HHS Public Access

Author manuscript

J Am Chem Soc. Author manuscript; available in PMC 2018 June 28.

Published in final edited form as:

J Am Chem Soc. 2017 June 28; 139(25): 8522–8536. doi:10.1021/jacs.7b02046.

Insight into the Complexity of the i-Motif and G-Quadruplex DNA Structures Formed in the *KRAS* Promoter and Subsequent Drug-Induced Gene Repression

Christine E. Kaiser[†], **Natalie A. Van Ert**[†], **Prashansa Agrawal**[†], **Reena Chawla**[§], **Danzhou Yang**[†,‡,§,‖], and **Laurence H. Hurley**[†,‡,§,*]

[†]College of Pharmacy, University of Arizona, Tucson, Arizona 85721, United States

[‡]University of Arizona Cancer Center, University of Arizona, Tucson, Arizona 85724, United States

[§]BIO5 Institute, University of Arizona, Tucson, Arizona 85721, United States

Abstract

Activating *KRAS* mutations frequently occur in pancreatic, colorectal, and lung adenocarcinomas. While many attempts have been made to target oncogenic KRAS, no clinically useful therapies currently exist. Most efforts to target KRAS have focused on inhibiting the mutant protein; a less explored approach involves targeting KRAS at the transcriptional level. The promoter element of the *KRAS* gene contains a GC-rich nuclease hypersensitive site with three potential DNA secondary structure-forming regions. These are referred to as the Near-, Mid-, and Far-regions, on the basis of their proximity to the transcription start site. As a result of transcription-induced negative superhelicity, these regions can open up to form unique DNA secondary structures: G-quadruplexes on the G-rich strand and i-motifs on the C-rich strand. While the G-quadruplexes have been well characterized, the i-motifs have not been investigated as thoroughly. Here we show that the i-motif that forms in the C-rich Mid-region is the most stable and exists in a dynamic equilibrium with a hybrid i-motif/hairpin species and an unfolded hairpin species. The transcription factor heterogeneous nuclear ribonucleoprotein K (hnRNP K) was found to bind selectively to the i-motif species and to positively modulate *KRAS* transcription. Additionally, we identified a benzophenanthridine alkaloid that dissipates the hairpin species and destabilizes the interaction of hnRNP K with the Mid-region i-motif. This same compound stabilizes the three existing *KRAS* G-quadruplexes. The combined effect of the compound on the Mid-region i-motif

[*]Corresponding Author: hurley@pharmacy.arizona.edu.

[‖]Present Address

Purdue University, College of Pharmacy, Medicinal Chemistry and Molecular Pharmacology; Purdue University Center for Cancer Research; and Purdue Institute for Drug Discovery, West Lafayette, Indiana 47907, United States

ORCID

Laurence H. Hurley: 0000-0002-8522-450X

Notes

The authors declare the following competing financial interest(s): Laurence Hurley has a financial interest in the G-quadruplex/i-motif-targeting companies Reglagene and Tetragene.

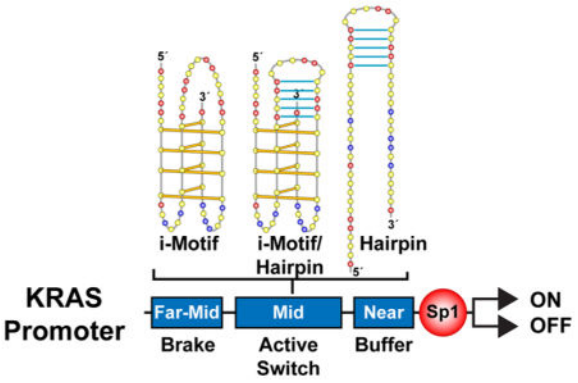
Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.7b02046.

CD spectra of binding of Nitidine to BCL2 and c-MYC i-motifs (PDF)

and the G-quadruplexes leads to downregulation of *KRAS* gene expression. This dual i-motif/G-quadruplex-interactive compound presents a new mechanism to modulate gene expression.

Graphical Abstract



INTRODUCTION

KRAS belongs to the RAS family of small GTPases, which function as switch proteins that cycle between the active GTP-bound form and the inactive GDP-bound form. KRAS primarily signals through the RAF/MEK/ERK, PI3K/PDK1/AKT, and RalGDS pathways, influencing cellular processes such as proliferation, differentiation, and survival.[1–3] Activating mutations in the *KRAS* gene impair GTP hydrolysis and render the protein and its downstream signaling cascades constitutively active, helping to drive tumorigenesis.[4,5] *KRAS* is most frequently mutated in lung (30%), colorectal (44%), and pancreatic (97%) adenocarcinomas.[4,6,7] Suppression of KRAS expression inhibits cellular proliferation in established cancer cell lines and leads to tumor regression in *KRAS*-driven mouse models.[8–14] While there have been many attempts to inhibit mutant KRAS, no clinically relevant therapies currently exist.[15,16]

The promoter element of the human *KRAS* gene (base pairs −500 to 0) is highly GC-rich and contains two nuclease hypersensitive sites.[17–19] Negative superhelicity induced by transcription can result in the local unwinding and opening of GC-rich regions of DNA, allowing for the formation of non-B-form DNA secondary structures, referred to as G-quadruplexes on the G-rich strand and i-motifs on the complementary C-rich strand.[20–23] These structures have been shown to exist in a mutually exclusive fashion: while G-quadruplexes largely act as transcription repressors, recent studies suggest that i-motifs act as transcription activators.[24–27] Within the *KRAS* promoter, three potential DNA secondary structure-forming regions exist. On the basis of their proximity to the transcription start site, they are termed the Near-region (−129 to −159), Mid-region (−174 to −228), and Far-region (−238 to −275) (Figure 1A).[28] (Regions on the G-rich strand will be referred to as Near-G, Mid-G, and Far-G while regions on the C-rich strand will be referred to as Near-C, Mid-C, and Far-C).

The promoter G-quadruplexes have been previously characterized, and both the Near- and Mid-region G-quadruplexes are suggested to play a role in modulating transcription.[28–33] The C-rich region of the murine *KRAS* promoter was one of the first oncogene promoter i-motifs to be characterized.[34]

While the characterization of G-quadruplex structures and their role in cellular processes such as transcription have been the primary focus of past research, more recent work includes a focus on the complementary i-motif structures and their role in transcription regulation.[22,35–38] This expanded focus reflects an effort to better understand the complexity of the entire gene promoter system. The i-motifs comprise two parallel duplexes with intercalated hemiprotonated cytosine$^+$–cytosine (C^+–C) base pairs.[39] These structures are most stable at acidic pH levels; however, they have been shown to exist at neutral pH values under conditions of negative supercoiling and molecular crowding, indicating that they are much more dynamic than their G-quadruplex counterparts.[22,40] A number of i-motifs in the promoter element of various oncogenes have been characterized to date, including *c-MYC*,[41] *BCL2*,[37] *PDGFR-β*,[38] *VEGF*,[42] *RET*,[43] *Rb*,[44] *HRAS*,[45] and *n-myc*.[46] In the case of the *BCL2* i-motif, recent work has shown that this structure can be targeted with small molecules to differentially modulate gene expression.[26] This differential effect on gene expression depends on a dynamic equilibrium between the i-motif and a flexible hairpin, each of which can be independently targeted with steroid molecules. In the case of *n-myc*, both the i-motif and G-quadruplex formed in the promoter have a long lateral loop that is proposed to house a hairpin structure that contributes to the stability of the secondary structure.[46] Most recently it has been shown that it is possible to target both the G-quadruplex and the i-motif with different ligands to downregulate PDGFR-β gene expression.[38] Furthermore, in order to explain the effects of mutation on the G-quadruplexes in the G-rich strand, it is necessary to take into account the corresponding effects on the i-motifs in the C-rich strand.[38]

In this work we show that the *KRAS* Mid-region i-motif is in dynamic equilibrium with a hairpin species. The transcription-activating protein hnRNP K was found to selectively bind to the i-motif species and positively modulate *KRAS* transcription. Furthermore, we identified a dual i-motif/G-quadruplex-interactive compound, the benzophenanthridine plant alkaloid Nitidine (NSC146397), that dissipates the hairpin in the Mid-region hybrid i-motif/hairpin species to destabilize the complex between the Mid-region oligomer and hnRNP K. We propose that the dual mechanism of G-quadruplex stabilization and destabilization of transcription-activator protein:DNA complexes contributes to the observed suppression of *KRAS* gene expression by Nitidine.

MATERIALS AND METHODS

DNA Oligonucleotides

DNA oligonucleotides were purchased from Eurofins MWG Operon (Huntsville, AL) as salt-free sequences. Strand concentrations were calculated using the Beer–Lambert Law: $A = \varepsilon \cdot C l$. Due to the ability of these sequences to form higher order secondary structures, we checked absorbance at 260 nm at 95 °C to ensure that the DNA was in a single-stranded form. The extinction coefficients for each strand were determined using the nearest neighbor

method and were in good agreement with the base composition method. The Mid-C FAM oligomer, with a single 5′-FAM label, and the Mid-C fluorescence resonant energy transfer (FRET) oligomer, with a 5′-FAM and 3′-TAMRA quencher, were purchased from Integrated DNA Technologies (Coralville, IA). The sequences are provided in Table 1.

Compounds

The NCI Diversity Set III and Mechanistic Set of compounds (2480 total) were obtained from the National Institutes of Health, National Cancer Institute Developmental Therapeutics Program (Bethesda, MD). All compounds were dissolved in 100% dimethyl sulfoxide (DMSO) to a stock concentration of 10 mM. Stock compounds were then diluted to working concentrations with deionized water.

Circular Dichroism (CD)

CD analyses were conducted on a Jasco J-810 spectropolarimeter (Jasco, Easton, MD) with a quartz cell of 1 mm optical path length. The various i-motif-forming sequences were diluted to 5 μM strand concentration in 50 mM sodium cacodylate buffer (pH = 4.6, 5.0, 5.5, 6.0, 6.3, 6.5, 6.7, 7.0, 7.5, or 7.9). The G-quadruplex-forming oligonucleotides were prepared at a 5 μM strand concentration in a buffer of 10 mM Tris-HCl (pH 7.4) with 100 mM KCl buffer. Spectra were recorded three times over a wavelength range of 230–350 nm, averaged, smoothed, and baseline corrected to remove signal contribution from buffers. Molar ellipticities were recorded at 286 nm over a temperature range of 5–80 °C and plotted versus temperature to determine the melting temperature (T_m) for i-motif-forming sequences. The Mid-region i-motif mutation analysis was conducted at pH 6.0. Molar ellipticities were recorded at 260 nm for parallel G-quadruplex-forming sequences and 295 nm for antiparallel G-quadruplex-forming sequences over a temperature range of 5–95 °C and plotted versus temperature to determine the T_m. The i-motif-forming samples containing small molecule compounds were prepared with a DNA strand concentration of 5 μM and molar concentration equivalents of compound ranging from 0 to 25 μM in sodium cacodylate buffer (pH = 6.0, 6.5, 7.0). The G-quadruplex-forming samples containing compounds were prepared with a DNA strand concentration of 5 μM and molar concentration equivalents of compound ranging from 0 to 10 μM in a 10 mM Tris-HCl (pH 7.4) + 100 mM KCl buffer.

End-Labeling and Purification of DNA Oligonucleotides

DNA oligonucleotides were 5′-end-labeled with [γ-^{32}P] ATP via incubation with T4 polynucleotide kinase at 37 °C for 1 h, followed by heat inactivation (5 min at 95 °C) of the kinase. Bio-Spin 6 columns (Bio-Rad) were used to remove unincorporated labeled ATP. Oligomers were then mixed with denaturing loading dye (95% formamide, 10 mM NaOH, 10 mM EDTA, 0.1% bromophenol blue) and subjected to denaturing gel electrophoresis (8% PAGE). The bands were excised and the crush and soak method was used to isolate DNA from the gel overnight at 4 °C.

Bromine Footprinting

The bromine footprinting assay was adapted from a previously established protocol.[47] The Mid-C oligomer was 5′-end labeled with [γ-^{32}P] ATP as previously described and incubated in a 10 mM sodium cacodylate buffer (pH 6.0). The samples were brominated via molecular bromine formed in situ by the addition of an equimolar concentration (2 mM) of KBr and $KHSO_5$ for 20 min and terminated by the addition of 80 μL of a 0.3 M sodium acetate and calf thymus DNA (5 μg) solution. Samples were precipitated with ethanol, dried, and resuspended in 50 μL of a 20% piperidine solution. The mixture was then heated to 90 °C for 20 min to induce bromine-specific cleavage, dried, and washed with ddH_2O twice. Samples were resuspended in an alkaline loading dye and visualized on a denaturing sequencing gel (16% PAGE, 7 M urea). A pyrimidine-specific reaction, using hydrazine, and a purine-specific reaction, using formic acid, were performed to generate sequencing makers.

NMR Studies

The DNA oligonucleotides were purchased from Eurofins MWG Operon. The final NMR samples were prepared in a 90% H_2O/10% D_2O solution at varying pH values. The stock solutions of Nitidine were prepared in DMSO-d_6. One-dimensional ^{1}H NMR experiments were performed on a Bruker DRX-600 MHz spectrometer at 25 °C, and the WATERGATE technique was used to suppress the water signal.

Purification of Recombinant hnRNP K

The cDNA of hnRNP K was purchased from Open Biosystems (Thermo Scientific) and subsequently cloned into the pET28a(+) protein expression vector (Novagen). After sequencing analysis to confirm the pET28a–hnRNP K, this expression construct was transformed into Rosetta-gami B (DE3) cells (Novagen). The expression of hnRNP K was induced by 1 mM IPTG (isopropyl β-D-1-thiogalactopyranoside) for 6 h at room temperature. Harvested cells were resuspended in a lysis buffer (50 mM NaH_2PO_4 [pH 8.0], 300 mM NaCl, 1% Triton X-100, 1 mg/mL lysozyme, and 1× protease inhibitor cocktail) (Sigma, #8465) and underwent 10 cycles of the following: incubation on ice for 15 min, vortexing, and sonication. Cell debris was removed by centrifugation at 14 000 rpm for 30 min at 4 °C, and the supernatant was removed and incubated with HisPur Cobalt Resin (Thermo Scientific) while rotating for 1 h at 4 °C to allow for the selective binding of histidine-tagged hnRNP K. The resin was washed with a wash buffer (50 mM NaH_2PO_4 [pH 8.0], 300 mM NaCl, 20 mM imidazole). An elution buffer (50 mM NaH_2PO_4 [pH 8.0], 300 mM NaCl, 250 mM imidazole) was used to separate hnRNP K from the resin. Purified hnRNP K was concentrated and buffer exchanged into 20 mM HEPES-NaOH (pH 7.5), 50 mM NaCl, 20% glycerol, and 0.05% Tween-20 using a Centricon Centrifugal Filter (Millipore). The purity of hnRNP K was confirmed by coomassie blue staining, and the concentration was determined by Bradford assay. The Arizona Proteomics Consortium at the University of Arizona confirmed the identity of the protein using mass spectrometry.

Electrophoretic Mobility Shift Assay (EMSA)

The 5′-FAM-labeled Mid-C oligomer (10 ng) was combined with hnRNP K protein (0, 25, 50, 100, 150, 200, 250 ng) in binding buffer: 20 mM HEPES (pH 6.8), 100 mM KCl, 2 mM MgCl2, 1 mM EDTA, 1 mM DTT, 0.1 μg BSA, 0.1% Tween-20, 10% glycerol, and 10 ng of poly(dI-dC) to a final volume of 11 μL. Binding reactions proceeded for 20 min on ice. The DNA–protein complex and free DNA were visualized following 4% native polyacrylamide gel electrophoresis (0.5× TBE and 1.25% glycerol).

FRET Protein Binding Assay

The dual-labeled Mid-C FRET oligomer (40 nM) was incubated in hnRNP K binding buffer (pH 6.5 or 7.5) for 2 min. hnRNP K (400 nM) was then added, and the reaction was incubated at room temperature for 15 min. The fluorescence intensity was evaluated using a BioTek Synergy HT plate reader with an excitation wavelength of 495 nm and an emission wavelength of 528 nm. The data were corrected by subtracting background signal from the buffer and normalized to the no-protein control. Statistical significance was evaluated using a Student's *t* test, and results were considered statistically significant when $P < 0.05$.

FRET Melt Assay

FRET melt was performed as previously described.[48] The Mid-C FRET probe was prepared at a concentration of 300 nM in 10 mM sodium cacodylate buffer (pH 6.5), heated to 95 °C for 10 min, and allowed to cool to room temperature. The annealed probe was aliquoted (18 μL/well) into a 96-well Mini Skirt PCR Plate and combined with compounds from the NCI Diversity Set III and Mechanistic Set (1 μM) in a final volume of 20 μL. Samples were prepared in single wells according to the NCI plate setup. Fluorescence was read on a Bio-Rad iQ real-time PCR thermocycler from 25 to 94 °C, at a ramp rate of 1 °C/min, with an excitation wavelength of 495 nm and an emission wavelength of 528 nm. NCI plate maps were used to determine compound identity. Controls with no compound were performed in triplicate, and a positive control at pH 5.2 with no compound was also used as a control for i-motif formation. Further compound characterization was performed as above on a pH gradient (pH 6.0, 65, 7.0) with an equal volume of different concentrations of compound (0, 150, 300, 600, 1500 nM).

EMSA with Nitidine

To determine the effect of Nitidine on the association of hnRNP K with the Mid-C oligomer, the oligomer was 5′-end labeled with [γ-^{32}P] ATP as described above. The oligomer (5000 CPM per sample) was combined with equal volumes of Nitidine at varying concentrations (0, 10, 25, 50, and 100 μM) in binding buffer and incubated at room temperature for 3 h. hnRNP K protein (500 ng) and poly(dI-dC) (20 ng) were subsequently added to the mixture, and samples were incubated on ice for 20 min. Free DNA and the protein:DNA complex were visualized following 4% native polyacrylamide gel electrophoresis (0.5× TBE and 1.25% glycerol).

Cell Culture

The AsPC-1, BxPC-3, MIA PaCa-2, PANC-1, and MCF-7 cell lines were obtained from the University of Arizona Cancer Center Experimental Mouse Shared Resource. The AsPC-1, BxPC-3, and MCF-7 cell lines were cultured in RPMI-1640 and MIA PaCa-2 and PANC-1 cells were cultured in DMEM (Corning Cellgro). All media stocks were supplemented with 10% FBS (Fisher Scientific) and 1% penicillin-streptomycin (GE Healthcare). All cell lines were cultured at 37 °C in a humidified atmosphere of 5% CO_2. Cells were counted and assessed for viability using trypan blue prior to use in experimental assays. Cells were assessed for mycoplasma contamination with MycoAlert (Lonza) upon initial culture and at two-month intervals during continuous culture.

siRNA Knockdown Assay

MCF7 (1×10^4 per well of a 12-well plate) cells were seeded in RPMI media supplemented with 10% FBS and allowed to rest overnight. At 50% confluence, cells were transfected with ON-TARGETplus Nontargeting Pool siRNA or ON-TARGETplus hnRNP K SMARTpool siRNA (Dharmacon) diluted to a final concentration of 50 nM using FuGENE HD (Promega) in serum-free RPMI media. For the untreated control, only media with FuGENE HD were used. To determine the knockdown effect of hnRNP K, total RNA was isolated 48 h after transfection using the Qiagen RNeasy Kit (Valencia, CA) according to the manufacturer's protocol. Reverse transcription was performed using the QuantiTect Reverse Transcription Kit (Qiagen) according to the manufacturer's protocol. Real-time PCR was conducted using the Rotor-Gene Q (Qiagen), and Ct values were normalized to GAPDH as an internal control. TaqMan probes used for PCR amplification were hnRNP K (Hs03989611_gH), KRAS (Hs00364284_g1), and GAPDH (Hs02758991_g1). Statistical significance was evaluated using a Student's *t* test, and results were considered statistically significant when $P < 0.05$.

Cellular Cytotoxicity Assay

Cellular cytotoxicity was determined by the MTS colorimetric assay. On the first day cells were harvested, counted, and seeded at 5000 cells per well into 96-well plates. On the second day cells were treated, in triplicate wells, with 10 μL of 10× compound or media for negative control and incubated for 72 h. An MTS/PMS solution (2 mg MTS/mL in PBS with 5% PMS) was prepared, and 20 μL was added to each well followed by incubation for 3 h. Absorbance at 490 nm was measured on a BioTek Synergy HT plate reader. The absorbance of wells containing only media and compound was subtracted from experimental wells to correct for background. Cytotoxicity was determined as the half maximal inhibitory concentration (IC_{50}) of compounds for each cell line.

Real-Time RT-PCR Assay

AsPC-1 cells (0.175×10^6 per well) were plated in a 6-well plate and allowed to rest overnight. On the following day cells were treated with equal volumes of Nitidine (0.5 μM or 5 μM) harvested at the indicated time points. Untreated cells and those treated with DMSO vehicle control were used to determine basal KRAS expression levels. Total RNA was isolated using the Qiagen RNeasy Kit (Valencia, CA) according to the manufacturer's

protocol. Reverse transcription was performed using the QuantiTect Reverse Transcription Kit (Qiagen) according to the manufacturer's protocol. Real-time PCR was conducted using the Rotor-Gene Q (Qiagen). Ct values were normalized to GAPDH and compared to the untreated controls. TaqMan probes used for PCR amplification were KRAS (Hs00364284_g1) and GAPDH (Hs02758991_g1). Statistical significance was evaluated using a Student's *t* test, and results were considered statistically significant when $P < 0.05$.

RESULTS

The Mid-Region i-Motif Is the Most Stable of Those Found in the Promoter Element

In the three *KRAS* promoter regions, runs of three or more consecutive cytosines may play a role in i-motif formation: the Near-C and Far-C have four cytosine runs, whereas the Mid-C has seven cytosine runs (Table 1). As i-motif stability is directly related to pH dependence and temperature, our initial experiment subjected the three single-stranded C-rich DNA oligomers to CD spectral and thermal analyses on a pH gradient.[49,50] The CD spectral signature for an i-motif, which is characteristic of the hemiprotonated C^+–C base pairs, consists of a positive peak within 285–290 nm and a negative peak near 260 nm.[34,41,51,52] At low pH values all three sequences exhibited this characteristic spectrum (Figure 1, B–D). As the pH increased, a decrease in molar ellipticity and a shift toward 277 nm, a signal characteristic of an unordered random coil, can be seen.[53,54] This effect is more apparent in the Near-C and Far-C sequences, suggesting that these i-motifs were less stable than the i-motif in the Mid-C sequence. The transitional pHs for the Near-C (6.6), Mid-C (6.9), and Far-C (6.2) oligomers were determined as the inflection point on a plot of pH versus the maximum molar ellipticity at 286 nm. The T_m of each oligomer at pH 6.0 was determined, on the basis of CD thermal curves, to be 39.8 °C for the Near-C, 50.3 °C for the Mid-C, and 33.1 °C for the Far-C (Table 2). The thermal stability of the Far-C was not able to be assessed above pH 6.0, suggesting that the i-motif formed is relatively weak. Taken together, these data indicate that the Mid-region i-motif is the most stable of the three i-motifs that form in the promoter element.

Potential Intermolecular Interactions among the Three i-Motifs in the *KRAS* Promoter Element

The three DNA secondary structure-forming regions in the *KRAS* proximal promoter are separated by a relatively small number of bases: 17 bases between the Mid- and Near-regions and 12 bases between the Mid- and Far-regions. This close proximity led us to consider the possibility of interactions among the three regions, either in the form of two separate structures interacting or as a single large structure. There is precedence for this idea, as MyoD has been shown to selectively bind to bimolecular G-quadruplex structures in the regulatory sequences of muscle-specific genes.[55]

CD spectral and thermal analyses were used to investigate the formation of i-motifs in oligonucleotides containing both the Mid- and Near-regions (Mid-Near-C) and the Far- and Mid-regions (Far-Mid-C) (Table 1) on a pH gradient. Both sequences displayed CD spectra consistent with i-motif formation, which dissipates at higher pH values, and gave a transitional pH value of 6.5 (Figure 2). T_m values for the composite sequences were

invariant across a concentration range of 5–20 μM (data not shown), suggesting intramolecular folding of the oligomers. The Far-Mid-C sequence had a slightly higher T_m at pH 6.0 of 43.1 °C, compared to 40.8 °C for the Mid-Near-C sequence (Table 2). The stability of each composite oligomer was greater than either the Near or Far counterparts alone, but lower than the Mid. The peak molar ellipticity values (286 nm, pH 4.6) were larger for the composite oligomers compared to the individual oligomers alone (Table 2). The Mid-Near-C oligomer had a molar ellipticity value 2.98 times greater than Mid-C alone and 2.99 times greater than Near-C alone. The Far-Mid-C oligomer had a molar ellipticity value 1.94 times greater than Mid-C alone and 3.84 times greater than Far-C alone. We suggest that, on the basis of the short distances between each set of i-motifs and the much higher molar ellipticities of the composite sequences, the composite forms might play a role in the control of *KRAS* gene expression.

The Mid-Region i-Motif Is in a Dynamic Equilibrium with a Hairpin Species

On the basis of its greater stability and inherent complexity, further characterization efforts were focused on the Mid-region i-motif. To determine the cytosines involved in C^+–C base-pairing, the T_m values of oligomers containing single C-to-T base substitutions were evaluated at pH 6.0 (Figure 3A). Results show that oligomers with mutations in runs II, III, VI, and VII produced the largest decrease in T_m. It is interesting to note that mutating C_{18} to a thymine stabilizes the i-motif by about 8.8 °C. This could be due to reducing the possibility of the cytosines shifting/slipping; however, we did not observe the same effect when mutating C_7 or C_{38}, both of which also border the C^+–C core. The more likely cause of this stabilizing effect is the formation of a T–T base pair between the mutant T_{18} and the natural T_{49}, highlighting the important contribution that interactions among bases of the lateral loops can have on the stability of the i-motif. This mutation analysis suggests a folding pattern for the Mid-region i-motif that consists of eight intercalated C^+–C base pairs and a 7:16:7 loop configuration (Figure 3B). Bromine footprinting at pH 6.0 was also used to investigate those cytosines involved in C^+–C base pairing. In this assay, bromine preferentially modifies cytosines that are not involved in base pairing at their C_5 position, making them more susceptible to piperidine cleavage compared to unmodified bases.[47] The cleavage products can then be visualized via PAGE to determine the protection pattern. Despite carrying out this assay at pH 6.0 and lower, we were unable to show a protection pattern for the cytosines involved in i-motif formation as suggested by the T_m analysis. Instead, the results show protection for cytosines that were posited to be in the 16-base central loop of the Mid-region i-motif (C_{24}, C_{25}, C_{34}, C_{35}, and C_{38}) (Figure 3C). On the basis of the sequence, this protection pattern would suggest the formation of a DNA hairpin with five Watson–Crick base pairs (Figure 3D). This inconsistency between the mutational analysis and the bromine footprinting led us to next investigate the formation of the i-motif by 1D ^{1}H NMR.

We evaluated the 1D ^{1}H NMR spectra of the Mid-C oligomer at varying pH values. The imino regions of the proton NMR spectra show a clear transition from an i-motif species at low pH (5.2) to a hairpin species at high pH (7.8) (Figure 3E). Signals in the imino region between 15 and 16 ppm, characteristic of C^+–C base pairing, clearly indicate the formation of an i-motif at pH 5.2. As the pH increases, this i-motif undergoes a slow transition to a

hairpin species, as shown by the increase in imino proton signals between 12.5 and 13.5 ppm, characteristic of Watson–Crick base pairing. Complete conversion of the i-motif species to the hairpin species is clear at pH 7.8. However, at intermediate pH values (6.0 and 6.7) there is clearly a mixture of equilibrating species. On the basis of these results, we propose a two-species equilibrium: at low pH values an i-motif predominates and at higher pH values an unfolded hairpin that contains only Watson–Crick base pairs predominates (Figure 3F). Multiple isoforms likely exist along this equilibrium, suggested by the presence of broad peaks in the 12.5–13.5 ppm range in the spectra at pH 6.0 and pH 6.7 and especially by the small peak at 13.5 ppm.

Mutations in the Hairpin Suggest That There Is a Composite Intermediate Hybrid i-Motif/Hairpin Species in Equilibrium with the i-Motif and Hairpin Species Previously Identified

The long central loop of the i-motif houses the entire hairpin-forming sequence, which led us to consider whether the i-motif and hairpin could coexist in a single hybrid i-motif/hairpin species. To investigate this idea we evaluated the T_m of Mid-C oligomers containing multiple C-to-T base substitutions located in the upper and lower portions of the hairpin at pH 6.0 (Figure 4A). Mutations in the lower part of the stem (C_{24}, C_{25}, C_{38}; Mid-C Lower Stem T) and upper part of the stem (C_{25}, C_{34}, C_{35}; Mid-C Upper Stem T) decrease the T_m by 5.7 and 5.9 °C respectively, relative to the wild-type sequence. In contrast, mutations of those cytosines in the loop of the hairpin (C_{28}, C_{29}, C_{30}; Mid-C Loop T) resulted in no change in T_m. Knocking out all Watson–Crick base pairing in the stem via C-to-T mutations (Mid-C Full Stem T) or full guanine replacement (Mid-C Full Stem G) also reduced the T_m by 6.3 and 12.6 °C respectively.

We also examined the Mid-C Lower Stem T mutant oligonucleotide by 1D ^{1}H NMR to assess the presence of each species compared to the wild-type oligonucleotide. The distribution is similar for all values except pH 6.8. While the wild-type Mid-C oligonucleotide still has an i-motif species present (Figure 4B), the Mid-C Lower Stem T mutant oligonucleotide has no i-motif species present at pH 6.8 (Figure 4C). Taken together, these results suggest that the presence of a hairpin in the central loop contributes to both the formation and stability of the i-motif. Therefore, we propose that there is an intermediate species consisting of a hybrid i-motif/hairpin that is in equilibrium with the i-motif and hairpin previously identified (Figure 4D). We posit that this coexistence of the i-motif and hairpin contributed to our inability to obtain the expected bromine footprint as noted previously.

hnRNP K Is a Transcription Activator of KRAS and Binds to the Mid-Region i-Motif

hnRNP K is a transcription activator that binds to the poly(C) and poly(CT) sequences in single-stranded DNA and RNA via three K homology (KH) domains.[56–59] This protein has been shown to bind to the i-motif formed in the CT element of the *c-MYC* promoter in addition to two cytosine runs downstream of the i-motif-forming region to positively regulate *c-MYC* transcription. [27,57] Recent studies on both the c-MYC and BCL2 promoter i-motifs have suggested that the preferential interaction of some transcription factors with i-motif structures over single-stranded DNA may be due to the constrained presentation of protein binding sequences located in the lateral loops.[25,27] Similar to the *c-MYC* promoter,

the two lateral loops of the *KRAS* Mid-region i-motif contain CT stretches of DNA, to which hnRNP K is known to bind. We investigated the possible role of hnRNP K in the modulation of KRAS transcription by evaluating the effects of siRNA knockdown. It had previously been shown that siRNA knockdown of hnRNP K results in c-MYC depletion.[60,61] Relative mRNA levels of hnRNP K and KRAS were determined by qPCR after treatment with hnRNP K siRNA. The level of KRAS mRNA decreased upon hnRNP K knockdown (Figure 5A). While the reduction in KRAS mRNA levels upon hnRNP K knockdown is not large, this may be because the knockdown of hnRNP K was only 51% relative to the control. This was also seen with the BCL2/hnRNP LL system,[25] suggesting that other proteins may also be involved in the activation of transcription. In order to gain further support for the involvement of hnRNP K, additional experiments were carried out, as described below.

EMSA (Figure 5B, left) show that in the absence of hnRNP K, the *KRAS* Mid-C oligomer at pH 6.8 separates into two different species, consisting of the hairpin and an equilibrium mixture of the i-motif and the hybrid i-motif/hairpin (Figure 4D). As the i-motif and the hybrid i-motif/hairpin species, with their greater positive charge, would not migrate toward the positive anode as quickly as the hairpin, we propose that this is the higher and more predominant species, whereas the DNA hairpin is the lower species.[25] Upon addition of hnRNP K, the putative i-motif and the hybrid i-motif/hairpin-forming species are preferentially depleted to form the hnRNP K–DNA shifted complex, while the hairpin species is not affected (Figure 5B). The K_D for this interaction was estimated to be 340 nM by quantifying the intensity of EMSA bands. This value is close to reported K_D values for the binding of hnRNP A1 (200 nM)[33] and MAZ (320 nM)[62] to the Near-region G-quadruplex. This suggests that hnRNP K preferentially binds to the KRAS Mid-region i-motif. Since hnRNP K and hnRNP LL have both been shown to unfold their respective i-motifs,[25,27] we also investigated if this binding leads to an unfolding of the Mid-region i-motif. FRET experiments with a dual-labeled (5′-FAM and 3′-TAMRA) Mid-C oligomer and hnRNP K at pHs 6.5 and 7.5 were carried out (Figure 5C). When the i-motif is present at pH 6.5, the addition of hnRNP K increases the fluorescence intensity 1.6-fold, suggesting that protein unfolds the DNA. However, at pH 7.5, when the hairpin predominates, there is a small but significant decrease in fluorescence intensity, indicating a different, perhaps nonspecific, binding event compared to pH 6.5. Together these results suggest that hnRNP K binds to *and* unfolds the *KRAS* Mid-region i-motif in a similar fashion to the i-motifs in the c-MYC and BCL2 promoters.[25,27] In further support of an unfolding mechanism for the Mid-region i-motif that parallels that of hnRNP K on the *c-MYC* i-motif, there is a third KH binding sequence on the 3′-end a similar distance from the Mid-region i-motif.

Identification of Nitidine, a Dual i-Motif/G-Quadruplex-Interactive Compound That Downregulates *KRAS* Gene Expression

Recent work with the BCL2 i-motif highlighted the ability of compounds to affect the dynamic equilibrium between the i-motif and an unfolded DNA hairpin, which resulted in a differential effect on gene transcription.[25,26] Accordingly, we sought to identify compounds capable of affecting the dynamic equilibrium of the Mid-C oligomer using a medium-throughput FRET melt screening assay.[25] A dual-labeled (5′-FAM and 3′-TAMRA) Mid-C

oligomer was used to identify interactive compounds from the NCI/NIH Diversity Set III and Mechanistic Set. Compounds that were able to shift the melting point by 3 °C or more were considered potential hits; this cutoff provided an interactive hit rate of 2.3% (56/2480). Further characterization of the hits led us to select Nitidine (Figure 6A), a known DNA-interactive compound that has been shown to bind preferentially to DNA containing alternating GC base pairs and to partially induce G-quadruplex formation.[63–65] We carried out both FRET melt and CD spectral analyses to gain initial insight into how Nitidine binds to the Mid-C oligomer, which has two potential interaction sites (the lateral loops and the hairpin in the central loop of the i-motif). Because there is a pH-dependent equilibrium among the three different species shown in Figure 4D, we suspected that the greatest effects of Nitidine on the FRET and CD spectra would occur at the pHs favoring one or more of the equilibrating species that served as the target for this drug. At 5 equiv, Nitidine shifted the melting curve of the FRET Mid-C oligomer at varying pH values (Figure 6B). This effect is especially apparent at pH 7.0, where the peak is significantly broadened, suggesting dissipation of the major species and the rise of multiple isoforms. Additional characterization by CD spectral analyses at varying pH values showed that Nitidine produced a concentration-dependent decrease in molar ellipticity and shifted the maximal peak to a lower wavelength compared to the no-compound control (Figure 6C). Again, this effect was more evident at higher pH values. Since the higher pHs favor the hairpin species over the i-motif, both these results suggest that it is the hairpin in the central loop rather than the lateral loops in the i-motif that is the target for Nitidine. We next used 1D ^{1}H NMR to investigate the effect of Nitidine on the Mid-C oligomer. At pH 6.5 the 1D ^{1}H NMR spectra of the imino region for the free DNA showed signal peaks for an i-motif species (15–16 ppm) and a hairpin species (~13 ppm) in equilibrium (Figure 6D). Upon titration with Nitidine, the hairpin peaks at ~13 ppm were markedly decreased, and at 4 equiv they were barely present. Collectively these results suggest that Nitidine specifically destabilizes the hairpin structure within the equilibrating species of the Mid-C oligomer (i.e., in both the hairpin and the hybrid i-motif/hairpin species). We also investigated the effect of Nitidine on the ability of hnRNP K to bind to and interact with the Mid-C oligomer. EMSA analysis revealed that at lower concentrations of Nitidine, a mobility-shifted complex is formed (Figure 6E). However, as the concentration of Nitidine increases, the complex formed becomes more diffuse as Nitidine destabilizes the binding of hnRNP K to the Mid-C oligomer. We believe that this destabilization is due to dissipation of the hybrid i-motif/hairpin species, which we propose is the more favorable substrate for hnRNP K binding. Our assumption is that the stem of the hairpin is disrupted by binding of the Nitidine and that this destabilizes the hybrid structure recognized by hnRNP K. However, a second possibility is that it is the hairpin formation that drives the formation of the i-motif in a somewhat analogous way to how the cooperative folding of the hairpin in the TERT promoter drives the formation of the G-quadruplex. In this case the presence of Nitidine prevents formation of the i-motif, which is required for binding of hnRNP K. Since Nitidine is proposed to bind to the DNA hairpin/i-motif in the KRAS promoter, we next examined the binding selectivity for different promoter i-motifs that either lack a hairpin structure (c-MYC) or have an equilibrating hairpin structure (BCL2) that might be targeted by Nitidine.[26,27] For this determination we used the same CD method as that shown in Figure 6C, since this had shown binding of Nitidine to the hybrid i-motif/hairpin species associated with the KRAS

promoter at pH 7.0. The results show that Nitidine did not bind to either the BCL2 or the c-MYC i-motif at the three different pHs (6.0, 6.5, and 7.0) (Supporting Information Figure S1). Therefore, these results suggest that Nitidine has selectivity for i-motifs that have hairpin species in one of the loops. Since Nitidine disrupts the binding of hnRNP K to the i-motif in the hybrid i-motif/hairpin structure and is known to bind to GC hairpin duplex structures,[63] it is highly likely that the target for Nitidine is this species rather than the hairpin species, similar to that found in equilibrium with the BCL2 i-motif at pH 7.0.

On the basis of Nitidine's reported interactions with G-quadruplexes, [66] we investigated the effect of Nitidine on the three KRAS promoter G-quadruplexes. As previously reported, [28] the Near-G forms an all-parallel G-quadruplex, while the Mid-G and Far-G form mixed parallel/antiparallel G-quadruplexes, with the Far-G being much weaker than the Mid-G, similar to our analyses of the i-motifs. CD spectral analysis of the three G-quadruplex-forming sequences in 100 mM KCl shows that at 1 and 2 equiv Nitidine increases the molar ellipticity at 260 nm for the Near-G (Figure 7A) and at 295 nm for the Mid-G and Far-G (Figure 7, B and C). Evaluation of melting temperature showed that at 2 equiv Nitidine stabilizes the Near-region G-quadruplex by 13.9 °C, the Mid-region G-quadruplex by 10.1 °C, and the Far-region G-quadruplex by 3.9 °C. Thus, it is clear that Nitidine interacts with secondary structures on both the C- and G-rich strands in the KRAS promoter, a unique feature that may be a new strategy for targeting gene transcription, although this potentially raises questions about the relative role of the two target structures. In a more recent paper it has been shown that Chelerythrine, a positional isomer of Nitidine, downregulates KRAS in MCF7 cells, which was attributed to the interaction with a G-quadruplex in the promoter, and this is in agreement with our studies reported here.[67] However, the G-quadruplex interaction studies were carried out with the single murine KRAS G-quadruplex structure, which is different to the three G-quadruplexes in the human KRAS promoter.

We assessed the biological activity of Nitidine by MTS analysis in a panel of pancreatic cancer cell lines (AsPC-1, BxPC-3, MIA PaCa-2, and PANC-1), which resulted in IC_{50} values of 616, 523, 1340, and 3530 nM respectively (Figure 8A). This differential sensitivity may be attributed to the varying status of different genes and pathways among the four pancreatic cancer cell lines. Nitidine significantly downregulated KRAS mRNA levels in AsPC-1 cells at 5.0 μM by 33% at 2 h and 62% at 8 h (Figure 8B).

DISCUSSION

While most research on DNA secondary structures has focused on G-quadruplexes, more attention is now being given to i-motifs on the complementary strand, as both structures are being recognized as modulators of gene transcription. Within the KRAS promoter system all three regions are capable of forming G-quadruplexes on the G-rich strand and i-motifs on the C-rich strand. The Mid-region secondary structures are the most stable on both strands. The Mid-C sequence was shown to be in dynamic equilibrium between an i-motif species (at low pH) and a hairpin species (at high pH). On the basis of mutational studies, the long central loop of the Mid-region i-motif, which is capable of forming a hairpin structure, has been shown to contribute to the stability of the i-motif and may serve as a capping structure. This hairpin structure may be important in the initial formation of the Mid-region i-motif,

similar to the hairpin in the hTERT G-quadruplex, which contributes to the cooperative folding of the structure.[68,69] The transcription activator hnRNP K specifically interacts with the Mid-C i-motif to modulate gene transcription.

Furthermore, we have identified Nitidine as a dual i-motif/G-quadruplex-interactive compound capable of downregulating *KRAS* gene expression. In the C-rich strand Nitidine is able to dissipate the hairpin and hybrid i-motif/hairpin species of the Mid-C oligomer and destabilize the binding of hnRNP K to the Mid-C oligomer. The disruption of this interaction is proposed to be one mechanism by which Nitidine downregulates *KRAS* gene transcription. Additionally, Nitidine stabilizes all three G-quadruplexes in the *KRAS* promoter, thereby providing a second potential mechanism for downregulating *KRAS* gene expression. In this respect, Nitidine represents the first compound proposed to have a dual mechanism of action by targeting both the G-quadruplex and i-motif in a promoter element. While we have investigated the effect of Nitidine on the Mid-region i-motif and G-quadruplex separately, it would be valuable to evaluate the effect of the compound on these structures in the same system to gain a better understanding of what may occur in a cellular context. We propose further studies involving single-molecule laser tweezer experiments as previously described.[26,27,70] In this way we would first examine the effect of Nitidine on each strand of the Mid-region and its associated DNA secondary structure. This will be followed by an investigation of the duplex Mid-region, from which both structures can form, in the presence of hnRNP K. This set up will allow us to simultaneously examine the effect of G-quadruplex stabilization by Nitidine and the disruptive effect of Nitidine on the binding of hnRNP K.

In the BCL2 and c-MYC promoters, studies have shown that the G-quadruplex and i-motif exist in a mutually exclusive fashion.[26,27] Assuming this holds true for the KRAS promoter system, Nitidine would interact with only one secondary structure at a given time. In both cases Nitidine-mediated stabilization of the G-quadruplex and destabilization of the transcription-activating protein hnRNP K and the i-motif complex work to turn off transcription. Therefore, if a mutually exclusive model exists between the G-quadruplex and i-motif in the Mid-region, these Nitidine-mediated effects are complementary rather than antagonistic. Additionally, since Nitidine can bind to all three G-quadruplexes, this compound will likely affect the overall equilibrium of the three regions in the KRAS promoter system, especially in view of our finding that the largest increase in melting point is exhibited by the binding of Nitidine to the Near-region G-quadruplex. Overall, we have identified a compound that is able to turn off KRAS gene transcription most probably via two mechanisms: first by destabilizing the interaction of the Mid-region i-motif with the transcription activator hnRNP K and second by stabilizing the G-quadruplexes in the KRAS promoter.

The KRAS proximal promoter is a complex and dynamic system, as it harbors three DNA secondary structure-forming regions in close proximity. Due to their close clustering, adjacent regions may interact to form larger, stable intramolecular i-motif species. In particular, for the Far-Mid-C sequence the significant increase in molecular ellipticity of the longer oligomer containing both regions over the sum of the individual Far-C and Mid-C regions (Table 2), together with the observation that the T_m values for the composite

sequence were invariant across a concentration range, suggests intramolecular folding of a composite Far-Mid-region i-motif that is distinct from the independent Far-C and Mid-C regions. Thus, this composite species may also play a role in the regulation of *KRAS* gene expression.

We have recently shown that the degree of transcription-induced negative superhelicity plays a role in *c-MYC* transcription activation via a mechanism in which increased negative supercoiling results in access of hnRNP K to an additional KH binding sequence and a higher transcription firing rate.[27] A similar array of three KH binding elements is present in the Mid-C KRAS promoter element, but whether this mechanism operates in this tandem set of three DNA secondary structure-forming elements is not clear. The presence of three tandem G-quadruplex/i-motif-forming sequences in the KRAS promoter begs the question of why they are all conserved with at least two configurations having been shown to be important in KRAS transcription control.[28] We propose that the population dynamics of the three DNA secondary structure-forming regions could provide a "fail-safe" mechanism to regulate transcription and prevent over-expression of KRAS, as the protein product of this gene plays an important role in numerous cellular processes, including cell proliferation, differentiation, and survival.[71] We propose a model for the functional roles of the DNA secondary structure-mediated control of *KRAS* transcription (Figure 9). First we propose that the Near-region predominates under basal conditions and acts as a "buffer" that prevents access to the Mid-region until hnRNP A1 dissolves the G-quadruplex structure and converts it back to duplex (Figure 9A).[33] In accord with our proposed model shown in Figure 9, the action of hnRNP A1 in dissolving the Near-region G-quadruplex structure causes transcription activation by allowing access to the Mid-region and the associated activating i-motif–hnRNP K complex.[72] During active transcription, the Mid-region element is engaged and acts as an "active switch" (Figure 9B). The more stable G-quadruplex predominates until hnRNP K is present and can actively compete with the protein that stabilizes the G-quadruplex (ON/OFF). This model assumes a mutual exclusivity for the G-quadruplex and i-motif, as has been previously shown for the secondary DNA structures in the BCL2 and c-MYC promoters.[25,27] Finally, because the Far-region G-quadruplex and i-motif are much less stable than the similar Mid- and Near-region structures, we propose that under the highest level of Sp1-induced supercoiling the Far-Mid composite intramolecular species will predominate due to its greater stability and act as a "brake" to prevent access of hnRNP K to the Mid-region i-motif (Figure 9C) (OFF). This tandem system shown in Figure 9 relies on a three-stage mechanosensor mechanism and has similarities to the cruise control system present in the c-MYC promoter,[73] which also tightly controls the transcription firing rate. In this case the mechanosensor system uses two competing proteins (the far upstream sequence element (FUSE)–binding protein (FBP) and the FBP interacting repressor (FIR)) in conjunction with a target sequence called FUSE to control the transcription firing rate.[74] However, for c-MYC an ON/OFF switch located in the NHE III_1 is encoded in a separate element downstream of the FUSE element, which also responds to torsional stress.[27] Thus, we propose that nature has evolved different dynamic systems to respond in a real-time way to torsional stress imposed by transcriptionally induced negative superhelicity that include safeguards to prevent aberrant transcription activation and ensure a precise transcription firing rate necessary for cellular growth. While we admit that this model is still tentative, in

the absence of identification of the proteins that bind to the different secondary DNA structures, as has been achieved with the c-MYC system,[27] we believe it is important to propose how these apparently diverse transcription control systems have underlying general principles centered around variable torsional stress imposed by transcription activation by different levels of Sp1.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. David Bishop for his considerable contribution in the preparation and editing of the final version of the text and figures presented in the article. We also acknowledge Dr. Hyun-Jin Kang for valuable technical assistance. This research was supported by the National Institutes of Health (T32 GM008804), an American Chemical Society Division of Medicinal Chemistry MEDI Predoctoral Fellowship, and the National Foundation for Cancer Research (VONHOFF-15-01).

References

1. Castellano E, Santos E. Genes Cancer. 2011; 2:216. [PubMed: 21779495]
2. Eser S, Schnieke A, Schneider G, Saur D. Br J Cancer. 2014; 111:817. [PubMed: 24755884]
3. Karnoub AE, Weinberg RA. Nat Rev Mol Cell Biol. 2008; 9:517. [PubMed: 18568040]
4. Friday BB, Adjei AA. Biochim Biophys Acta, Rev Cancer. 2005; 1756:127.
5. Pylayeva-Gupta Y, Grabocka E, Bar-Sagi D. Nat Rev Cancer. 2011; 11:761. [PubMed: 21993244]
6. Baines AT, Xu D, Der CJ. Future Med Chem. 2011; 3:1787. [PubMed: 22004085]
7. Cox AD, Fesik SW, Kimmelman AC, Luo J, Der CJ. Nat Rev Drug Discovery. 2014; 13:828. [PubMed: 25323927]
8. Brummelkamp TR, Bernards R, Agami R. Cancer Cell. 2002; 2:243. [PubMed: 12242156]
9. Chin L, Tam A, Pomerantz J, Wong M, Holash J, Bardeesy N, Shen Q, O'Hagan R, Pantginis J, Zhou H, Horner JW II, Cordon-Cardo C, Yancopoulos GD, DePinho RA. Nature. 1999; 400:468. [PubMed: 10440378]
10. Collins MA, Bednar F, Zhang Y, Brisset J-C, Galbán S, Galbán CJ, Rakshit S, Flannagan KS, Adsay NV, Pasca di Magliano M. J Clin Invest. 2012; 122:639. [PubMed: 22232209]
11. Collins MA, Brisset JC, Zhang Y, Bednar F, Pierre J, Heist KA, Galbán CJ, Galbán S, di Magliano MP. PLoS One. 2012; 7:e49707. [PubMed: 23226501]
12. Podsypanina K, Politi K, Beverly LJ, Varmus HE. Proc Natl Acad Sci U S A. 2008; 105:5242. [PubMed: 18356293]
13. Singh A, Greninger P, Rhodes D, Koopman L, Violette S, Bardeesy N, Settleman J. Cancer Cell. 2009; 15:489. [PubMed: 19477428]
14. Ying H, Kimmelman AC, Lyssiotis CA, Hua S, Chu GC, Fletcher-Sananikone E, Locasale JW, Son J, Zhang H, Coloff JL, Yan H, Wang W, Chen S, Viale A, Zheng H, Paik J-h, Lim C, Guimaraes AR, Martin ES, Chang J, Hezel AF, Perry SR, Hu J, Gan B, Xiao Y, Asara JM, Weissleder R, Wang YA, Chin L, Cantley LC, DePinho RA. Cell. 2012; 149:656. [PubMed: 22541435]
15. Huynh MV, Campbell SL. Mini-Rev Med Chem. 2016; 16:383. [PubMed: 26423694]
16. Wang Y, Kaiser CE, Frett B, Li H-y. J Med Chem. 2013; 56:5219. [PubMed: 23566315]
17. Jordano J, Perucho M. Nucleic Acids Res. 1986; 14:7361. [PubMed: 3763406]
18. Jordano J, Perucho M. Oncogene. 1988; 2:359. [PubMed: 3283654]
19. Yamamoto F, Perucho M. Oncogene Res. 1988; 3:125. [PubMed: 3067187]
20. Kouzine F, Gupta A, Baranello L, Wojtowicz D, Ben-Aissa K, Liu J, Przytycka TM, Levens D. Nat Struct Mol Biol. 2013; 20:396. [PubMed: 23416947]

21. Kouzine F, Sanford S, Elisha-Feil Z, Levens D. Nat Struct Mol Biol. 2008; 15:146. [PubMed: 18193062]
22. Sun D, Hurley LH. J Med Chem. 2009; 52:2863. [PubMed: 19385599]
23. Zhang C, Liu H-H, Zheng K-W, Hao Y-h, Tan Z. Nucleic Acids Res. 2013; 41:7144. [PubMed: 23716646]
24. Cui Y, Kong D, Ghimire C, Xu C, Mao H. Biochemistry. 2016; 55:2291. [PubMed: 27027664]
25. Kang HJ, Kendrick S, Hecht SM, Hurley LH. J Am Chem Soc. 2014; 136:4172. [PubMed: 24559432]
26. Kendrick S, Kang HJ, Alam MP, Madathil MM, Agrawal P, Gokhale V, Yang D, Hecht SM, Hurley LH. J Am Chem Soc. 2014; 136:4161. [PubMed: 24559410]
27. Sutherland C, Cui Y, Mao H, Hurley LH. J Am Chem Soc. 2016; 138:14138.
28. Morgan RK, Batra H, Gaerig VC, Hockings J, Brooks TA. Biochim Biophys Acta, Gene Regul Mech. 2016; 1859:235.
29. Cogoi S, Paramasivam M, Filichev V, Géci I, Pedersen EB, Xodo LE. J Med Chem. 2009; 52:564. [PubMed: 19099510]
30. Cogoi S, Paramasivam M, Spolaore B, Xodo LE. Nucleic Acids Res. 2008; 36:3765. [PubMed: 18490377]
31. Cogoi S, Xodo LE. Nucleic Acids Res. 2006; 34:2536. [PubMed: 16687659]
32. Paramasivam M, Cogoi S, Xodo LE. Chem Commun (Cambridge, U K). 2011; 47:4965.
33. Paramasivam M, Membrino A, Cogoi S, Fukuda H, Nakagama H, Xodo LE. Nucleic Acids Res. 2009; 37:2841. [PubMed: 19282454]
34. Manzini G, Yathindra N, Xodo LE. Nucleic Acids Res. 1994; 22:4634. [PubMed: 7984411]
35. Brazier JA, Shah A, Brown GD. Chem Commun (Cambridge, U K). 2012; 48:10739.
36. Choi J, Kim S, Tachikawa T, Fujitsuka M, Majima T. J Am Chem Soc. 2011; 133:16146. [PubMed: 21882887]
37. Kendrick S, Akiyama Y, Hecht SM, Hurley LH. J Am Chem Soc. 2009; 131:17667. [PubMed: 19908860]
38. Brown RV, Wang T, Chappeta R, Wu G, Onel B, Chawla R, Quijada H, Camp SM, Chiang ET, Lassiter QR, Lee C, Phanse S, Turnidge MA, Zhao P, Garcia JGN, Gokhale V, Yang D, Hurley LH. J Am Chem Soc. 2017; 139:7456–7475. [PubMed: 28471683]
39. Gehring K, Leroy JL, Guéron M. Nature. 1993; 363:561. [PubMed: 8389423]
40. Cui J, Waltman P, Le VH, Lewis EA. Molecules. 2013; 18:12751. [PubMed: 24132198]
41. Simonsson T, Pribylova M, Vorlícková M. Biochem Biophys Res Commun. 2000; 278:158. [PubMed: 11071868]
42. Guo K, Gokhale V, Hurley LH, Sun D. Nucleic Acids Res. 2008; 36:4598. [PubMed: 18614607]
43. Guo K, Pourpak A, Beetz-Rogers K, Gokhale V, Sun D, Hurley LH. J Am Chem Soc. 2007; 129:10220. [PubMed: 17672459]
44. Xu Y, Sugiyama H. Nucleic Acids Res. 2006; 34:949. [PubMed: 16464825]
45. Miglietta G, Cogoi S, Pedersen EB, Xodo LE. Sci Rep. 2016; 5:18097.
46. Benabou S, Ferreira R, Aviñó A, González C, Lyonnais S, Solà M, Eritja R, Jaumot J, Gargallo R. Biochim Biophys Acta, Gen Subj. 2014; 1840:41.
47. Ross SA, Burrows CJ. Tetrahedron Lett. 1997; 38:2805.
48. De Cian A, Guittat L, Kaiser M, Saccà B, Amrane S, Bourdoncle A, Alberti P, Teulade-Fichou MP, Lacroix L, Mergny JL. Methods. 2007; 42:183. [PubMed: 17472900]
49. Jin KS, Shin SR, Ahn B, Rho Y, Kim SJ, Ree M. J Phys Chem B. 2009; 113:1852. [PubMed: 19173566]
50. Phan AT, Mergny JL. Nucleic Acids Res. 2002; 30:4618. [PubMed: 12409451]
51. Pataskar SS, Dash D, Brahmachari SK. J Biomol Struct Dyn. 2001; 19:307. [PubMed: 11697735]
52. Vorlíčková M, Kejnovská I, Bednářová K, Renčiuk D, Kypr J. Chirality. 2012; 24:691. [PubMed: 22696273]
53. Cantor CR, Warshaw MM, Shapiro H. Biopolymers. 1970; 9:1059. [PubMed: 5449435]

54. Mathur V, Verma A, Maiti S, Chowdhury S. Biochem Biophys Res Commun. 2004; 320:1220. [PubMed: 15249220]
55. Etzioni S, Yafe A, Khateb S, Weisman-Shomer P, Bengal E, Fry M. J Biol Chem. 2005; 280:26805. [PubMed: 15923190]
56. Michelotti EF, Michelotti GA, Aronsohn AI, Levens D. Mol Cell Biol. 1996; 16:2350. [PubMed: 8628302]
57. Takimoto M, Tomonaga T, Matunis M, Avigan M, Krutzsch H, Dreyfuss G, Levens D. J Biol Chem. 1993; 268:18249. [PubMed: 8349701]
58. Tomonaga T, Levens D. J Biol Chem. 1995; 270:4875. [PubMed: 7876260]
59. Tomonaga T, Levens D. Proc Natl Acad Sci U S A. 1996; 93:5830. [PubMed: 8650178]
60. Lynch M, Chen L, Ravitz MJ, Mehtani S, Korenblat K, Pazin MJ, Schmidt EV. Mol Cell Biol. 2005; 25:6436. [PubMed: 16024782]
61. Wen F, Shen A, Shanas R, Bhattacharyya A, Lian F, Hostetter G, Shi J. Ann Surg Oncol. 2010; 17:2619. [PubMed: 20499280]
62. Cogoi S, Zorzet S, Rapozzi V, Geci I, Pedersen EB, Xodo LE. Nucleic Acids Res. 2013; 41:4049. [PubMed: 23471001]
63. Bai LP, Zhao ZZ, Cai Z, Jiang ZH. Bioorg Med Chem. 2006; 14:5439. [PubMed: 16730995]
64. Yang S, Xiang J, Yang Q, Li Q, Zhou Q, Zhang X, Tang Y, Xu G. Chin J Chem. 2010; 28:771.
65. Yang S, Xiang J, Yang Q, Zhou Q, Zhang X, Li Q, Tang Y, Xu G. Fitoterapia. 2010; 81:1026. [PubMed: 20624448]
66. Ghosh S, Jana J, Kar RK, Chatterjee S, Dasgupta D. Biochemistry. 2015; 54:974. [PubMed: 25566806]
67. Jana J, Mondal S, Bhattacharjee P, Sengupta P, Roychowdhury T, Saha P, Kundu P, Chatterjee S. Sci Rep. 2017; 7:40706. [PubMed: 28102286]
68. Kang HJ, Cui Y, Yin H, Scheid A, Hendricks WPD, Schmidt J, Sekulic A, Kong D, Trent JM, Gokhale V, Mao H, Hurley LH. J Am Chem Soc. 2016; 138:13673.
69. Yu Z, Gaerig V, Cui Y, Kang H, Gokhale V, Zhao Y, Hurley LH, Mao H. J Am Chem Soc. 2012; 134:5157. [PubMed: 22372563]
70. Cui Y, Koirala D, Kang H, Dhakal S, Yangyuoru P, Hurley LH, Mao H. Nucleic Acids Res. 2014; 42:5755. [PubMed: 24609386]
71. Adjei AA. J Natl Cancer Inst. 2001; 93:1062. [PubMed: 11459867]
72. Chu PC, Yang MC, Kulp SK, Salunke SB, Himmel LE, Fang CS, Jadhav AM, Shan YS, Lee CT, Lai MD, Shirley LA, Bekaii-Saab T, Chen CS. Oncogene. 2016; 35:3897. [PubMed: 26616862]
73. Michelotti GA, Michelotti EF, Pullner A, Duncan RC, Eick D, Levens D. Mol Cell Biol. 1996; 16:2656. [PubMed: 8649373]
74. Liu J, Kouzine F, Nie Z, Chung HJ, Elisha-Feil Z, Weber A, Zhao K, Levens D. EMBO J. 2006; 25:2119. [PubMed: 16628215]

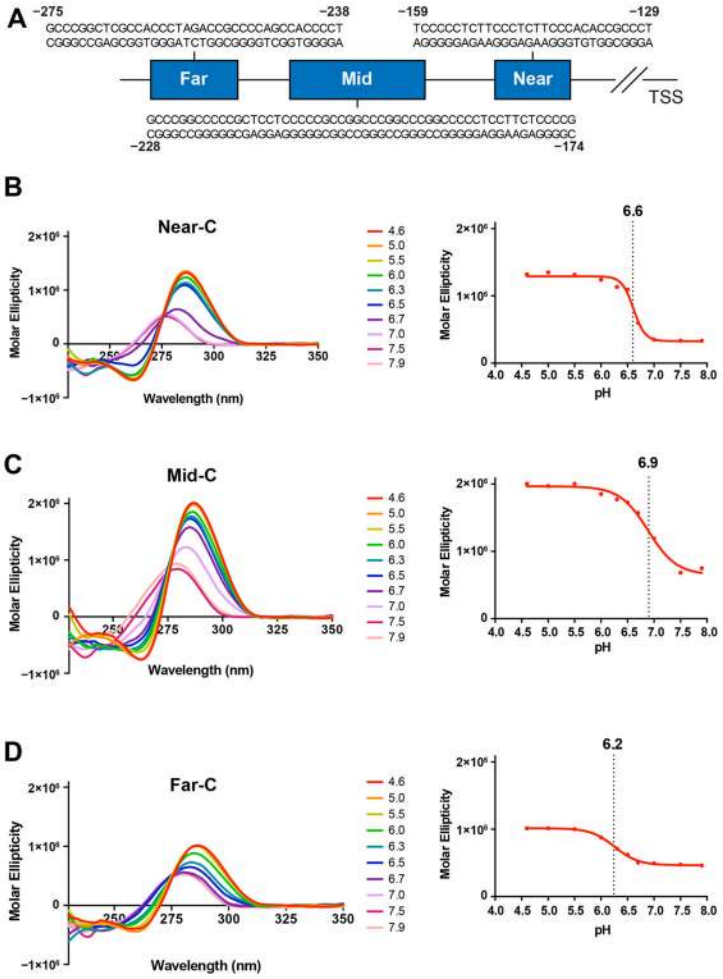


Figure 1.

(A) Schematic highlighting three potential DNA secondary structure-forming regions within the promoter element of the human *KRAS* gene relative to the transcription start site (TSS). (B) CD spectra of the Near-C (B), Mid-C (C), and Far-C (D) oligomers at varying pH values with respective plots of the pH versus molar ellipticity at 286 nm used to determine the transitional pH.

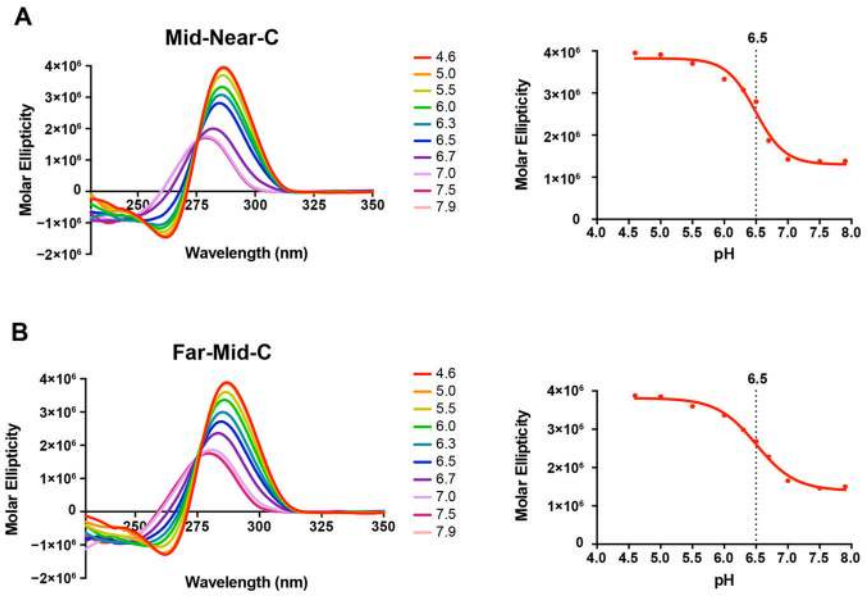


Figure 2.

CD spectra for the Mid-Near-C (A) and Far-Mid-C (B) oligomers at varying pH values with respective plots of the pH versus molar ellipticity at 286 nm used to determine the transitional pH.

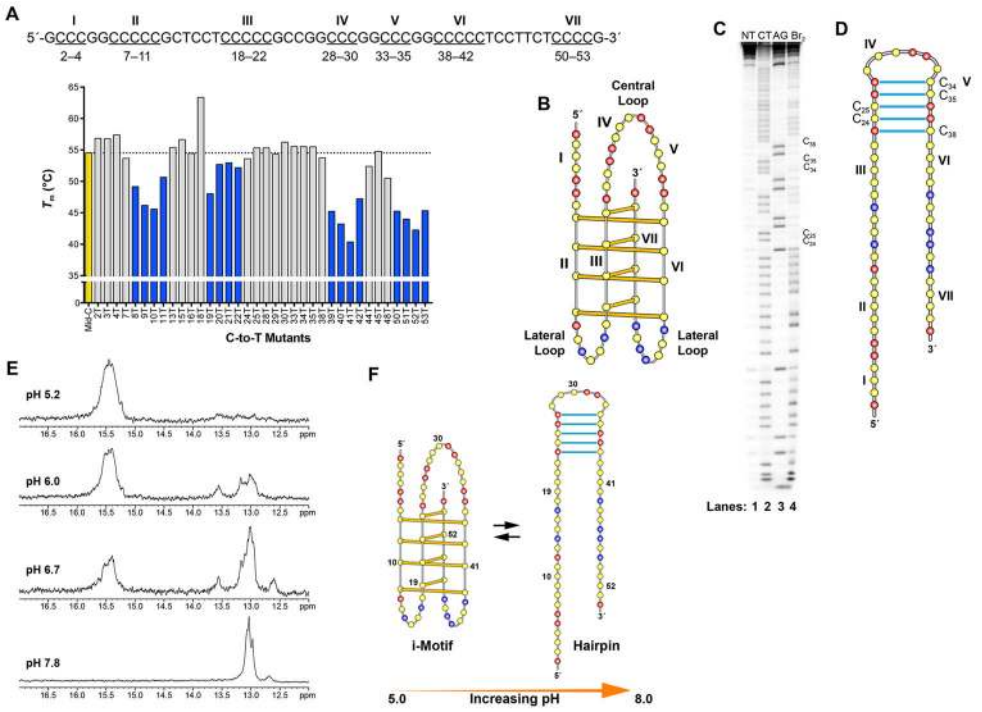


Figure 3.

(A) Histogram representation of the CD melting temperatures for the Mid-C single cytosine-to-thymine mutants at pH 6.0. The wild-type sequence is shown above with all seven cytosine runs underlined. (B) Illustration of the proposed 7:16:7 folding pattern for the Mid-region i-motif based on mutant thermal analyses with C-runs indicated by roman numerals (red = guanine, yellow = cytosine, blue = thymine). (C) Bromine footprint protection pattern of the Mid-C oligomer at pH 6.0. Lanes 1–3 are controls: no treatment (NT), pyrimidine sequencing (CT), and purine sequencing (AG). Lane 4 represents the Br_2 footprinting pattern. (D) Illustration of the proposed folding pattern for the Mid-C hairpin loop based on bromine footprinting, with GC base pairing indicated by blue lines. (E) The imino proton region of the 1D ^1H NMR spectra of free Mid-C oligomer at 25 °C on a pH gradient. (F) Proposed schematic for equilibrating i-motif and hairpin species in the Mid-C oligomer. The i-motif predominates at low pH while the hairpin predominates at higher pH.

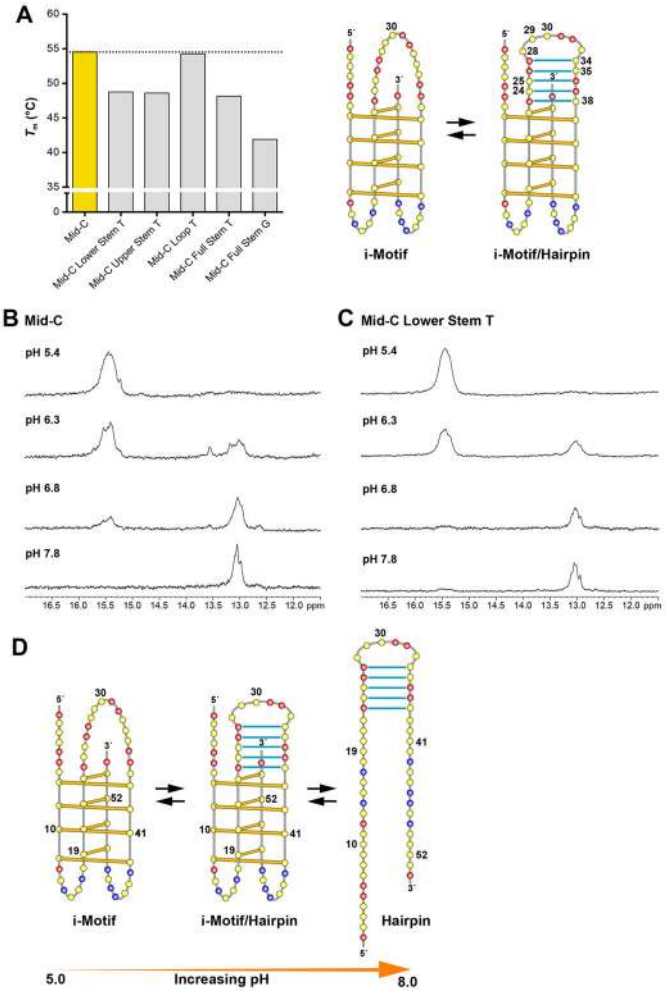


Figure 4.

(A) Histogram representation of the CD melting temperatures for the Mid-C multiple cytosine-to-thymine hairpin mutants at pH 6.0. Folding patterns to the right show the equilibrating species, with the location of mutated bases shown in the hybrid i-motif/hairpin species. (B) The imino proton region of the 1D ^{1}H NMR spectra at 25 °C and varying pH values for the Mid-C (B) and Mid-C Lower Stem T (C) oligomers. Spectra at pH 6.8 highlight the persistence of the i-motif species in the wild-type sequence relative to the mutant sequence. (D) Proposed schematic for the equilibrating i-motif, the hybrid i-motif/hairpin, and the hairpin species in the Mid-C oligomer.

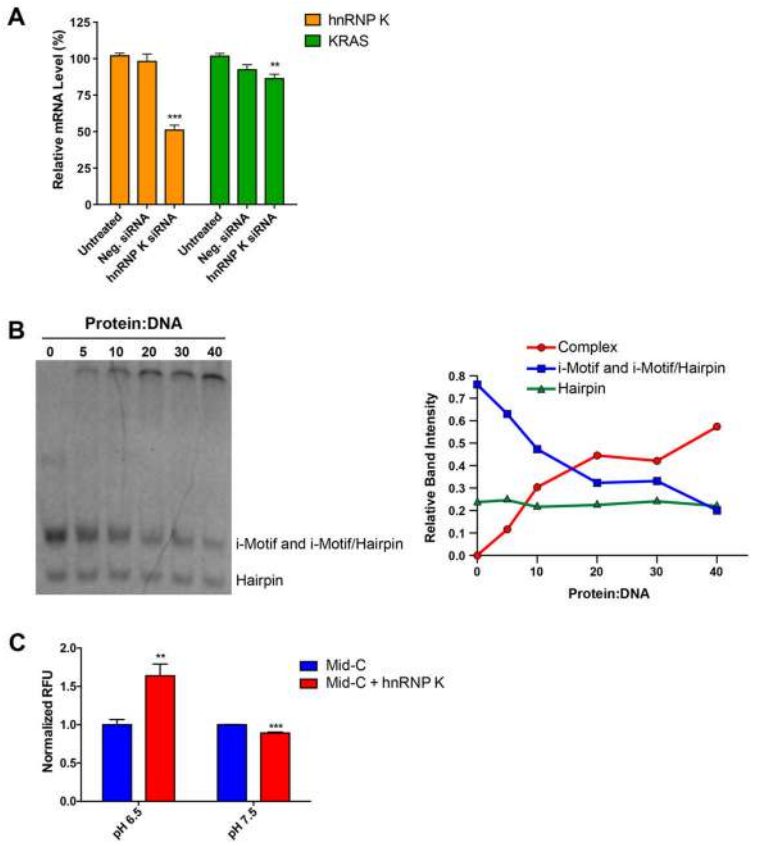


Figure 5.

(A) The effects of siRNA knockdown of hnRNP K on the levels of hnRNP K and KRAS mRNA in MCF-7 cells (**$P < 0.01$, ***$P < 0.001$). (B) EMSA analysis of concentration-dependent binding of hnRNP K to the Mid-C oligomer at pH 6.8 (left) and densitometric analysis of the EMSA gel (right). The i-motif, hybrid i-motif/hairpin, and hairpin species are those shown in Figure 4D. (C) FRET protein binding experiment highlighting the differential unfolding activity of the Mid-region i-motif (pH 6.5) compared to the hairpin (pH 7.5) by hnRNP K (**$P < 0.01$, ***$P < 0.001$).

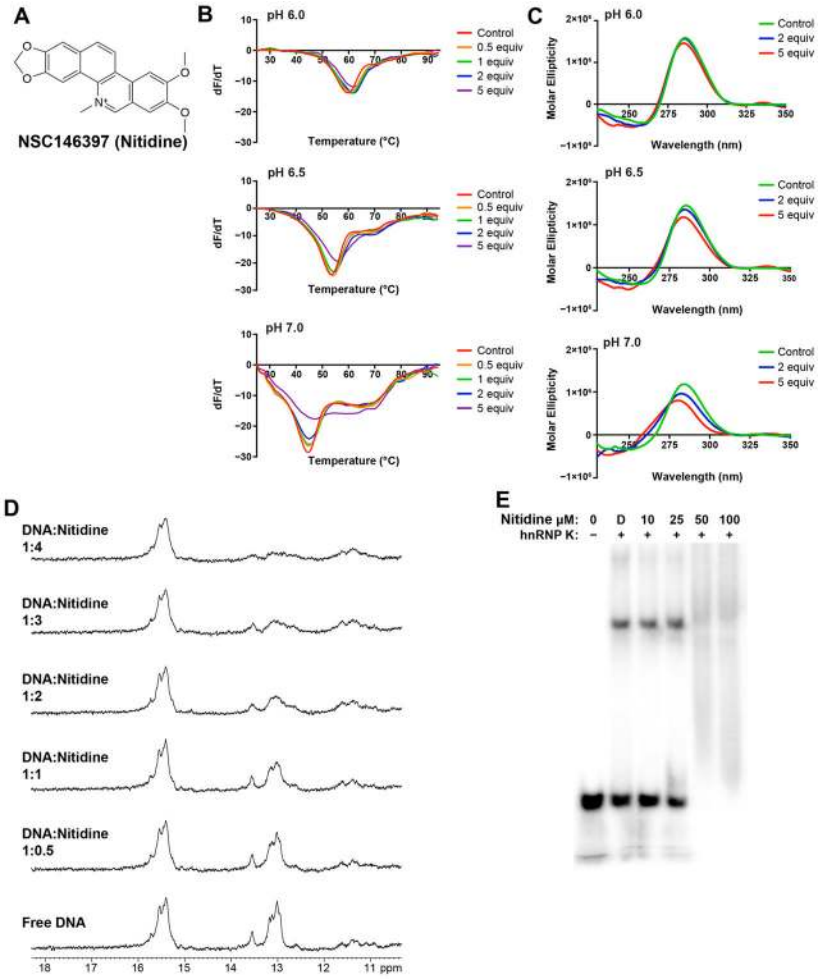


Figure 6.

(A) Chemical structure of Nitidine. (B) Effect of Nitidine on the Mid-C FRET oligomer pH 6.0 (top), pH 6.5 (middle), and pH 7.0 (bottom). Graphs represent the first derivative of the melting curves. (C) CD spectra for Mid-C oligomer with varying concentrations of Nitidine at pH 6.0 (top), pH 6.5 (middle), and pH 7.0 (bottom). (D) The imino proton regions of the 1D ^{1}H NMR spectra at 25 °C and pH 6.5 for the Mid-C oligomer with increasing equivalents of Nitidine. (E) EMSA analysis of the concentration-dependent interaction of Nitidine with the *KRAS* Mid-C oligomer and hnRNP K at pH 6.8 (D = DMSO control).

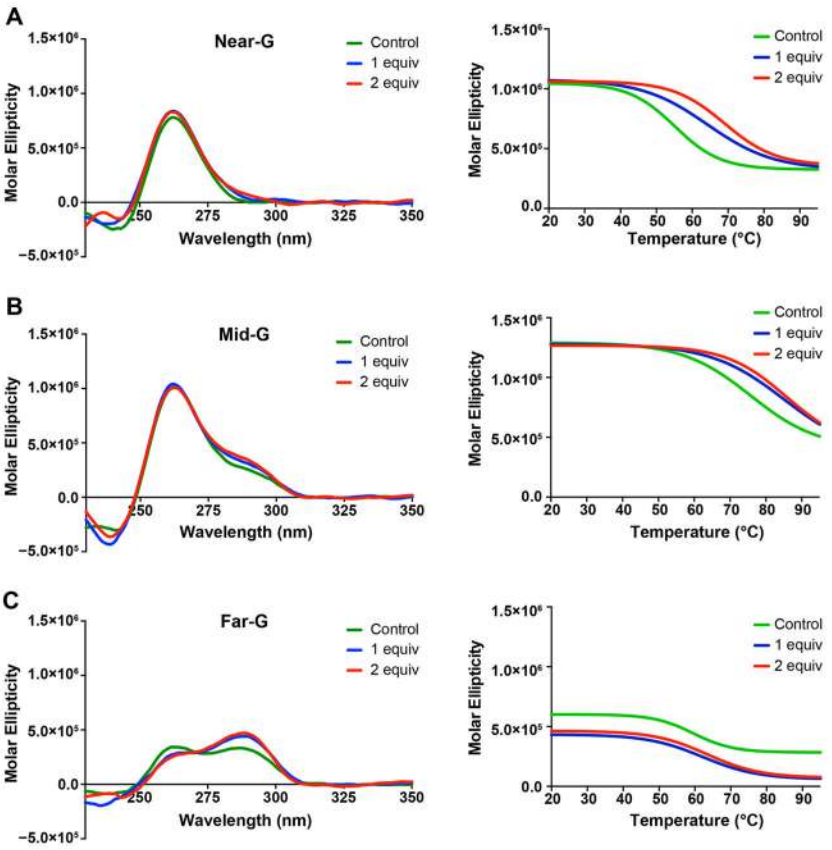


Figure 7.

CD spectra and thermal melting curves for the Near-G (A), Mid-G (B), and Far-G (C) oligomers with varying concentrations of Nitidine in 100 mM KCl.

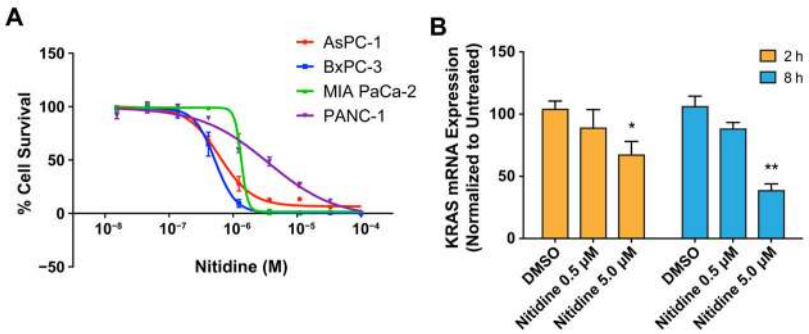


Figure 8.

(A) MTS analysis of Nitidine on four pancreatic cancer cell lines indicating IC_{50} values of 616, 523, 1340, and 3530 nM for AsPC-1, BxPC-3, MIA PaCa-2, and PANC-1 respectively (curves are representative of three individual assays). (B) An analysis of KRAS mRNA expression levels in AsPC-1 cells following treatment with Nitidine at the indicated times and concentrations, showing significant downregulation of KRAS (*$P < 0.05$, **$P < 0.01$).

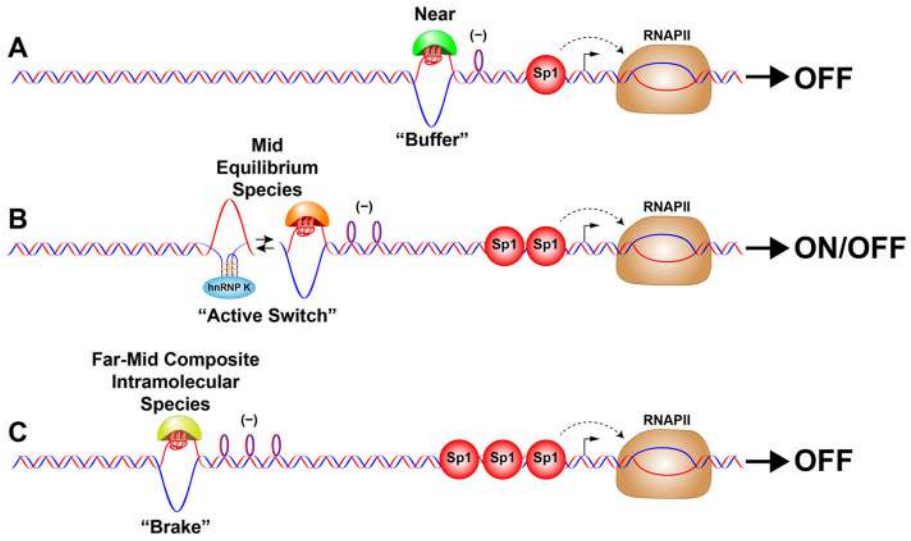


Figure 9.

Proposed model for the functional roles of the Far-, Mid- and Near-region DNA secondary structures in a three-stage Sp1-induced mechanosensor control of *KRAS* transcription. (A) At the lowest level of Sp1-induced negative superhelicity, the Near-region is accessed at basal conditions and acts as a "buffer" that prevents access to the Mid-region until hnRNP A1 dissolves the G-quadruplex structure and converts it back to duplex.[33] (B) At an increased level of Sp1-induced negative superhelicity, the Mid-region element is engaged and acts as an "active switch" in which the G-quadruplex and i-motif are in dynamic equilibrium. The more stable G-quadruplex (OFF) predominates until hnRNP K is present and can actively compete with the protein that stabilizes the G-quadruplex (ON). This model assumes a mutual exclusivity for the G-quadruplex and i-motif. (C) Finally, the Far-Mid composite intramolecular species, formed by combining the Far- and Mid-regions, is accessed at the highest level of Sp1-induced supercoiling and acts as a "brake" to prevent access of hnRNP K to the Mid i-motif.

Table 1

Sequences of Single-Stranded Oligonucleotides Used in This Study

DNA oligomer	sequences (5′–3′)
Near-C	TCCCCCTCTTCCCTCTTCCCACACCGCCCT
Mid-C	GCCCGGCCCCCGCTCCTCCCCCGCCGGCCCGGCCCGGCCCCCTCCTTCTCCCCG
Far-C	GCCCGGCTCGCCACCCTAGACCGCCCCAGCCACCCCT
Near-G	AGGGCGGTGTGGGAAGAGGGAAGAGGGGGA
Mid-G	CGGGGAGAAGGAGGGGGCCGGGCCGGGCCGGCGGGGGAGGAGCGGGGGCCGGGC
Far-G	AGGGGTGGCTGGGGCGGTCTAGGGTGGCGAGCCGGGC
Mid-Near-C	GCCCGGCCCCCGCTCCTCCCCCGCCGGCCCGGCCCGGCCCCCTCCTTCTCCCCGCCGGCGCTCGCTGCCTCCCCCTCTTCCCTCTTCCCACACCGCCCT
Far-Mid-C	GCCCGGCTCGCCACCCTAGACCGCCCCAGCCACCCCTTCCTCCGCCGGCCCGGCCCCCGCTCCTCCCCCGCCGGCCCGGCCCGGCCCCCTCCTTCTCCCCG
Mid-C Lower Stem T	GCCCGGCCCCCGCTCCTCCCCCGTTGGCCCGGCCCGGTCCCCTCCTTCTCCCCG
Mid-C Upper Stem T	GCCCGGCCCCCGCTCCTCCCCCGCTGGCCCGGCTTGGCCCCCTCCTTCTCCCCG
Mid-C Loop T	GCCCGGCCCCCGCTCCTCCCCCGCCGGTTTGGCCCGGCCCCCTCCTTCTCCCCG
Mid-C Full Stem T	GCCCGGCCCCCGCTCCTCCCCCGTTGGCCCGGCTTGGTCCCCTCCTTCTCCCCG
Mid-C Full Stem G	GCCCGGCCCCCGCTCCTCCCCCGGGGGCCCGGCGGGGGCCCCTCCTTCTCCCCG
BCL2 i-Motif	CAGCCCCGCTCCCGCCCCCTTCCTCCCGCGCCCGCCCCT
MYC 4CT	TCCCCACCTTCCCCACCCTCCCCACCCTCCCCA

J Am Chem Soc. Author manuscript; available in PMC 2018 June 28.

Table 2

Peak Molar Ellipticity Values for i-Motif-Forming Sequences (286 nm, pH 4.5)

DNA oligomer	peak molar ellipticity	T_m at pH 6.0 (°C)
Near-C	1.32×10^6	39.8
Mid-C	2.00×10^6	50.3
Far-C	1.01×10^6	33.1
Mid-Near-C	3.95×10^6	40.8
Far-Mid-C	3.88×10^6	43.1