti-hnRNP M (clone A-12, monoclonal antibody, IgM mouse, Santa Cruz Biotechnology, Dallas, TX, USA), anti-hnRNP F/H (clone 1G11, monoclonal antibody, IgG Mouse, Santa Cruz Biotechnology, Dallas, TX, USA), anti-hnRNP A2/B1 (clone B-7, monoclonal antibody, IgG mouse, Santa Cruz Biotechnology, Dallas, TX, USA), anti-HRAS (clone C-20, polyclonal antibody, IgG rabbit, Santa Cruz Biotechnology, Dallas, TX, USA), anti-NRAS (clone F155-227, monoclonal antibody, IgG mouse, Calbiochem, San Diego, CA, USA), anti-KRAS (clone 3B10-2F2, mouse monoclonal, IgG mouse, Merck Life Science, Milano, Italy), anti-ILK (polyclonal antibody, IgG rabbit, Cusabio Technology LLC, Houston, TX, USA), anti-nucleoporin (polyclonal antibody, IgG Rabbit, Abcam, Cambridge, UK), and anti- β -actin (monoclonal antibody, IgG Mouse, Merck Life Science, Milano, Italy). The membranes were incubated overnight at 4 °C with the primary antibodies, washed with 0.1% Tween in PBS, and then incubated for 1 h with the secondary antibodies conjugated to horseradish peroxidase: anti-mouse IgG (diluted 1:5000), anti-rabbit IgG (diluted 1:5000), and anti-mouse IgM (diluted 1:5000) (Merck Life Science, Milano, Italy). The signal was developed with Super Signal West PICO and FEMTO (Thermo Fisher Scientific, Waltham, MA, USA) and detected with the ChemiDOC XRS, Quantity One 4.6.5 software (Bio-Rad Laboratories, Segrate, (Milano), Italy).

Gene Expression Analysis. Data set GSE15471 was downloaded from GEO.⁴⁵ CEL files were processed using standard tools available within the R affy package.⁴⁶ The normalization step was done with the standard RMA algorithm,⁴⁷ while the Jetset scoring was used to identify the optimal microarray probe set for each gene.⁴⁸ The impact of gene expression on patient survival in the PDAC data set from the Cancer Genome Atlas (TCGA-PAAD) was evaluated. mRNA expression data from 178 samples (normalized by the RNAseq by the Expectation–Maximization (RSEM) method) and patients' clinical data were retrieved from TCGA in May 2021 using the R package cgdscr.⁴⁹ The whole gene signature was taken into account: every patient's median expression value was determined, and all the patients were divided into "high" and "low" expression groups based on the optimal cutoff. This is the value that creates the largest survival separation between groups with the highest significance. For this purpose, we used the `surv_cutpoint` function in `survminer` package.⁵⁰ Overall survival (OS) of the two groups was compared by using the Kaplan–Meier plots, with *p* values calculated via log-rank test, using the R survival package in R.⁵¹

Statistics. Vertical bar graphs report mean values \pm standard error (SE). Statistical analyses were carried out by using the Sigma Plot software. Group differences were analyzed by Student's *t* test or one-way analysis of variance (ANOVA). Groups are considered different when *P* < 0.05.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsomega.1c05538>.

Structures of 32R and oxidized 32R from DMS footprinting; FRET experiments with G25T/G9T and

hnRNPA1; and confocal microscopy of Panc-1 and Nrf2 knockout Panc-1 cells (koA1_1) (PDF)

AUTHOR INFORMATION

Corresponding Authors

Gilmar F. Salgado – ARNA Laboratory, Université de Bordeaux, Pessac 33607, France; orcid.org/0000-0002-0296-5979; Email: gilmar.salgado@u-bordeaux.fr

Luigi E. Xodo – Department of Medicine, Laboratory of Biochemistry, Udine 33100, Italy; orcid.org/0000-0003-3344-7207; Email: luigi.xodo@uniud

Authors

Annalisa Ferino – Department of Medicine, Laboratory of Biochemistry, Udine 33100, Italy

Julien Marqueville – ARNA Laboratory, Université de Bordeaux, Pessac 33607, France

Himanshi Choudhary – Department of Medicine, Laboratory of Biochemistry, Udine 33100, Italy

Giorgio Cinque – Department of Medicine, Laboratory of Biochemistry, Udine 33100, Italy

Coralie Robert – ARNA Laboratory, Université de Bordeaux, Pessac 33607, France

Anne Bourdoncle – ARNA Laboratory, Université de Bordeaux, Pessac 33607, France

Raffaella Picco – Department of Medicine, Laboratory of Biochemistry, Udine 33100, Italy

Jean-Louis Mergny – ARNA Laboratory, Université de Bordeaux, Pessac 33607, France; Laboratoire d'Optique et Biosciences, Ecole Polytechnique, CNRS, INSERM, Institut Polytechnique de Paris, Palaiseau 91128, France; orcid.org/0000-0003-3043-8401

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acsoomega.1c05538>

Funding

This work has been financed by AIRC (Associazione Italiana per la Ricerca sul Cancro, Project No. 19898) and La Ligue contre le Cancer, Comité de Gironde, France.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The work on the 800 MHz NMR spectrometer (IECB) was possible thanks to support from TGIR-RMN-THC Fr3050 CNRS, CNRS UMS3033, Inserm US001, Univ. Bordeaux.

ABBREVIATIONS

KRAS *Kirsten ras*; hnRNPA1 heterogeneous nuclear ribonucleoprotein A1; UP1 unfolding protein 1; ILK integrin linked kinase; PARP1 poly(ADP-ribose) polymerase 1; MAZ Myc associated zinc finger protein; Ku70 ATP-dependent DNA helicase II, 70 kDa subunit; G4 G-quadruplex; TF transcription factor; EMSA electrophoresis mobility shift assay; FRET fluorescence resonance energy transfer; ITC isothermal titration calorimetry; DMS dimethyl sulfate; PDAC pancreatic ductal adenocarcinoma; 8OG 7,8-dihydro-8-oxoguanine

REFERENCES

- (1) Cogo, S.; Xodo, L. E. G4 DNA in ras genes and its potential in cancer therapy. *Biochim. Biophys. Acta* **2016**, *1859*, 663–674.
- (2) Cogo, S.; Xodo, L. E. G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription. *Nucleic Acids Res.* **2006**, *34*, 2536–2549.
- (3) Cogo, S.; Paramasivam, M.; Spolaore, B.; Xodo, L. E. Structural polymorphism within a regulatory element of the human KRAS promoter: Formation of G4-DNA recognized by nuclear proteins. *Nucleic Acids Res.* **2008**, *36*, 3765–3780.
- (4) Cogo, S.; Paramasivam, M.; Membrino, A.; Yokoyama, K. K.; Xodo, L. E. 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.
- (5) Cogo, S.; Zorzet, S.; Rapozzi, V.; Géci, I.; Pedersen, E. B.; Xodo, L. E. MAZ-binding G4-decoy with locked nucleic acid and twisted intercalating nucleic acid modifications suppresses KRAS in pancreatic cancer cells and delays tumor growth in mice. *Nucleic Acids Res.* **2013**, *41*, 4049–4064.
- (6) Cinque, G.; Ferino, A.; Pedersen, E. B.; Xodo, L. E. Role of Poly [ADP-ribose] Polymerase 1 in Activating the Kirsten ras (KRAS) Gene in Response to Oxidative Stress. *Int. J. Mol. Sci.* **2020**, *21*, 6237–6259.
- (7) Lau, J. S.; Baumeister, P.; Kim, E.; Roy, B.; Hsieh, T. Y.; Lai, M.; Lee, A. S. Heterogeneous nuclear ribonucleoproteins as regulators of gene expression through interactions with the human thymidine kinase promote. *J. Cell. Biochem.* **2000**, *79*, 395–406.
- (8) Xia, H. Regulation of γ -fibrinogen chain expression by heterogeneous nuclear ribonucleoprotein A1. *J. Biol. Chem.* **2005**, *280*, 13171–13178.
- (9) Campillos, M.; Lamas, J. R.; García, M. A.; Bullido, M. J.; Valdivieso, F.; Vázquez, J. Specific interaction of heterogeneous nuclear ribonucleoprotein A1 with the -219T allelic form modulates APOE promoter activity. *Nucleic Acids Res.* **2003**, *31*, 3063–3070.
- (10) Tavanez, J. P.; Madl, T.; Kooshapur, H.; Sattler, M.; Valcárcel, J. hnRNP A1 proofreads 3' splice site recognition by U2AF. *Mol. Cell* **2012**, *45*, 314–329.
- (11) Mili, S.; Shu, H. J.; Zhao, Y.; Piñol-Roma, S. Distinct RNP complexes of shuttling hnRNP proteins with pre-mRNA and mRNA: candidate intermediates in formation and export of mRNA. *Mol. Cell Biol.* **2001**, *21*, 7307–7319.
- (12) Izaurralde, E.; Jarmolowski, A.; Beisel, C.; Mattaj, I. W.; Dreyfuss, G.; Fischer, U. A role for the M9 transport signal of hnRNP A1 in mRNA nuclear export. *J. Cell Biol.* **1997**, *137*, 27–35.
- (13) Bonnal, S.; Pileur, F.; Orsini, C.; Parker, F.; Pujol, F.; Prats, A. C.; Vagner, S. Heterogeneous nuclear ribonucleoprotein A1 is a novel internal ribosome entry site trans-acting factor that modulates alternative initiation of translation of the fibroblast growth factor 2 mRNA. *J. Biol. Chem.* **2005**, *280*, 4144–4153.
- (14) Cammas, A.; Pileur, A. F.; Bonnal, S.; Lewis, S. M.; Lévêque, N.; Holcik, M.; Vagner, S. Cytoplasmic relocalization of heterogeneous nuclear ribonucleoprotein A1 controls translation initiation of specific mRNAs. *Mol. Biol. Cell* **2007**, *18*, 5048–5059.
- (15) Ford, L. P.; Wright, W. E.; Shay, J. W. A model for heterogeneous nuclear ribonucleoproteins in telomere and telomerase regulation. *Oncogene* **2002**, *21*, 580–583.
- (16) He, Y.; Smith, R. Nuclear functions of heterogeneous nuclear ribonucleoproteins A/B. *Cell. Mol. Life Sci.* **2009**, *66*, 1239–1256.
- (17) Barraud, P.; Allain, F. H.-T. Solution structure of the two RNA recognition motifs of hnRNP A1 using segmental isotope labeling: how the relative orientation between RRM influences the nucleic acid binding topology. *J. Biomol. NMR* **2013**, *55*, 119–138.
- (18) Xu, R. M.; Jokhan, L.; Cheng, X.; Mayeda, A.; Krainer, A. R. Crystal structure of human UP1, the domain of hnRNP A1 that contains two RNA-recognition motifs. *Structure* **1997**, *5*, 559–570.
- (19) Shamoo, Y.; Krueger, U.; Rice, L. M.; Williams, K. R.; Steitz, T. A. Crystal structure of the two RNA binding domains of human hnRNP A1 at 1.75 Å resolution. *Nat. Struct. Biol.* **1997**, *4*, 215–222.
- (20) Ding, J.; Hayashi, M. K.; Zhang, Y.; Manche, L.; Krainer, A. R.; Xu, R. M. Crystal structure of the two-RRM domain of hnRNP A1

(UP1) complexed with single-stranded telomeric DNA. *Genes Dev.* **1999**, *13*, 1102–1115.

(21) Das, S.; Ward, S. V.; Markle, D.; Samuel, C. E. DNA damage-binding proteins and heterogeneous nuclear ribonucleoprotein A1 function as constitutive KCS element components of the interferon-inducible RNA-dependent protein kinase promoter. *J. Biol. Chem.* **2004**, *279*, 7313–7321.

(22) Fukuda, H.; Katahira, M.; Tsuchiya, N.; Enokizono, Y.; Sugimura, T.; Nagao, M.; Nakagama, H. Unfolding of quadruplex structure in the G-rich strand of the minisatellite repeat by the binding protein UP1. *Proc. Natl. Acad. Sci.* **2002**, *99*, 12685–12690.

(23) Paramasivam, M.; Membrino, A.; Cogoi, S.; Fukuda, H.; Nakagama, H.; Xodo, L. E. Protein hnRNP A1 and its derivative Up1 unfold quadruplex DNA in the human KRAS promoter: implications for transcription. *Nucleic Acids Res.* **2009**, *37*, 2841–2853.

(24) Krüger, A. C.; Raarup, M. K.; Nielsen, M. M.; Kristensen, M.; Besenbacher, F.; Kjems, J.; Birkedal, V. Interaction of hnRNP A1 with telomere DNA G-quadruplex structures studied at the single molecule level. *Eur. Biophys. J.* **2010**, *39*, 1343–1350.

(25) Ghosh, M.; Singh, M. RGG-box in hnRNPA1 specifically recognizes the telomere G-quadruplex DNA and enhances the G-quadruplex unfolding ability of UP1 domain. *Nucl. Acids Res.* **2018**, *46*, 10246–10261.

(26) Ghosh, M.; Singh, M. Structure specific recognition of telomeric repeats containing RNA by the RGG-box of hnRNPA1. *Nucl. Acids Res.* **2020**, *48*, 4492–4506.

(27) Cogoi, S.; Rapozzi, V.; Cauci, S.; Xodo, L. E. Critical role of hnRNP A1 in activating KRAS transcription in pancreatic cancer cells: A molecular mechanism involving G4 DNA. *Biochim. Biophys. Acta, Gen. Subj.* **1861**, *2017*, 1389–1398.

(28) Chu, P. C.; Yang, M. C.; Kulp, S. K.; Salunke, S. B.; Himmel, L. E.; Fang, C. S.; Jadhav, A. M.; Shan, Y. S.; Lee, C. T.; Lai, M. D.; Shirley, L. A.; Bekaii-Saab, T.; Chen, C.-S. Regulation of oncogenic KRAS signaling via a novel KRAS-integrin-linked kinase-hnRNPA1 regulatory loop in human pancreatic cancer cells. *Oncogene* **2016**, *35*, 3897–3908.

(29) Marquevielle, J.; Robert, C.; Lagrabette, O.; Wahid, M.; Bourdoncle, A.; Xodo, L. E.; Mergny, J. L.; Salgado, G. F. Structure of two G-quadruplexes in equilibrium in the KRAS promoter. *Nucleic Acids Res.* **2020**, *48*, 9336–9345.

(30) Paramasivam, M.; Cogoi, S.; Xodo, L. E. Primer extension reactions as a tool to uncover folding motifs within complex G-rich sequences: analysis of the human KRAS NHE. *Chem. Commun.* **2011**, *47*, 4965–4967.

(31) Kerkour, A.; Marquevielle, J.; Ivashchenko, S.; Yatsunyk, L. A.; Mergny, J. L.; Salgado, G. F. High-resolution three-dimensional NMR structure of the KRAS proto-oncogene promoter reveals key features of a G-quadruplex involved in transcriptional regulation. *J. Biol. Chem.* **2017**, *292*, 8082–8091.

(32) Cogoi, S.; Ferino, A.; Miglietta, G.; Pedersen, E. B.; Xodo, L. E. The regulatory G4 motif of the Kirsten ras (KRAS) gene is sensitive to guanine oxidation: implications on transcription. *Nucleic Acids Res.* **2018**, *46*, 661–676.

(33) Myers, J. C.; Shamoo, Y. Human UP1 as a model for understanding purine recognition in the family of proteins containing the RNA recognition motif (RRM). *J. Mol. Biol.* **2004**, *342*, 743–756.

(34) Ferino, A.; Rapozzi, V.; Xodo, L. E. The ROS-KRAS-Nrf2 axis in the control of the redox homeostasis and the intersection with survival-apoptosis pathways: Implications for photodynamic therapy. *J. Photochem. Photobiol. B.* **2020**, *202*, 111672.

(35) Camelo, F.; Le, A. (2021). The Intricate Metabolism of Pancreatic Cancers. In *Advances in Experimental Medicine and Biology*; (pp. 77–88). (Advances in Experimental Medicine and Biology; Vol. 1311). Springer. DOI: 10.1007/978-3-030-65768-0_5.

(36) Eser, S.; Schnieke, A.; Schneider, G.; Saur, D. Oncogenic KRAS signalling in pancreatic. *Cancer. Br. J. Cancer.* **2014**, *111*, 817–822.

(37) Eser, S.; Reiff, N.; Messer, M.; Seidler, B.; Gottschalk, K.; Dobler, M.; Hieber, M.; Arbeiter, A.; Klein, S.; Kong, B.; et al. Selective requirement of PI3K/PDK1 signaling for Kras oncogene-

driven pancreatic cell plasticity and cancer. *Cancer Cell* **2013**, *23*, 406–420.

(38) Soldatenkov, V. A.; Vetcher, A. A.; Duka, T.; Ladame, S. First evidence of a functional interaction between DNA quadruplexes and poly(ADP-ribose) polymerase-1. *ACS Chem. Biol.* **2008**, *3*, 214–219.

(39) Pollock, C. B.; Shirasawa, S.; Sasazuki, T.; Kolch, W.; Dhillon, A. S. Oncogenic K-RAS Is Required to Maintain Changes in Cytoskeletal Organization, Adhesion, and Motility in Colon Cancer Cells. *Cancer Res.* **2005**, *65*, 1251–1250.

(40) Collins, M. A.; Bednar, F.; Zhang, Y.; Brisset, J.-C.; Galbán, S.; Galbán, C. J.; Rakshit, S.; Flannagan, K. S.; Volkan Adsay, N.; Pasca di Magliano, M. Oncogenic Kras is required for both the initiation and maintenance of pancreatic cancer in mice. *J. Clin. Invest.* **2012**, *122*, 639–653.

(41) Ying, H.; Kimmelman, A. C.; Lyssiotis, C. A.; Hua, S.; Chu, G. C.; Fletcher-Sanankone, E.; Locasale, J. W.; Son, J.; Zhang, H.; Coloff, J. L.; et al. Oncogenic Kras maintains pancreatic tumors through regulation of anabolic glucose metabolism. *Cell* **2012**, *149*, 656–670.

(42) Zhu, X.; Luo, W.; Liang, W.; Tang, F.; Bei, C.; Ren, Y.; Qin, L.; Tan, C.; Zhang, Y.; Tan, S. Overexpression and clinical significance of MYC-associated zinc finger protein in pancreatic carcinoma. *Oncotargets Ther.* **2016**, *Volume 9*, 7493–7501. eCollection 2016

(43) Nishikawa, T.; Kuwano, Y.; Takahara, Y.; Nishida, K.; Rokutan, K. HnRNPA1 interacts with G-quadruplex in the TRA2B promoter and stimulates its transcription in human colon cancer cells. *Sci. Rep.* **2019**, *9*, 10276.

(44) Chu, P. C.; Kulp, S. K.; Bekaii-Saab, T.; Chen, C. S. Targeting integrin-linked kinase to suppress oncogenic KRAS signaling in pancreatic cancer. *Small GTPases* **2018**, *9*, 452–456.

(45) Barrett, T.; Troup, D. B.; Wilhite, S. E.; Ledoux, P.; Evangelista, C.; Kim, I. F.; Tomashevsky, M.; Marshall, K. A.; Phillippy, K. H.; Sherman, P. M.; et al. NCBI GEO: archive for functional genomics data sets—10 years on. *Nucleic Acids Res.* **2011**, *39*, D1005–D1010.

(46) Gautier, L.; Cope, L.; Bolstad, B. M.; Irizarry, R. A. affy-analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **2004**, *20*, 307–315.

(47) Irizarry, R. A.; Hobbs, B.; Collin, F.; Beazer-Barclay, Y. D.; Antonellis, K. J.; Scherf, U.; Speed, T. P. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **2003**, *4*, 249–264.

(48) Li, Q.; Birkbak, N. J.; Györffy, B.; Szallasi, Z.; Eklund, A. C. Jsetset: selecting the optimal microarray probe set to represent a gene. *BMC Bioinformatics* **2011**, *12*, 474.

(49) Jacobsen, A. *cgdsr: R-Based API for Accessing the MSKCC Cancer Genomics Data Server (CGDS)*; 2015. <https://cran.r-project.org/web/packages/cgdsr/index.html> (accessed 22 Apr 2016).

(50) Kassambara, A.; Kosinski, M.; Biecek, P.; Fabian, S. *survminer: Drawing Survival Curves using 'ggplot2'* 0.4.9 ed2021; (2021). Available online at: <https://cran.r-project.org/web/packages/survminer/>.

(51) Therneau, T., 2015, *A Package for Survival Analysis in S*; version 2.38, <https://CRAN.R-project.org/package=survival>;