

Expanding the Functions of KHSRP Protein: Insights into DNA G-Quadruplex Binding

Pasquale Russomanno, Pasquale Zizza, Linda Cerofolini, Federica D'Aria, Sara Iachettini, Serena Di Vito, Annamaria Biroccio, Jussara Amato, Marco Fragai, and Bruno Pagano*

KHSRP (KH-type splicing regulatory protein) is a multifunctional nucleic acid-binding protein that regulates various cellular processes, with critical roles in controlling gene expression. G-quadruplexes (G4s) are noncanonical nucleic acid structures involved in essential cellular activities, including gene expression, and are recognized as potential therapeutic targets in cancer. The biological functions of G4s are mediated by proteins making their formation highly dynamic within cells. Therefore, the recognition of G4s by specific proteins is crucial for modulating physiological and pathological pathways. Given the growing interest in DNA G4s, a deeper understanding of the proteins that interact with them and their molecular recognition is imperative. This study demonstrates that KHSRP binds to these DNA structures. Biophysical analyses provide insights into the thermodynamics, kinetics, and structural aspects of these interactions, showing that G4 structural variability significantly influences KHSRP binding, in which the KH3 protein domain plays a key role. Validation of these interactions in cancer cells further highlights their biological relevance. Notably, the G4 ligand pyridostatin affects KHSRP/G4 interactions both in vitro and in cells, suggesting that small molecules can modulate this molecular recognition. These findings underscore KHSRP's potential role in regulating cellular mechanisms through binding to G4-forming DNA, positioning it as a possible therapeutic target in cancer.

1. Introduction

KHSRP (KH-type splicing regulatory protein), also known as KSRP or FUBP2 (far upstream element-binding protein 2), is a multifunctional nucleic acid-binding protein present in both the nucleus and cytoplasm, that has emerged as a multifaceted regulator of diverse cellular processes with critical roles in fine-tuning gene expression.^[1,2]

KHSRP protein structure is organized in three distinct regions: two N- and C-terminal low sequence complexity regions that contain several sites for post-translational modifications as well as protein-interaction motifs, and a mainly structured central region comprising four hnRNPK homology (KH) domains responsible for nucleic acid binding (Figure 1).^[3] The modular structure of KHSRP and its consequent binding flexibility enables the protein to interact with multiple targets.^[1]

Originally identified as a factor interacting with an enhancer element upstream of the *c-MYC* oncogene promoter thereby stimulating its expression,^[4] KHSRP

has since been implicated in a plethora of post-transcriptional processes, highlighting its functional diversity and versatility.^[5,6]

The RNA-binding capability of KHSRP allows it to interact with a wide range of RNA species, including messenger RNAs (mRNAs), non-coding RNAs, and viral RNAs.^[1,5-8] Moreover, KHSRP has been shown to participate in alternative splicing regulation, contributing to the generation of transcriptomic diversity and the production of functionally distinct protein isoforms.^[5,6,9,10] The multifaceted functions of KHSRP are not limited to RNA metabolism but also extend to broader cellular processes. Indeed, KHSRP has been implicated in cell proliferation, apoptosis, development, and neuronal processes.^[5,10-13] For example, it was recently shown that KHSRP controls monocytic differentiation by interacting with specific genomic sites in promoter and enhancer regions to regulate the expression of several genes through transcriptional activation and bind to pre-mRNA intronic regions to modulate alternative splicing, thus showing the dual DNA- and RNA-binding activity of the protein.^[14]

Given its diverse functions, altered KHSRP expression has been associated with various human diseases, including cancer,

P. Russomanno, F. D'Aria, J. Amato, B. Pagano
Department of Pharmacy
University of Naples Federico II
Naples 80131, Italy
E-mail: bruno.pagano@unina.it

P. Russomanno, L. Cerofolini, M. Fragai
CERM-CIRMMP and Department of Chemistry "Ugo Schiff"
University of Florence
Sesto Fiorentino (FI) 50019, Italy

P. Zizza, S. Iachettini, S. Di Vito, A. Biroccio
Translational Oncology Research Unit
IRCCS-Regina Elena National Cancer Institute
Rome 00144, Italy

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/advs.202410086>

© 2024 The Author(s). Advanced Science published by Wiley-VCH GmbH. This is an open access article under the terms of the [Creative Commons Attribution License](#), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/advs.202410086

cardiovascular diseases, and neurological disorders, underscoring its clinical relevance.^[5,6,15–17]

G-quadruplexes (G4s) are nucleic acid secondary structures that can arise under physiological conditions from guanine(G)-rich sequences usually containing four or more consecutive G-runs of at least two guanines.^[18] These structures consist of the stacking of at least two G-tetrads, the planar association of four guanines held together by Hoogsteen hydrogen bonds. Therefore, they are markedly different from canonical double-helical DNA.^[19] Moreover, G4s can exhibit a high degree of structural polymorphism: depending on the sequence, loop length, and base composition, the relative orientation of G-runs can be parallel, antiparallel, or mixed (referred to as hybrid), giving rise to different conformational topologies. G4 folding and stability are strongly influenced by the presence of cations.^[20] Cation coordination is indeed required to form G4s, to stabilize the G-tetrad stacks. Cation type and concentration can also significantly impact G4 conformations, an especially relevant factor given the polymorphic nature of these structures, potentially affecting their functional roles in biological systems.

It is well established that G4-forming sequences are present in the human telomeric DNA and in the promoters of several oncogenes, such as *c-MYC*, *c-KIT*, and *BCL-2*,^[21] playing important roles in various biological functions.^[22] In the cellular environment, the formation of G4 structures is heavily influenced by interactions with a multitude of protein factors.^[23–26] These protein interactors can alter the equilibrium between unfolded G-rich sequences and folded G4 structures, either stabilizing or destabilizing them, which can impact gene expression. As a result, the formation of G4 structures in living cells exhibits remarkable dynamism. Interestingly, numerous single-stranded RNA-binding proteins are known to directly interact with G4s.^[27–31]

Biochemical studies conducted more than 20 years ago revealed that KHSRP is a component of a multiprotein complex that includes heterogeneous nuclear ribonucleoprotein F/H1 (hnRNPF/hnRNPH1) and polypyrimidine tract binding protein (PTBP1),^[10,32] with the first two more recently shown to interact also with G4 structures formed by RNA sequences.^[27,33,34] It has also been observed that KHSRP can associate with hnRNPA1 in different cell types and that the two proteins show similar binding preferences to RNA sequences,^[35,36] with a preferred binding target for both proteins including G-rich sequences.^[11,36,37] KHSRP and hnRNPA1 have also been proposed to have antagonistic roles in the post-transcriptional regulation of miRNA expression.^[36] According to the literature reports, hnRNPA1 also participates in the resolution of G4 DNA structures at the end of telomeres to stimulate telomere elongation.^[38,39] Furthermore, it is able to recognize and unfold G4s in some gene promoter regions, including those found in *KRAS* and *TRA2B* gene promoters, thereby modulating their transcription in cancer cells.^[40–42] Indeed, accumulating evidence indicates that G4 DNA formed by promoter sequences is involved in the regulation of gene expression.^[23,43–45]

In a previous chemoproteomic-driven study, some of us identified KHSRP as a potential interacting partner of human telomeric G4 DNA.^[46] Understanding the intricate functions and regulatory mechanisms of KHSRP is essential for unraveling the complexities of gene regulation. In this study, we demonstrated KHSRP's ability to interact with the G4-forming sequences located in oncogene promoter regions. By integrating biophysical

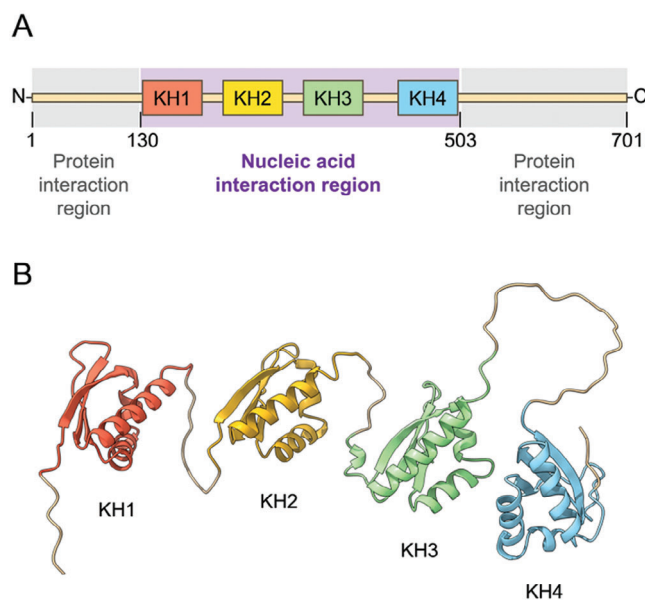


Figure 1. A) Domain organization of full-length human KHSRP (KHSRP_{fl}), and B) structural model of the central nucleic acid binding region consisting of the four KH domains (KHSRP₁₃₀₋₅₀₃) obtained with AlphaFold DB. The Protein Data Bank codes used for the four KH domains are 2opu (KH1), 2opv (KH2), 2hh3 (KH3), and 2hh2 (KH4).

and biological analyses, we gained valuable insights into the energetic aspects and molecular characteristics of KHSRP binding to G4s from the *c-MYC* and *c-KIT* gene promoters, to which the protein showed a preference. Noteworthy, the validation of these interactions in the cellular context highlighted their biological relevance. Our findings shed light on the involvement of protein domains in recognizing these DNA structures, as well as on the G4 structural determinants that contribute to this interaction.

2. Results

2.1. Protein Expression, Purification, and Assignment of Protein Residues

The full-length KHSRP protein (KHSRP_{fl}, residues 1–711, Figure S1, Supporting Information) contains a folded central region that includes the four KH domains (residues 130–503) responsible for nucleic acid binding,^[37,47,48] and two unfolded regions located at the N- and C-terminus (1–129 and 504–711, respectively) (Figure 1). So far, eukaryotic systems have been utilized to express KHSRP_{fl}, followed by protein purification using the ammonium sulfate precipitation method and subsequent refolding.^[10,49,50] Here, we have devised an alternative approach to obtain the protein in its native folding, consisting of using *Escherichia coli* as the expression system, and a purification method based on a tandem combination of GST affinity and size-exclusion chromatography (see Figure S2 for protein construct, Supporting Information). The identity and purity of KHSRP_{fl} were ascertained by SDS-PAGE analysis and 1D ¹H-NMR spectroscopy (Figure S3, Supporting Information). The NMR spectrum of KHSRP_{fl} obtained through the tandem purification method was compared with that of the protein purified

exclusively by GST affinity chromatography. Both purification methods successfully yielded properly folded protein. However, only the tandem method allowed for obtaining the protein completely free of bacterial DNA, as evidenced by the absence of signals in the 5–6 ppm region of the spectrum, which corresponds to the resonances of DNA sugar protons (Figure S3, Supporting Information). However, due to the full-length protein's low stability and high spectral complexity, its utility in the study of the protein-G4 interaction was limited. Consequently, we opted to proceed with the investigation using the truncated form of the protein (KHSRP₁₃₀₋₅₀₃) comprising solely the four KH domains (see Figure S2 for protein construct, Supporting Information). The expression and purification of the KHSRP₁₃₀₋₅₀₃ protein were conducted using the procedures described by Gherzi et al.^[49] and Kung et al.^[50] with slight modifications. These modifications included the implementation of protein overexpression in various *Escherichia coli* strains and the optimization of the protein solubility by using a low growth temperature (see the Experimental Section for details, Supporting Information). SDS-PAGE and 1D ¹H-NMR analysis confirmed the identity and purity of KHSRP₁₃₀₋₅₀₃ (Figure S4, Supporting Information).

The spreading of the signals in the 2D ¹H-¹⁵N HSQC spectrum acquired on [U-¹⁵N] KHSRP₁₃₀₋₅₀₃ is consistent with the presence of well-structured protein domains (Figure S5, Supporting Information). Interestingly, the 2D ¹H-¹⁵N HSQC spectrum showed sharp and well-resolved signals for a protein of 40 kDa, corroborating the presence of interdomain flexibility. This finding is also supported by the signals' overlap observed in the central region of the spectrum since most of the residues of the linkers are expected to resonate at these frequencies. The backbone assignment of the protein has been obtained from the analysis of triple resonance spectra recorded on samples of [U-¹³C, ¹⁵N] KHSRP₁₃₀₋₅₀₃ (see the Experimental Section and Figure S6, Supporting Information). ≈83% of the total expected non-proline spin systems have been assigned (Figure S5, Supporting Information); most of the unassigned residues are in the loop regions, mainly in the long linker connecting KH3 and KH4 domains (Gln275-Lys281, Arg320-Gly324, Arg340-Glu343, Ile386-Leu390, Pro401-Trp420). Twenty-two proline residues, out of a total of 34, have been assigned by integrating the data obtained from the ¹H- and ¹³C-detected NMR spectra. The protein resonance assignment has now been reported in the BMRB database under the accession code: 52573. The prediction of the secondary structure elements, obtained by TALOS+ analysis using the resonances of KHSRP₁₃₀₋₅₀₃ as input (Figure S7, Supporting Information), is in agreement with the folding of protein reported in the AlphaFold structural model.^[51]

2.2. NMR Investigation of KHSRP₁₃₀₋₅₀₃ Interaction with DNA G4s

1D ¹H macromolecule-based NMR experiments were performed as a first biophysical assay to detect the interaction between KHSRP₁₃₀₋₅₀₃ and several DNA G4s which differ in topology and/or loop length and composition and/or presence of flanking residues. In particular, three parallel-stranded G4s, formed by the G-rich sequences of *c-KIT* (*c-Kit1* and *c-Kit2*) and *c-MYC* (*c-Myc*) oncogene promoters,^[52-54] and two hybrid topology G4s

from the promoter region of the *BCL-2* (*Bcl-2*) oncogene and human telomeric DNA (*mTel₂₄*),^[55,56] were chosen for the investigation. Since we were dealing with two macromolecules, both protein-based and DNA-based assays have been performed, where changes in the line width of the ¹H nuclei belonging to the protein or DNA have been analyzed, respectively.

Simple 1D ¹H NMR experiments can be used to demonstrate the protein-ligand interaction by observing signals in the aliphatic part of the protein's spectrum (typically below 0.7 ppm) both in the absence and presence of the potential binder.^[57] Indeed, changes in intensity and/or chemical shift of these signals provide immediate evidence of ligand binding, and qualitative information on its affinity by analyzing the evolution of the chemical shift of the signals upon the addition of increasing amounts of the investigated compound. Instead, for the experiments based on the observation of the signals of DNA, information on ligand binding can be obtained by monitoring protons resonating in specific regions, according to the specific DNA secondary structure. Concerning a G4 structure, the peaks of the imino protons assigned to the guanines involved in G-tetrad formation, exhibit distinct chemical shifts typically ranging from 10.5 to 12 ppm.^[58] Analysis of the changes affecting these signals can be used to monitor structural alterations of G4 structures, as well as reveal interaction with ligands and proteins.

Therefore, 1D ¹H NMR spectra were recorded on KHSRP₁₃₀₋₅₀₃ and on the different DNA G4s in the absence and presence of increasing amounts of the investigated G4s or the protein, respectively. As for the former, a significant decrease in the intensity of the proton signals of the protein located in the aliphatic region of the spectrum was observed upon the addition of G4s, clearly indicating the formation of the corresponding protein-DNA complexes (Figure S8, Supporting Information). Interestingly, these variations were more significant for *c-Kit1*, *c-Kit2*, and *c-Myc* (all G4s with parallel topology) than for *Bcl-2* and *mTel₂₄* (which adopt hybrid G4 conformations). Additional information on ligand binding can be obtained from the 1D ¹H spectra by observing the region ≈10 ppm where NH proton signal of the Trp indole rings resonate and no signals of DNA are present. Indeed, if a Trp residue is close to the binding site, this region can be used to monitor ligand binding.^[57] Interestingly, we observed clear changes in chemical shift and/or intensity of such a signal when each of the investigated G4s was added. These changes varied thus suggesting different affinities of the DNA molecules for the protein or different binding modes.

Comparable results were obtained in DNA-based experiments. A general decrease in the intensity of imino proton signals was observed for all G4s (Figure 2), thus confirming the ability of KHSRP₁₃₀₋₅₀₃ to interact with all of them, albeit with a different affinity. Indeed, greater variations were observed for *c-Kit1*, *c-Kit2*, and *c-Myc* G4s than for the others, again suggesting stronger interactions with parallel-stranded than hybrid G4s. Furthermore, the DNA-based experiments also provided insight into the G4 residues mostly affected by the interaction with the protein. For *c-Kit1*, *c-Kit2*, and *c-Myc* G4s, the guanine peak intensities of the inner G-tetrads resulted less altered than those of the outer G-tetrads, which is consistent with external protein binding without structural changes to the G4s.

Moreover, in the case of *c-Myc*, significant shifts were observed for the aromatic protons of adenine base (A12)

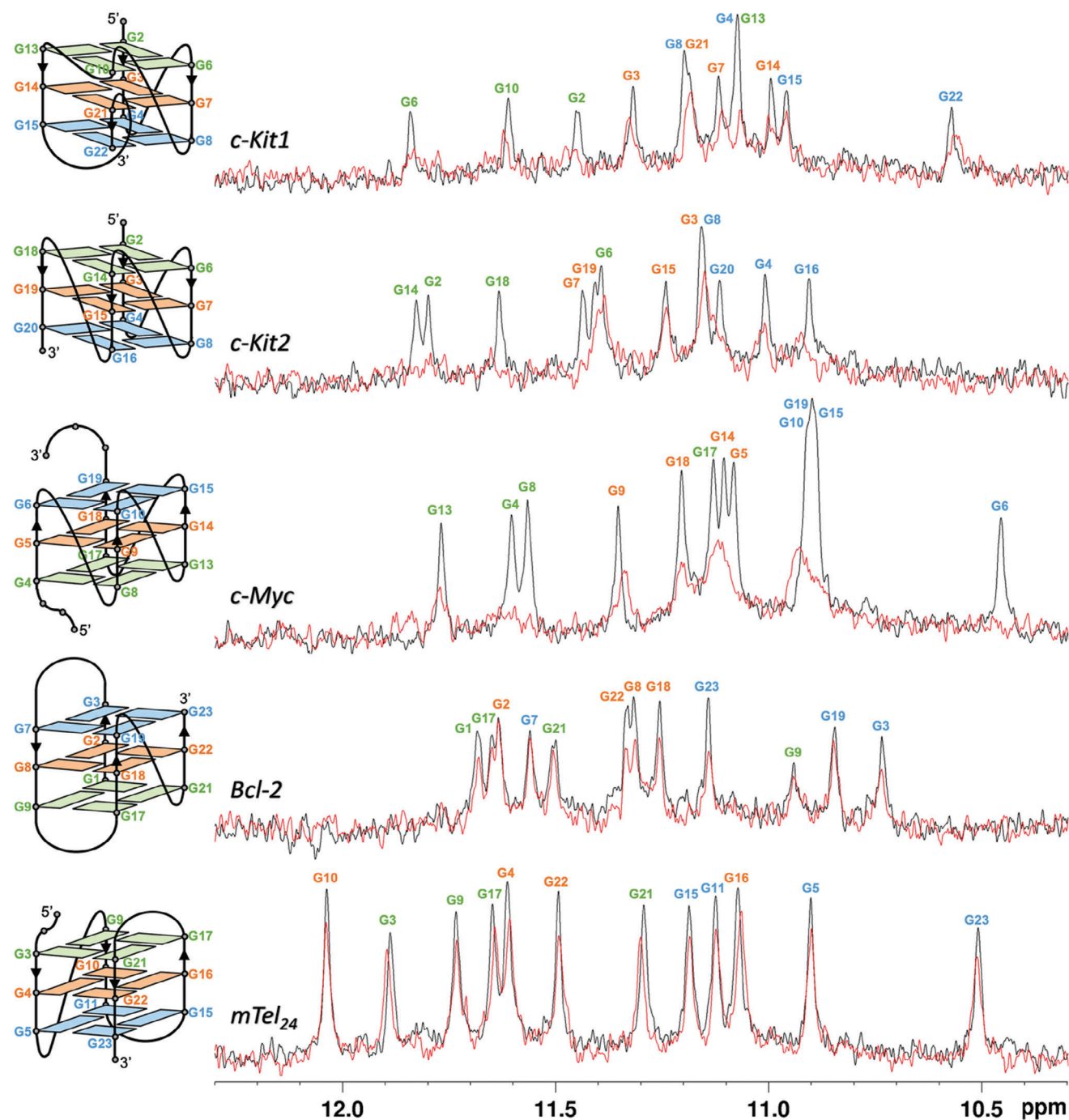


Figure 2. (Left) Schematic representation of the folding topology of the investigated G4s. The guanines forming the G-tetrads are depicted as colored boxes. Strand directionalities are indicated by arrows. (Right) Imino proton spectra of G4s (20 μ M) at 25 $^{\circ}$ C in the absence (black lines) and presence (red lines) of the KHSRP₁₃₀₋₅₀₃ protein (1:1 ratio).

belonging to the double nucleotide loop (Figure S9, Supporting Information), which is located toward the G-tetrad groove pointing to the top of the G4 structure.^[54] This observation suggests that the interaction with the protein leads to the displacement of this base. These findings may also explain the minor effects observed in the interaction between the protein and hybrid G4s, where the external G-tetrads are rather

occupied by the loops, making them less prone to interact.

To better assess differences in protein binding to the external G-tetrads, the data from the 1D NMR titrations of G4s with KHSRP₁₃₀₋₅₀₃ were further analyzed. Specifically, in each spectrum, the intensities of the imino proton signals were normalized relative to an isolated signal unaffected by the addition of

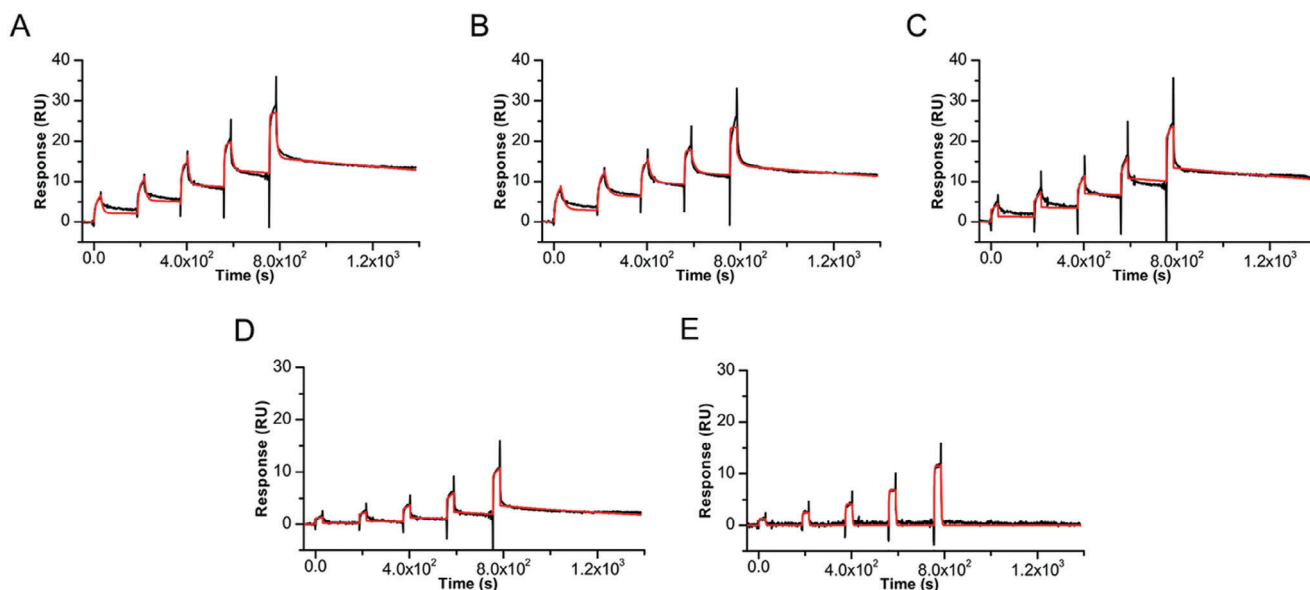


Figure 3. Time evolution SPR sensorgrams obtained at 25 °C by injections of various concentrations (from 1 to 20 μM) of A) *c-Kit1*, B) *c-Kit2*, C) *c-Myc*, D) *Bcl-2*, and E) *mTel*₂₄ G4s on the chip-immobilized KHSRP₁₃₀₋₅₀₃, with a contact time of 30 s and a flow rate of 30 $\mu\text{L min}^{-1}$. The sensorgrams are shown as black lines and their respective fits as red lines.

KHSRP₁₃₀₋₅₀₃. Normalization was performed on the signal of G3 for *c-Kit1*, G15 for *c-Kit2*, G9 for *c-Myc*, G7 for *Bcl-2*, and G5 for *mTel*₂₄, respectively. The normalized intensities of the signals in the presence of KHSRP₁₃₀₋₅₀₃ were compared to the normalized intensities of the same signals in the absence of the protein. The residues most affected by the interaction with KHSRP₁₃₀₋₅₀₃ (G2, G4-G13, G6, and G10 for *c-Kit1*; G2, G7, G14, and G18 for *c-Kit2*; G4, G5, G6, and G8, for *c-Myc*; G1 and G3 for *Bcl-2*; G15 and G21 for *mTel*₂₄) suggest different binding modes of KHSRP₁₃₀₋₅₀₃ for the G4s (Figure S10, Supporting Information). In particular, the G-tetrad at the 5'-end is primarily affected in *c-Kit1* and *c-Kit2*, whereas for *c-Myc* the 5'-end G-tetrad together with a side of the G4 are mostly affected by protein binding. On the other hand, the interaction of KHSRP₁₃₀₋₅₀₃ with *Bcl-2* and *mTel*₂₄ appears weaker and less specific.

2.3. Assessment of Binding Kinetics and Affinity of KHSRP₁₃₀₋₅₀₃ Toward G4s

Surface plasmon resonance (SPR) experiments were conducted to quantitatively assess the binding affinity of KHSRP₁₃₀₋₅₀₃ to the studied G4s and obtain insights into the kinetic parameters of the interaction. SPR is a powerful technique for real-time investigation of DNA-protein interactions as it allows measurement of the association (k_{on}) and dissociation (k_{off}) rate constants, enabling the determination of the equilibrium dissociation constant (K_{d}) values.^[59–61]

Figure 3 shows representative SPR sensorgrams obtained for the interaction of immobilized KHSRP₁₃₀₋₅₀₃ with increasing concentrations of the investigated G4s using the single-cycle kinetics (SCK) method. Representative replicates for each DNA, along with their corresponding fits, are shown in Figure S11 (Supporting Information). SCK was chosen over multi-cycle kinetics due

Table 1. Kinetic parameters and dissociation constants for the interaction of the investigated G4s with KHSRP₁₃₀₋₅₀₃, determined by SPR.

G4	$k_{\text{on}} [\times 10^3 \text{ M}^{-1} \text{ s}^{-1}]$	$k_{\text{off}} [\times 10^{-4} \text{ s}^{-1}]$	$K_{\text{d}} [\text{nM}]$
<i>c-Kit1</i>	3.1 (± 0.6)	5.4 (± 1.7)	190 (± 90)
<i>c-Kit2</i>	5.0 (± 2.0)	2.8 (± 1.5)	56 (± 8)
<i>c-Myc</i>	2.2 (± 0.8)	2.8 (± 1.5)	130 (± 25)
<i>Bcl-2</i>	0.8 (± 0.5)	12.9 (± 2.0)	2000 (± 980)
<i>mTel</i> ₂₄	4.1 (± 2.7)	50.9 (± 12.8)	1470 (± 480)

to its advantage of using a single regeneration step at the end of the complete binding cycle. This single regeneration step is less detrimental, resulting in a longer lifespan for the chip surface and immobilized protein.

Almost all sensorgrams showed a response proportional to analyte concentration in solution, thus indicating the specificity of DNA/protein interactions. SPR data were then analyzed in the framework of the 1:1 binding model to derive the kinetic rate constants, from which the equilibrium dissociation constants were determined. The K_{d} values presented in Table 1 show a stronger interaction between the protein and the *c-Kit1*, *c-Kit2*, and *c-Myc* G4s, as compared to the *Bcl-2* and *mTel*₂₄ G4s. These findings are consistent with the NMR data.

Noteworthy, protein binding exhibits a difference in affinity of approximately ten-fold or more between parallel (*c-Kit1*, *c-Kit2*, and *c-Myc*) and hybrid (*Bcl-2* and *mTel*₂₄) G4s. The comparison of the kinetic constants indicates that the disparity is primarily attributed to a variation in the dissociation rate rather than the association rate. In fact, the protein demonstrates a relatively similar association with the different G4s, while the complex dissociates more rapidly in the case of hybrid G4s compared to parallel ones.

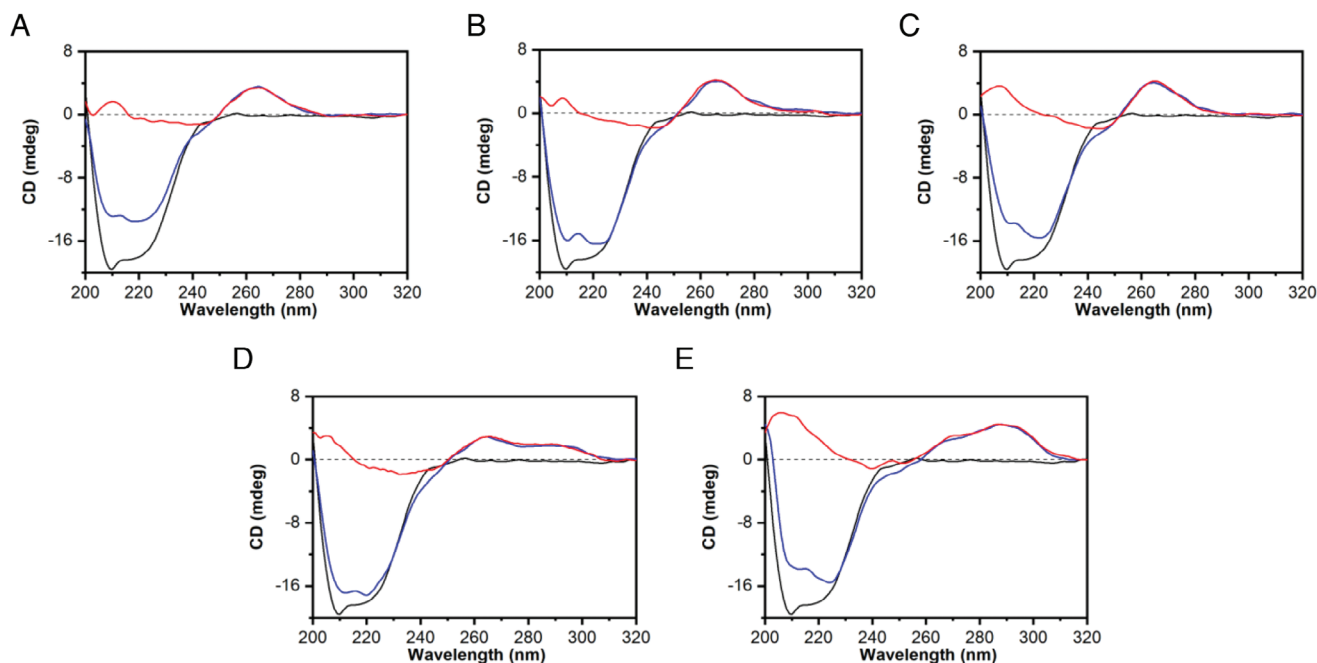


Figure 4. CD spectra of A) *c-Kit1*, B) *c-Kit2*, C) *c-Myc*, D) *Bcl-2*, and E) *mTel*₂₄ G4s in the absence (red lines) and presence (blue lines) of KHSRP₁₃₀₋₅₀₃ (1 molar equiv.) at 25 °C. The CD spectrum of the protein alone (black lines) is also shown in each panel for comparison.

Among the parallel G4s, KHSRP₁₃₀₋₅₀₃ exhibited the highest affinity for *c-Kit2*, followed by *c-Myc*. In this case, the analysis of the kinetic constants indicates that the difference lies in the rate of association.

2.4. Circular Dichroism (CD) Spectroscopy Analysis of DNA G4s in the Presence of KHSRP₁₃₀₋₅₀₃

To verify that the decrease in the intensity of the G4 imino protons observed in the DNA-based NMR experiments was not caused by DNA structure unfolding or secondary structure changes, CD experiments were conducted. CD spectroscopy is extremely sensitive to conformational changes of nucleic acids and provides characteristic spectra depending on the structure adopted.^[62] Consistent with the presence of a parallel G4 topology, the CD spectra of *c-Kit1*, *c-Kit2*, and *c-Myc* exhibited the typical signature of this structure, characterized by a positive band at 263 nm and a negative band \approx 240 nm (Figure 4).^[63] On the other hand, *Bcl-2* and *mTel*₂₄ displayed CD spectra with two positive bands at \approx 268 and 290 nm, along with a negative band \approx 240 nm, indicative of hybrid-type G4 topologies.^[63]

The CD spectra of KHSRP₁₃₀₋₅₀₃ alone, as well as those of protein/G4 (1:1 molar equiv.) mixtures, were also recorded and are shown in Figure 4. Notably, no unfolding or structural alterations of DNA G4s were observed upon protein interaction. Indeed, the presence of KHSRP₁₃₀₋₅₀₃ does not significantly affect, both in terms of signal intensity and shape, the CD spectrum of any of the G4s at wavelengths above 245 nm, which corresponds to the spectral region where only the G4 chromophores absorb. However, it should be noted that for interactions with higher K_d values (i.e., *Bcl-2* and *mTel*₂₄), the complex is only par-

tially formed under the assay conditions. These experiments were also conducted with protein concentrations two- and four-fold higher than that of the G4s (Figure S12, Supporting Information). No significant perturbations in the structures of the G4s were observed, even in the presence of an excess of protein. This confirms that KHSRP₁₃₀₋₅₀₃ interacts with the G4s without affecting their structure, at least under these experimental conditions.

Furthermore, CD melting experiments were performed to assess any potential stabilizing or destabilizing effect of the protein on the studied G4 structures. Interestingly, distinct effects on the thermal stability of G4 structures were observed upon the addition of KHSRP₁₃₀₋₅₀₃ (Figure S13, Supporting Information). No changes were observed in the thermal stability of *c-Kit1* and *Bcl-2* G4s. On the other hand, the CD melting curves showed a slight decrease in the stability of *mTel*₂₄ ($\Delta T_m = -2.2 \pm 0.5$ °C) and a significant increase in the stability of *c-Kit2* and *c-Myc* G4s (with ΔT_m values $\approx 5.3 \pm 1.0$ °C) in the presence of the protein.

2.5. Structural Characterization of KHSRP₁₃₀₋₅₀₃ Interaction with *c-Myc* and *c-Kit2* DNA G4s

Mapping of the surface of KHSRP₁₃₀₋₅₀₃ protein involved in the interaction with different G4 molecules has been obtained by monitoring changes in cross-peak intensity (I/I_0) and/or in chemical shift ($\Delta\delta$), occurring in the 2D ¹H-¹⁵N HSQC spectrum of the uniformly ¹⁵N-isotopically enriched protein upon the addition of increasing amounts of the G4-forming DNAs, versus the spectrum of the free protein in solution. In particular, the effect of the two parallel G4s showing the highest affinity for KHSRP₁₃₀₋₅₀₃, namely *c-Myc* and *c-Kit2*, was investigated

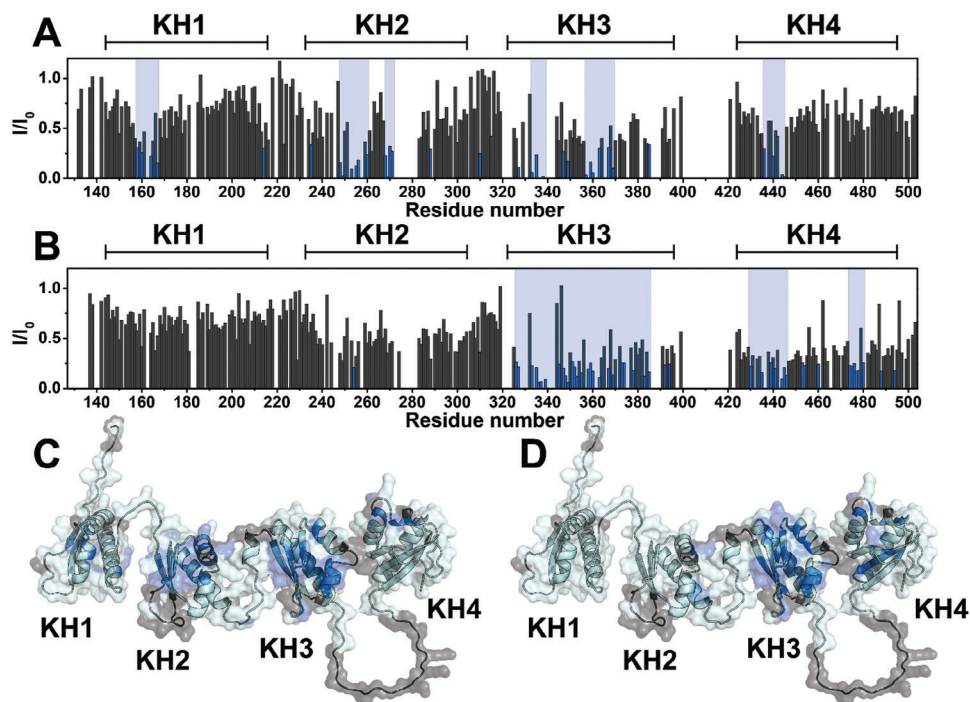


Figure 5. Graphical representation of the per-residue intensity changes of KHSRP₁₃₀₋₅₀₃ (50 μ M) in the presence of A,C) 25 μ M *c-Myc* and B,D) 6.25 μ M *c-Kit2*. The residues exhibiting the highest decreases in signal intensities (*c-Myc*: Leu158, Ile160, Gly164, Ile167, Arg214, Val235, Val248, Ile249, Gly253, Glu254, Thr255, Ile256, Lys257, Gln260, Ile268, Leu269, Ile270, Gln271, Asp288, Gly310, Val327, Ser333, Val334, Gly335, Val336, Val337, Ile338, Gly339, Lys347, Gln349, Gln357, Phe358, Lys359, Gln360, Gly363, Glu367, Ile369, His371, Ile385, Cys436, Ile440, Gly444, Glu445; *c-Kit2*: Leu247, Glu254, Asp326, Val327, Val329, Ser333, Val334, Gly335, Val336, Val337, Gly339, Ile345, Lys347, Ile348, Gln349, Asp351, Ala352, Gly353, Arg355, Gln357, Phe358, Lys359, Gln360, Gly363, Glu367, Ile369, Met373, Gly374, Arg378, Ala383, Ile385, Ser392, Arg394, Ser430, Lys435, Val439, Gly441, Gly444, Glu445, Asn446, Gln453, Thr454, Ile460, Leu474, Phe475, Ile476, Ile477, Arg478, Ser480, Lys488, Lys494) are colored in blue in the plots A,B) and on the AlphaFold model of the protein C,D). Residues that could not be assigned in the NMR spectra are colored in gray on the protein model.

and characterized. In the presence of both *c-Myc* and *c-Kit2*, we observed changes in signal intensity (Figure S14, Supporting Information). In the presence of low concentrations of G4s, largest effects are observed for *c-Kit2*, suggesting a higher affinity of the protein for this G4. Figure 5 shows the plot of the per-residue changes in signal intensity of the protein in the presence of each DNA molecule. Interestingly, the regions of KHSRP₁₃₀₋₅₀₃ affected by DNA binding are slightly different for *c-Myc* and *c-Kit2*. In particular, the signals of the residues located on the KH3 domain are severely decreased in intensity in the presence of both G4-forming DNAs. However, only in the presence of *c-Myc*, several signals corresponding to residues of the KH2 domain experience a large decrease in intensity, while they remain almost unchanged upon the addition of *c-Kit2*. Conversely, residues located on the KH4 domain experience larger effects in the presence of *c-Kit2*. Some more modest effects on the KH1 and KH4 domains due to *c-Myc* DNA were also visible. Moreover, in the presence of both DNA strands, some residues, mostly from the KH3 and KH4 domains, undergo a slight chemical shift perturbation (Figure S15, Supporting Information). The change in chemical shift observed for the residues of the linker between the KH3 and KH4 domains is also intriguing. This could be explained by a conformational change that this loop undergoes upon DNA binding (Figure S16, Supporting Information).

2.6. Thermal Properties of KHSRP₁₃₀₋₅₀₃ in the Presence of *c-Kit2* G4

Differential scanning calorimetry (DSC) is a valuable tool also for investigating the interaction between a biological macromolecule and a ligand.^[64,65] It is particularly useful for assessing how ligand and binding affects the stability of the macromolecule and allows correlation of the thermodynamics that drive the binding with any conformational changes occurring in the macromolecule due to the interaction. Le Chatelier's principle indicates that if a ligand binds preferentially to the folded form of a biological macromolecule, this will stabilize the folded state, and the unfolding of the macromolecule will become less favorable in the presence of the ligand. In the case of two biological macromolecules, their association to form a complex can result in the stabilization of both components. Examples in the literature of DSC heat capacity profiles show that on heating, protein/DNA complexes can unfold at a somewhat higher temperature than the free proteins and free DNA molecules.^[65]

We have shown above, through CD melting experiments, that the strongest G4-KHSRP₁₃₀₋₅₀₃ interactions, specifically with *c-Kit2* and *c-Myc* G4s, result in a significant increase in the stability of these DNA structures in the presence of the protein. Since it was not possible to use CD spectroscopy to investigate the thermal properties of the protein in the presence of G4s, due to the

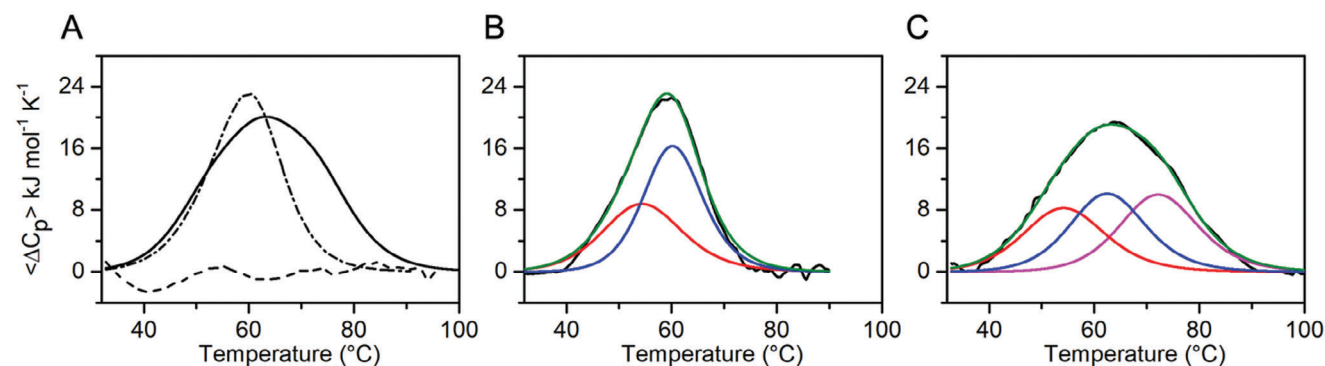


Figure 6. A) Comparison of DSC profiles for KHSRP₁₃₀₋₅₀₃ (dot-dashed line), KHSRP₁₃₀₋₅₀₃/*c-Kit2* mixture (solid line), and *c-Kit2* (dashed line). B,C) DSC thermograms for KHSRP₁₃₀₋₅₀₃ and KHSRP₁₃₀₋₅₀₃ in complex with *c-Kit2* G4 (black lines) with corresponding deconvolutions of the excess heat capacity functions. B) Two transitions are observed for free KHSRP₁₃₀₋₅₀₃ (red and blue lines), and their sum is shown as a green line. T_m values for the first (red line) and second (blue line) transitions are 54 and 60 °C, respectively. C) Three transitions are observed for complexed KHSRP₁₃₀₋₅₀₃ (red, blue, and magenta lines) and their sum is shown as a green line. T_m values for the first (red line), second (blue line), third (magenta line) transitions are \approx 54, 62, and 72 °C, respectively. Experiments were performed in 5 mM potassium phosphate buffer (pH 7.0), 20 mM KCl, using a heating rate of 1 °C min⁻¹.

simultaneous absorption of DNA at the wavelengths useful for this purpose, we employed DSC.

The DSC thermograms obtained for the denaturation of KHSRP₁₃₀₋₅₀₃ in the absence and presence of *c-Kit2* G4 are shown in **Figure 6A**. The DSC curves show that protein unfolding proceeds with significant heat absorption, i.e., enthalpy change. As for the free KHSRP₁₃₀₋₅₀₃, the apparent melting temperature, T_m , corresponding to the maximum of DSC peak (59.0 ± 0.5 °C), coincides with the T_m obtained by following the temperature dependence of the molar ellipticity at 222 nm (58.0 ± 1.0 °C) (Figure S17, Supporting Information). The latter shows that the helicity of the protein decreases sigmoidally over the range from \approx 35 to 75 °C, dropping to a value corresponding to a completely unfolded state, and that the ellipticity also changes at temperatures below 35 °C, suggesting some temperature-induced conformational change.

The analysis of the DSC profile of KHSRP₁₃₀₋₅₀₃ showed that the curve is not completely symmetrical and cannot be well described by a two-state melting process. Deconvolution analysis of the DSC curve revealed that the unfolding of the protein proceeds in two stages (Figure 6B), with the first transition ($T_m \approx 54$ °C) occurring with a smaller enthalpy change than the second ($T_m \approx 60$ °C). Considering the structure of the protein, one could conclude that some domains may be slightly less stable than others. Studies reported in the literature support this hypothesis and suggest that KH1 should be, among the four domains, the significantly less stable one.^[66]

In the presence of *c-Kit2* G4, it is observed that the complex appears more thermostable than the free protein (apparent $T_m = 63.7 \pm 0.5$ °C). The calorimetric enthalpy change, obtained by direct integration of the area under the DSC curve, is significantly larger than the value obtained by analyzing the curve of KHSRP₁₃₀₋₅₀₃ ($\Delta_{cal}H = 531$ and 402 kJ mol⁻¹, respectively). An important observation is that the thermal transition of free G4 is not detectable at the concentrations used for this experiment (dashed line in Figure 6) and therefore its contribution can be neglected. The DSC curve of the protein/G4 mixture clearly shows a multistate transition, therefore deconvolution of the experimental curve was performed. Such analysis showed that the

thermal transition can be modeled by a minimum of three transitions (Figure 6C), the first two of which have T_m values very close to those of the two transitions observed for the protein alone ($T_m \approx 54$ and 62 °C, respectively), while the third transition occurs at a higher temperature ($T_m \approx 72$ °C). These results suggest that the first two transitions may correspond to protein domains not involved in G4 binding, while the third should be due to the protein domains that interact with DNA in the complex.

2.7. Pyridostatin Hinders the Interaction of KHSRP₁₃₀₋₅₀₃ with *c-Myc* and *c-Kit2* G4s

We also evaluated the effect of pyridostatin, a first-in-class G4-binding ligand,^[67] on the interaction of KHSRP₁₃₀₋₅₀₃ with *c-Kit2* and *c-Myc* G4s. First, NMR titrations of the KHSRP₁₃₀₋₅₀₃/G4 complexes with the ligand were performed. After the addition of pyridostatin to KHSRP₁₃₀₋₅₀₃ in complex with G4 DNAs, the protein signals were clearly affected by intensity changes. In detail, a group of the signals that experienced a decrease in intensity after the addition of G4s, partially restored their original intensity upon pyridostatin addition (Figure 7). In the presence of *c-Myc*, these signals belong to residues located in the KH1 and KH3 domains, whereas in the presence of *c-Kit2*, the signals influenced by the addition of pyridostatin belong mainly to the KH4 domain. These effects imply that pyridostatin can affect the interaction of KHSRP₁₃₀₋₅₀₃ with both G4s.

To verify these data using an orthogonal biophysical assay, SPR experiments were again carried out on *c-Myc* and *c-Kit2* G4s when complexed to the ligand. SCK experiments were performed on KHSRP₁₃₀₋₅₀₃-immobilized sensor chip, flowing the G4s in the presence of pyridostatin. Increasing concentrations of the G4/pyridostatin complexes were flowed on the immobilized protein as reported above. A comparison between the resulting sensorgrams and those obtained in the absence of pyridostatin (Figure S18, Supporting Information) reveals that the presence of this G4 binder certainly has a significant impact on the recognition of G4s by the protein.

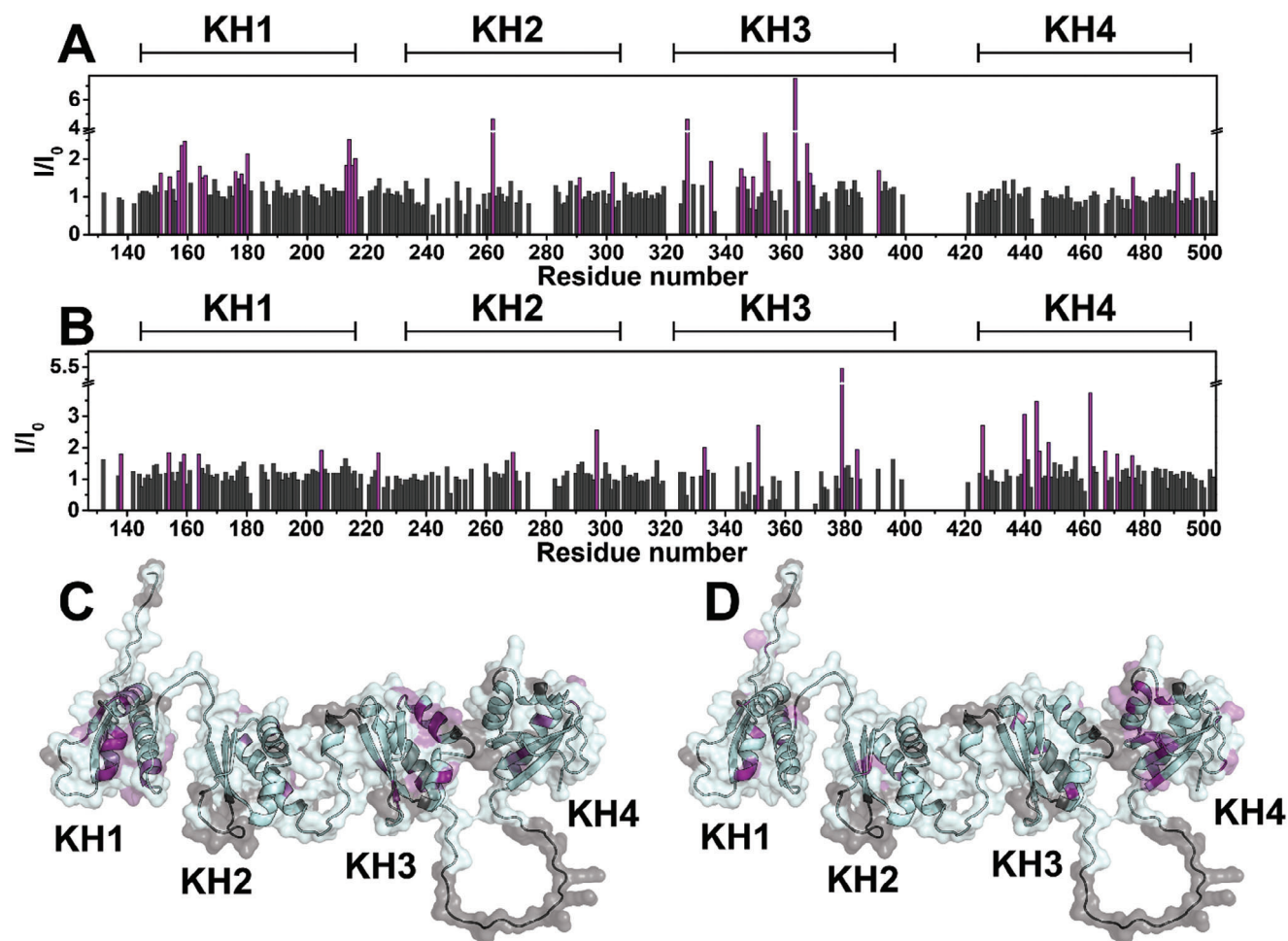


Figure 7. Graphical representation of the per-residue intensity changes of (A,C) KHSRP₁₃₀₋₅₀₃/*c-Myc* and (B,D) KHSRP₁₃₀₋₅₀₃/*c-Kit2* complexes (50 μ m) in the presence of pyridostatin (200 μ m for *c-Myc* and 100 μ m *c-Kit2*) compared to the protein in the presence of *c-Myc* or *c-Kit2* alone. The residues exhibiting the highest increases in signal intensities (*c-Myc*: Val151, Gly154, Gly157, Leu158, Ile159, Gly164, Glu165, Gln166, Cys176, Val178, Ile180, Ser213, Arg214, Gly215, Arg216, Arg262, Lys291, Ile302, Val327, Gly335, Ile345, Lys346, Gln349, Gly353, Val354, Gly363, Glu367, Lys368, Gln391, Ile476, Ile491, Glu496; *c-Kit2*: His138, Gly154, Ile159, Gly164, Lys205, His224, Leu269, Glu297, Ser333, Asp351, Cys379, Arg384, Glu426, Ile440, Gly444, Glu445, Lys448, Arg462, Asn467, Asn471, Ile476) are colored in violet in the plots (A,B) and on the AlphaFold model of the protein (C,D). Residues that could not be assigned in the NMR spectra are colored in gray on the protein model.

2.8. Biological Validation of KHSRP/G4 Interaction

In the latter part of this study, we aimed to validate the biophysical data by testing the ability of KHSRP to bind G4 structures in cells. For this purpose, BJ-EHLT cells – a model of transformed human foreskin fibroblasts expressing KHSRP (Figure S19A, Supporting Information) – were subjected to immunofluorescence (IF) experiments to evaluate the co-localization between the cell-endogenous KHSRP and the G4 structures. Of note, confocal microscopy analysis evidenced numerous co-localization spots within the cell nuclei (Figure 8A), indicating that endogenous KHSRP can interact with G4 structures spontaneously originating within the human genome, in line with what was already demonstrated with the truncated form of the protein in vitro. Furthermore, experiments performed by combining immunofluorescence of KHSRP and fluorescence in situ hybridization (FISH) assays – carried out by labeling *c-KIT* and *c-MYC* with specific probes – clearly

showed that KHSRP localizes in the proximity of these genes (Figure 8B).

Based on these findings, and taking into account the results of biophysical analyses, we focused on gaining deeper insights into the capability of KHSRP to bind the well-known G4s localized within the promoters of *c-KIT* and *c-MYC*. To this aim, we performed a chromatin immunoprecipitation (ChIP) assay to evaluate the amount of KHSRP associated with the G4 regions of interest. For these experiments, BJ-EHLT cells were first subjected to treatment with a cross-linking agent (formaldehyde) and, upon stabilization of protein-DNA interactions, KHSRP was immunoprecipitated and the G4 motifs localized within the promoter of *c-KIT* and *c-MYC* were amplified and quantified by real-time qPCR. Of note, the results evidenced a significant enrichment of KHSRP associated with the G4-forming sequences of each gene promoter (Figure 8C). For completeness, we also assessed the ability of KHSRP to bind telomeric G4s. Interestingly, the ChIP assays performed with primers designed to pair with

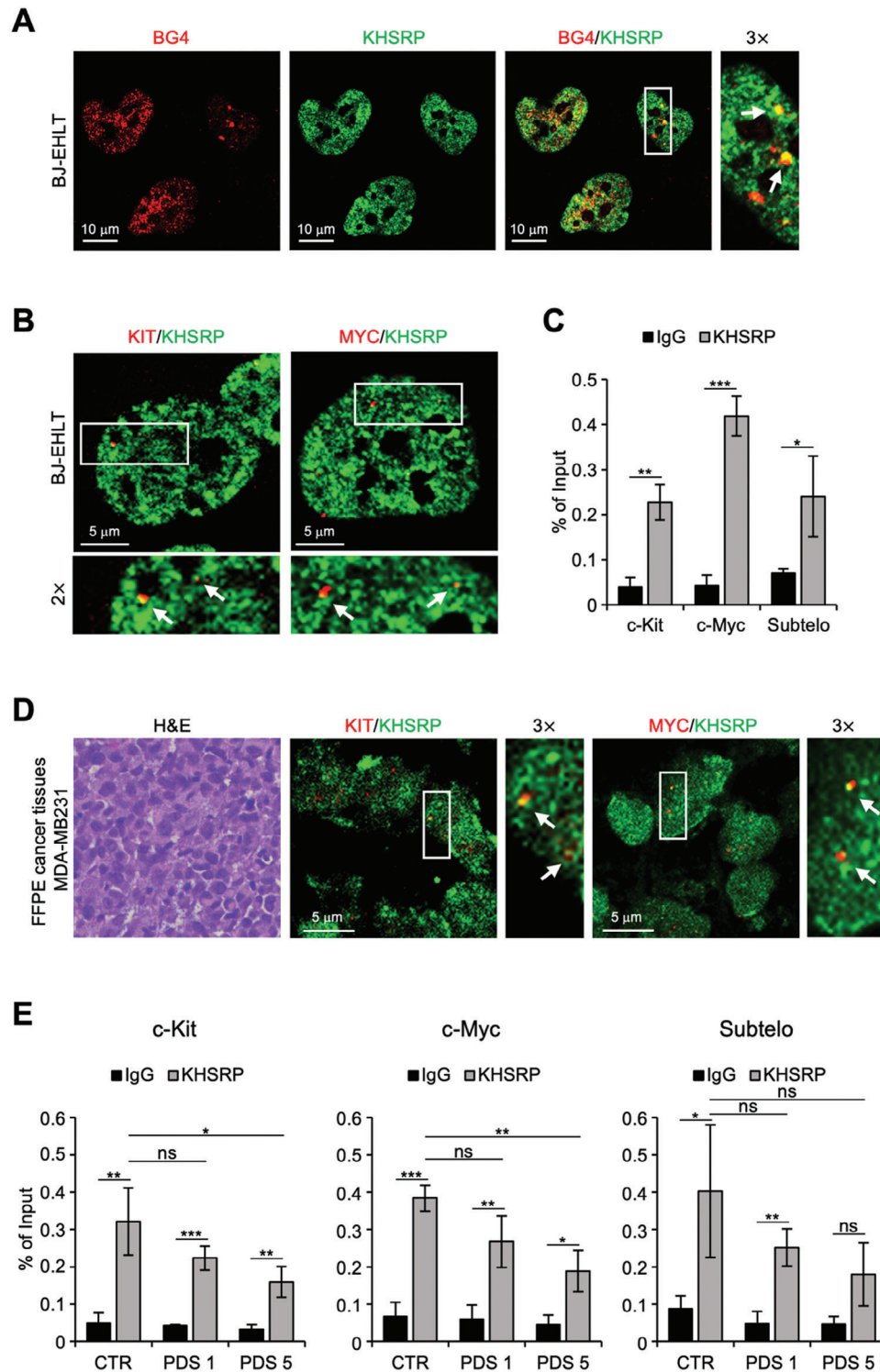


Figure 8. KHSRP binds to G4 structures located within the promoters of *c-KIT* and *c-MYC* in cells. A) Representative images of immunofluorescence (IF) analyses performed in human BJ-EHLT fibroblast using the antibodies against G4 structures (anti-BG4, red spots) and KHSRP (green spots). Arrows indicate the colocalization (yellow spots). Specific enlargement (2x) is shown. Scale bar: 10 μ m. B) Representative images of IF-combined FISH analysis (IF-FISH) performed in BJ-EHLT using specific probes for *c-Kit* and *c-Myc* sequences (red spots) and the antibody against KHSRP (green spots). Arrows indicate the colocalization (yellow spots). Specific enlargements (3x) are shown. Scale bar: 5 μ m. C) Chromatin Immunoprecipitation (ChIP) assays performed in BJ-EHLT fibroblasts by immunoprecipitating protein-DNA complex with the antibody against KHSRP. Rabbit IgG was used as a negative control. KHSRP enrichment at G4 forming sequence of the promoter of *c-KIT* and *c-MYC* and at subtelomeric regions was evaluated by qPCR Real-Time analyses. The histogram represents the quantification of KHSRP enrichment, expressed as the percentage of the input. D) Ex vivo analysis on FFPE

sub-telomeric regions (Table S1, Supporting Information) confirmed that KHSRP also effectively binds telomeric G4s in the cellular environment (Figure 8C), in agreement with our previous findings.^[46]

The same experiments were then extended to HeLa and MCF7 cells, two human cancer cell lines deriving from cervical carcinoma and breast adenocarcinoma, respectively, both expressing KHSRP (Figure S19A, Supporting Information). Interestingly, IF analyses (Figure S19B, Supporting Information), as well as ChIP assays (Figure S19C,D, Supporting Information), fully recapitulated what was previously observed in BJ-EHLT cells, indicating that the ability of KHSRP to bind the G4 structures located within the promoter regions of *c-KIT* and *c-MYC* is a generalizable phenomenon, independent of tumor histotype.

These findings were further confirmed in a xenograft model obtained by the intramuscular injection of MDA-MB-231, a breast cancer cell line of human origin, within the leg of immune-compromised mice. In particular, FISH assays performed on tumor sections highlighted that KHSRP, even in the complexity of an in vivo model, maintains its ability to localize within the sequences of *c-KIT* and *c-MYC* (Figure 8D), further underscoring the biological relevance of our findings.

To complete the biological characterization of the interaction between KHSRP and G4s in cells, ChIP assays were performed in BJ-EHLT cells exposed to increasing concentrations of pyridostatin (1 and 5 μM for 24 h). Interestingly, the G4 ligand was found to impair, in a dose-dependent manner, the binding of KHSRP to G4 structures (Figure 8E). This indicates that pyridostatin can displace KHSRP from G4s and suggests that they compete for the same binding region on the G4-forming DNA, consistent with the results of NMR experiments. Altogether, these findings corroborate the direct binding of KHSRP to G4 structures.

3. Discussion

G4 structures play a key role in the regulation of diverse cellular processes and therefore are one of the crucial secondary structures in nucleic acids. In particular, DNA G4s have been shown to play significant roles in the complex mechanism of gene expression, in which they participate, along with a network of interacting proteins, in the regulation of transcription. Precise orchestration of G4 formation, stabilization, and resolution is thus imperative for ensuring proper biological functions.

Within this context, the recognition and processing of G4s by nucleic acid-directed proteins emerge as pivotal events, influencing the activation or deactivation of physiological and pathological pathways. Some proteins act as mediators, detecting G4s and recruiting helicases or other proteins to facilitate the unfolding of noncanonical structures. At the same time, distinct proteins possess the ability to directly recognize and destabilize G4s, while others have the opposite function of stabilizing them.

These interactors alter the equilibrium between unfolded G-rich sequences and folded G4s and make G4 formation very dynamic within living cells.

To date, a wide range of G4-binding proteins has been discovered. Whilst there are multiple studies describing their identification, for most of them the molecular basis of interactions remains to be understood. Furthermore, while some proteins exhibit a broad nucleic acids-binding capability, others act in a more selective manner and with different affinity, but virtually all of them have multiple roles. KHSRP has been shown to be among those proteins capable of interacting with multiple RNA and DNA targets, thanks to its modular structure and binding flexibility.

In this investigation, we conducted biophysical and biological studies to gain insights into the G4 DNA recognition by the KHSRP protein. We focused on biologically relevant G4-forming sequences found in the promoter regions of *c-MYC*, *c-KIT*, and *BCL-2* oncogenes, as well as in human telomeric DNA. Initially, we optimized protein expression and purification, followed by an NMR study of KHSRP₁₃₀₋₅₀₃, which allowed us to assign backbone residues. Then, we focused on studying G4-protein interactions in vitro. Using various biophysical methodologies (NMR, SPR, CD melting, and DSC), we found that KHSRP binds to the investigated G4s, albeit with differences in terms of affinity and binding mode. Both protein-based and DNA-based 1D ¹H NMR experiments provided evidence of a better interaction with the parallel-stranded G4s (*c-Kit1*, *c-Kit2*, and *c-Myc*) compared to hybrid topology G4s (*Bcl-2* and *mTel₂₄*). As for *c-Kit1*, *c-Kit2*, and *c-Myc* G4s, the DNA residues experiencing the most significant perturbations were those belonging to the external G-tetrads, indicating the involvement of G-tetrad planes in the protein binding process. This could also explain the reduced effects observed in interactions between the protein and hybrid G4s in which the external G-tetrads are predominantly occupied by loops, decreasing their susceptibility to interaction.

Interestingly, no unfolding or structural alterations in DNA secondary structures were detected upon binding to KHSRP. Instead, stronger interactions resulted in enhanced thermal stability of the G4 structures induced by the protein, indicating that KHSRP preferentially binds to the folded form of DNA molecules. In this context, it cannot be excluded that KHSRP may serve as a mediator to recruit other proteins, including helicases, either directly or indirectly. For instance, as mentioned in the introduction, KHSRP interacts with hnRNPA1,^[35] a protein known to resolve G4 structures at telomeres and in the promoters of genes such as *KRAS* and *TRA2B*,^[40,42] thereby modulating their transcription in cancer cells. Additionally, KHSRP is also part of a multiprotein complex that includes hnRNPF and hnRNPH1, which have recently been shown to interact with RNA G4s.^[27] While the involvement of these proteins with DNA G4 structures has not yet been demonstrated, this possibility cannot be ruled out.

cancer tissues derived from MDA-MB-231 tumor tissue. *Left panel*: representative images of H&E staining; scale bar: 40 μm . *Central and Right panels*: Representative images of IF-FISH analysis of KHSRP and *c-Kit* and *c-Myc* genes performed as in (B). Arrows indicate the colocalization (yellow spots). Specific enlargements (3 \times) are shown. Scale bar: 5 μm . *E*) ChIP assays performed in BJ-EHLT fibroblasts treated with pyridostatin (PDS) for 24 h at the indicate doses and performed as in (C). All representative images of IF and IF-FISH were acquired by confocal microscopy (magnification 63 \times). All histograms represent mean values \pm S.D. of three independent experiments. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

To gain insights into the thermodynamic and kinetic aspects of binding, the interaction of KHSRP with the five G4s was examined by SPR. These experiments indicated specific DNA/protein interactions and showed that G4 structural variability can significantly influence protein recognition. Indeed, the affinity for parallel-stranded G4s was found to be at least one order of magnitude higher than that measured for hybrid G4s, in agreement with the NMR data. Despite the different binding affinities, these findings don't exclude the possibility that this protein might have a broad role as G4-interactor, not being selective for a single G4. The quantitative estimate of on/off rates of binding revealed that the difference in affinity between parallel and hybrid G4s arises mainly from dissociation rates rather than association rates. Given the structural differences between parallel and non-parallel G4s, the reduced protein affinity for non-parallel G4s can be attributed to the presence of loop residues above the G-tetrads. Although these loops do not appear to impede recognition, they may restrict the formation of more extensive or complete interactions. Among the parallel G4s, KHSRP showed the highest affinity for *c-Kit2*, followed by *c-Myc*, with kinetic analysis indicating that the difference lies in the association rates.

The 2D NMR data on the interaction of the protein with *c-Kit2* and *c-Myc* G4s suggested a major role for the KH3 domain, with varying participation from the other domains depending on the target. This is consistent with a multi-domain mechanism of KHSRP-G4 recognition, as also observed for KHSRP binding to RNA targets.^[37] Our calorimetric investigations on the stability of KHSRP in the absence and presence of *c-Kit2* G4, have shown that the stability of the protein, or rather of some of its domains, changes upon association with the DNA structure, as a consequence of the interaction and the specific involvement of the different domains in the binding to DNA.

Interestingly, the results of our experiments performed in the presence of pyridostatin, a well-characterized G4 ligand, clearly showed that it can affect G4/protein interactions in vitro and in cells, further corroborating that KHSRP is capable of directly binding to G4 structures. As documented in the literature, G4-targeting agents are known to modulate gene expression by altering the formation of G4s within gene regulatory elements, such as gene promoters and/or enhancers. However, the impact of pyridostatin on gene expression is quite complex. Indeed, it can cause DNA double-strand breaks,^[68] and displace G4-binding proteins.^[69,70] Together, these results suggest that pyridostatin may compete with the protein in binding the G4s, potentially influencing other processes resulting from protein-DNA recognition.

Overall, our results provided valuable insights into the structural and energetic aspects of the KHSRP-G4 interaction and shed light on the involvement of protein domains in the molecular recognition of these noncanonical DNA structures, as well as on the G4 structural determinants that contribute to the interaction. Additionally, validation of these findings in cancer cells and preclinical in vivo models highlighted their biological relevance, suggesting that KHSRP may play a significant role in cancer formation and progression. This paves the way for further investigations aimed at exploring potential therapeutic developments related to these findings.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

P.R. and P.Z. contributed equally to this work. The research leading to these results has received funding from the Italian Association for Cancer Research (IG 24590 to B.P., IG 21579 to A.B.). The authors acknowledge the support and the use of resources of Instruct-ERIC, a landmark ESFRI project, and specifically the CERM/CIRMMIP Italy centre, the "Progetto Dipartimenti di Eccellenza 2023–2027" to the Department of Chemistry "Ugo Schiff" of the University of Florence, the Recombinant Proteins JOYNLAB laboratory and the project "Potentiating the Italian Capacity for Structural Biology Services in Instruct Eric (ITACA.SB)" (IR0000009) within the call MUR 3264/2021 PNRR M4/C2/L3.1.1, funded by the European Union NextGenerationEU (CN00000041). This study was also supported by the NextGenerationEU project within the MUR PNRR "National Center for Gene Therapy and Drugs based on RNA Technology".

Open access publishing facilitated by Università degli Studi di Napoli Federico II, as part of the Wiley - CRUI-CARE agreement.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The protein resonance assignment has been deposited in the BMRB database under the accession code: 52573. The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

biophysical methods, gene promoters, G-quadruplex DNA, protein-DNA interactions

Received: August 22, 2024
Revised: December 3, 2024
Published online: January 6, 2025

- [1] P. Briata, C.-Y. Chen, A. Ramos, R. Gherzi, *Biochim. Biophys. Acta, Gene Regul. Mech.* **2013**, 1829, 689.
- [2] R. Gherzi, C.-Y. Chen, A. Ramos, P. Briata, *Semin. Cell Dev. Biol.* **2014**, 34, 2.
- [3] R. Gherzi, C. Chen, M. Trabucchi, A. Ramos, P. Briata, *Wiley Interdiscip. Rev. RNA* **2010**, 1, 230.
- [4] T. Davis-Smyth, R. C. Duncan, T. Zheng, G. Michelotti, D. Levens, *J. Biol. Chem.* **1996**, 271, 31679.
- [5] P. Briata, D. Bordo, M. Puppo, F. Gorlero, M. Rossi, N. Perrone-Bizzozero, R. Gherzi, *WIREs RNA* **2016**, 7, 227.
- [6] K.-A. Palzer, V. Bolduan, R. Käfer, H. Kleinert, M. Bros, A. Pautz, *Cell* **2022**, 11, 1482.
- [7] M. Giovarelli, G. Bucci, A. Ramos, D. Bordo, C. J. Wilusz, C.-Y. Chen, M. Puppo, P. Briata, R. Gherzi, *Proc. Natl. Acad. Sci. USA* **2014**, 111, E5023.
- [8] J.-Y. Lin, M.-L. Li, S.-R. Shih, *Nucleic Acids Res.* **2009**, 37, 47.
- [9] M. Puppo, G. Bucci, M. Rossi, M. Giovarelli, D. Bordo, A. Moshiri, F. Gorlero, R. Gherzi, P. Briata, *Cell Rep.* **2016**, 16, 967.
- [10] H. Min, C. W. Turck, J. M. Nikolic, D. L. Black, *Genes Dev.* **1997**, 11, 1023.

- [11] M. Trabucchi, P. Briata, M. Garcia-Mayoral, A. D. Haase, W. Filipowicz, A. Ramos, R. Gherzi, M. G. Rosenfeld, *Nature* **2009**, 459, 1010.
- [12] Y.-Y. Wang, X.-L. Gu, C. Wang, H. Wang, Q.-C. Ni, C.-H. Zhang, X.-F. Yu, L.-Y. Yang, Z.-X. He, G.-X. Mao, S.-Y. Yang, *Tumor Biol.* **2016**, 37, 9755.
- [13] R. Wang, B. Wei, J. Wei, Z. Li, Y. Tian, C. Du, *J. Clin. Neurosci.* **2016**, 33, 259.
- [14] J. Xu, D. Wang, H. Ma, X. Zhai, Y. Huo, Y. Ren, W. Li, L. Chang, D. Lu, Y. Guo, Y. Si, Y. Gao, X. Wang, Y. Ma, F. Wang, J. Yu, *Blood Sci.* **2022**, 4, 103.
- [15] M. Yan, L. Sun, J. Li, H. Yu, H. Lin, T. Yu, F. Zhao, M. Zhu, L. Liu, Q. Geng, H. Kong, H. Pan, M. Yao, *J. Exp. Clin. Cancer Res.* **2019**, 38, 478.
- [16] Q. Gou, L. Gao, X. Nie, W. Pu, J. Zhu, Y. Wang, X. Liu, S. Tan, J.-K. Zhou, Y. Gong, J. He, K. Wu, Y. Xie, W. Zhao, L. Dai, L. Liu, R. Xiang, Y.-Q. Wei, L. Zhang, Y. Peng, *Cancer Res.* **2018**, 78, 4163.
- [17] K. Taniuchi, M. Ogasawara, *Oncotarget* **2020**, 11, 131.
- [18] D. J. Patel, A. T. Phan, V. Kuryavyi, *Nucleic Acids Res.* **2007**, 35, 7429.
- [19] S. Neidle, *J. Biol. Chem.* **2021**, 296, 100553.
- [20] E. Largy, J.-L. Mergny, V. Gabelica, *The Alkali Metal Ions: Their Role for Life* (Eds.: A. Sigel, H. Sigel, R. Sigel), Springer, Switzerland **2016**, pp. 203–258.
- [21] F. Romano, A. Di Porzio, N. Iaccarino, G. Riccardi, R. Di Lorenzo, S. Laneri, B. Pagano, J. Amato, A. Randazzo, *Expert Opin. Ther. Pat.* **2023**, 33, 745.
- [22] D. Rhodes, H. J. Lipps, *Nucleic Acids Res.* **2015**, 43, 8627.
- [23] T. Tian, Y.-Q. Chen, S.-R. Wang, X. Zhou, *Chem* **2018**, 4, 1314.
- [24] H. Shu, R. Zhang, K. Xiao, J. Yang, X. Sun, *Biomolecules* **2022**, 12, 648.
- [25] O. Mendoza, A. Bourdoncle, J.-B. Boulé, R. M. Brosh, J.-L. Mergny, *Nucleic Acids Res.* **2016**, 44, 1989.
- [26] J. Amato, L. Cerofolini, D. Brancaccio, S. Giuntini, N. Iaccarino, P. Zizza, S. Iachettini, A. Biroccio, E. Novellino, A. Rosato, M. Fragai, C. Luchinat, A. Randazzo, B. Pagano, *Nucleic Acids Res.* **2019**, 47, 9950.
- [27] A. Decorsière, A. Cayrel, S. Vagner, S. Millevoi, *Genes Dev.* **2011**, 25, 220.
- [28] J.-F. Fiset, D. R. Montagna, M.-R. Mihailescu, M. S. Wolfe, *J. Neurochem.* **2012**, 121, 763.
- [29] A. C. Krüger, M. K. Raarup, M. M. Nielsen, M. Kristensen, F. Besenbacher, J. Kjems, V. Birkedal, *Eur. Biophys. J.* **2010**, 39, 1343.
- [30] A. von Hacht, O. Seifert, M. Menger, T. Schütze, A. Arora, Z. Konthur, P. Neubauer, A. Wagner, C. Weise, J. Kurreck, *Nucleic Acids Res.* **2014**, 42, 6630.
- [31] X. Zhang, J. Spiegel, S. Martínez Cuesta, S. Adhikari, S. Balasubramanian, *Nat. Chem.* **2021**, 13, 626.
- [32] V. Markovtsov, J. M. Nikolic, J. A. Goldman, C. W. Turck, M.-Y. Chou, D. L. Black, *Mol. Cell. Biol.* **2000**, 20, 7463.
- [33] T. Vo, T. Brownmiller, K. Hall, T. L. Jones, S. Choudhari, I. Grammatikakis, K. R. Ludwig, N. J. Caplen, *Nucleic Acids Res.* **2022**, 50, 6474.
- [34] H. Huang, J. Zhang, S. E. Harvey, X. Hu, C. Cheng, *Genes Dev.* **2017**, 31, 2296.
- [35] T. Ruggiero, M. Trabucchi, M. Ponassi, G. Corte, C.-Y. Chen, L. Al-Haj, K. S. Khabar, P. Briata, R. Gherzi, *BMC Mol. Biol.* **2007**, 8, 28.
- [36] G. Michlewski, J. F. Cáceres, *Nat. Struct. Mol. Biol.* **2010**, 17, 1011.
- [37] M. F. García-Mayoral, I. Díaz-Moreno, D. Hollingworth, A. Ramos, *Nucleic Acids Res.* **2008**, 36, 5290.
- [38] Q.-S. Zhang, L. Manche, R.-M. Xu, A. R. Krainer, *RNARNA* **2006**, 12, 1116.
- [39] M. Ghosh, M. Singh, *Nucleic Acids Res.* **2018**, 46, 10246.
- [40] M. Paramasivam, A. Membrino, S. Cogoi, H. Fukuda, H. Nakagama, L. E. Xodo, *Nucleic Acids Res.* **2009**, 37, 2841.
- [41] A. Ferino, J. Marquevielle, H. Choudhary, G. Cinque, C. Robert, A. Bourdoncle, R. Picco, J.-L. Mergny, G. F. Salgado, L. E. Xodo, *ACS Omega* **2021**, 6, 34092.
- [42] T. Nishikawa, Y. Kuwano, Y. Takahara, K. Nishida, K. Rokutan, *Sci. Rep.* **2019**, 9, 10276.
- [43] R. Rigo, M. Palumbo, C. Sissi, *Bioch. Biophys. Acta - Gen. Subj.* **2017**, 1861, 1399.
- [44] J. Amato, A. Pagano, D. Capasso, S. Di Gaetano, M. Giustiniano, E. Novellino, A. Randazzo, B. Pagano, *ChemMedChem* **2018**, 13, 406.
- [45] S. Pelliccia, J. Amato, D. Capasso, S. Di Gaetano, A. Massarotti, M. Piccolo, C. Irace, G. C. G. C. Tron, B. Pagano, A. Randazzo, E. Novellino, M. Giustiniano, *J. Med. Chem.* **2020**, 63, 2035.
- [46] B. Pagano, L. Margarucci, P. Zizza, J. Amato, N. Iaccarino, C. Cassiano, E. Salvati, E. Novellino, A. Biroccio, A. Casapullo, A. Randazzo, *Chem. Commun.* **2015**, 51, 2964.
- [47] R. Nicastro, M. F. García-Mayoral, D. Hollingworth, G. Kelly, S. R. Martin, P. Briata, R. Gherzi, A. Ramos, *Nat. Struct. Mol. Biol.* **2012**, 19, 1282.
- [48] M. F. García-Mayoral, D. Hollingworth, L. Masino, I. Díaz-Moreno, G. Kelly, R. Gherzi, C.-F. Chou, C.-Y. Chen, A. Ramos, *Structure* **2007**, 15, 485.
- [49] R. Gherzi, K.-Y. Lee, P. Briata, D. Wegmüller, C. Moroni, M. Karin, C.-Y. Chen, *Mol. Cell* **2004**, 14, 571.
- [50] Y.-A. Kung, C.-T. Hung, K.-Y. Chien, S.-R. Shih, *Nucleic Acids Res.* **2017**, 45, 271.
- [51] M. Varadi, S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, A. Židek, T. Green, K. Tunyasuvunakool, S. Petersen, J. Jumper, E. Clancy, R. Green, A. Vora, M. Lutfi, M. Figurnov, A. Cowie, N. Hobbs, P. Kohli, G. Kleywegt, E. Birney, D. Hassabis, S. Velankar, *Nucleic Acids Res.* **2022**, 50, D439.
- [52] A. T. Phan, V. Kuryavyi, S. Burge, S. Neidle, D. J. Patel, *J. Am. Chem. Soc.* **2007**, 129, 4386.
- [53] V. Kuryavyi, A. T. Phan, D. J. Patel, *Nucleic Acids Res.* **2010**, 38, 6757.
- [54] A. Ambrus, D. Chen, J. Dai, R. A. Jones, D. Yang, *Biochemistry* **2005**, 44, 2048.
- [55] J. Dai, D. Chen, R. A. Jones, L. H. Hurley, D. Yang, *Nucleic Acids Res.* **2006**, 34, 5133.
- [56] K. N. Luu, A. T. Phan, V. Kuryavyi, L. Lacroix, D. J. Patel, *J. Am. Chem. Soc.* **2006**, 128, 9963.
- [57] E. Barile, M. Pellecchia, *Chem. Rev.* **2014**, 114, 4749.
- [58] M. Adrian, B. Heddi, A. T. Phan, *Methods* **2012**, 57, 11.
- [59] J. Majka, C. Speck, *In Advances in Biochemical Engineering/Biotechnology*, Springer, Singapore **2007**, pp. 13–36.
- [60] P. Cavaliere, B. Pagano, V. Granata, S. Prigent, H. Rezaei, C. Giancola, A. Zagari, *Nucleic Acids Res.* **2013**, 41, 327.
- [61] A. M. Ogloblina, N. Iaccarino, D. Capasso, S. D. Gaetano, E. U. Garzarella, N. G. Dolinnaya, M. G. Yakubovskaya, B. Pagano, J. Amato, A. Randazzo, *Int. J. Mol. Sci.* **2020**, 21, 7781.
- [62] J. Kypr, I. Kejnovska, D. Renciuik, M. Vorlickova, *Nucleic Acids Res.* **2009**, 37, 1713.
- [63] A. I. Karsiotis, N. M. Hessari, E. Novellino, G. P. Spada, A. Randazzo, M. Webba da Silva, *Angew. Chem., Int. Ed.* **2011**, 50, 10645.
- [64] I. Jelesarov, H. R. Bosshard, *J. Mol. Recognit.* **1999**, 12, 3.
- [65] P. L. Privalov, A. I. Dragan, *Biophys. Chem.* **2007**, 126, 16.
- [66] I. Díaz-Moreno, D. Hollingworth, T. A. Frenkiel, G. Kelly, S. Martin, S. Howell, M. García-Mayoral, R. Gherzi, P. Briata, A. Ramos, *Nat. Struct. Mol. Biol.* **2009**, 16, 238.
- [67] S. Müller, S. Kumari, R. Rodriguez, S. Balasubramanian, *Nat. Chem.* **2010**, 2, 1095.
- [68] R. Rodriguez, K. M. Miller, J. V. Forment, C. R. Bradshaw, M. Nikan, S. Britton, T. Oelschlaegel, B. Xhemalce, S. Balasubramanian, S. P. Jackson, *Nat. Chem. Biol.* **2012**, 8, 301.
- [69] L. Li, P. Williams, Z. Gao, Y. Wang, *Nucleic Acids Res.* **2020**, 48, 11994.
- [70] L. Li, P. Williams, W. Ren, M. Y. Wang, Z. Gao, W. Miao, M. Huang, J. Song, Y. Wang, *Nat. Chem. Biol.* **2021**, 17, 161.