Review

Dynamics Studies of DNA with Non-canonical Structure Using NMR Spectroscopy

Kwang-Im Oh [1],†, Jinwoo Kim [2],†, Chin-Ju Park [2],* and Joon-Hwa Lee [1],*

[1]   Department of Chemistry and RINS, Gyeongsang National University, Gyeongnam 52828, Korea;
     orangekwang@gmail.com
[2]   Department of Chemistry, Gwangju Institute of Science and Technology, Gwangju 61005, Korea;
     wlsdn8810@gist.ac.kr
*   Correspondence: cjpark@gist.ac.kr (C.-J.P.); joonhwa@gnu.ac.kr (J.-H.L.); Tel.: +82-62-715-3630 (C.-J.P.);
     +82-55-772-1490 (J.-H.L.)
†   These authors contributed equally to this work.

check for
updates

Abstract: The non-canonical structures of nucleic acids are essential for their diverse functions during various biological processes. These non-canonical structures can undergo conformational exchange among multiple structural states. Data on their dynamics can illustrate conformational transitions that play important roles in folding, stability, and biological function. Here, we discuss several examples of the non-canonical structures of DNA focusing on their dynamic characterization by NMR spectroscopy: (1) G-quadruplex structures and their complexes with target proteins; (2) i-motif structures and their complexes with proteins; (3) triplex structures; (4) left-handed Z-DNAs and their complexes with various Z-DNA binding proteins. This review provides insight into how the dynamic features of non-canonical DNA structures contribute to essential biological processes.

Keywords: nucleic acids; NMR; G-quadruplex; I-motif; triple helix; Z-DNA; dynamics

1. Introduction

Nucleic acids play an important role in all biological processes related to genetic information such as replication, transcription, and translation. DNA duplexes typically form a right-handed B-form double-helical structure containing Watson-Crick A·T and G·C base-pairs. With two complementary strands are paired, RNA preferentially has A-form right-handed double helix. In other cases, RNA can adopt diverse non-canonical structures, such as left-handed helices, i-motifs, triplexes, and quadruplexes as major or minor conformations [1–5]. These structures contain unusual Hoogsteen and/or wobble base-pairs and are key structural components that are essential for the diverse functions of nucleic acids during various biological processes [6,7]. There are comprehensive studies and reviews of the structural studies of non-canonical RNA and base-pairing [8–11].

The non-canonical structures of nucleic acids can have multiple conformational folds, producing a highly dynamic structural ensemble. Even though the less populated conformations are not often visualized in structural studies, they can contribute to folding, stability, and biological function [12–15]. The exchange process can feature a transient snapshot of biologically important conformational states. Thus, in addition to structural studies of non-canonical nucleic acids, their dynamic features also should be investigated to fully understand the biological function of nucleic acids. Dynamics studies of protein-nucleic acid interactions are complementary to the complex structures, in that they provide detailed information about the binding interfaces, exchange rates, and conformational changes associated with the binding. In this review, we discuss several examples of non-canonically structured DNAs, focusing on their dynamic characterization by NMR spectroscopy: (1) G–quadruplex

(G4) structures and their complexes with target proteins; (2) i-motif structures and their complexes with proteins; (3) triplex structures; (4) left-handed Z-DNAs and their complexes with various Z-DNA binding proteins (ZBPs). This review provides insight into how the dynamic features of the non-canonical structures contribute to essential biological processes.

2. G-Quadruplex (G4)

The G-tetrad is a planar structure composed of four guanines with Hoogsteen base-pairs and large monovalent cations in the center of the plane (Figure 1A) [16]. G4s are structures with two or more stacked G-tetrads. Typically, four tracts of three Gs (G_3) are connected with various looping sequences, while variants with longer lengths of G tracts also form G4s [12,13,17]. Considering its high prevalence in the human genome, especially in oncogene promoters and telomere regions, the understanding of the dynamics of G4s and G4-protein (or ligand) complexes provides valuable information regarding cancers [18–20].
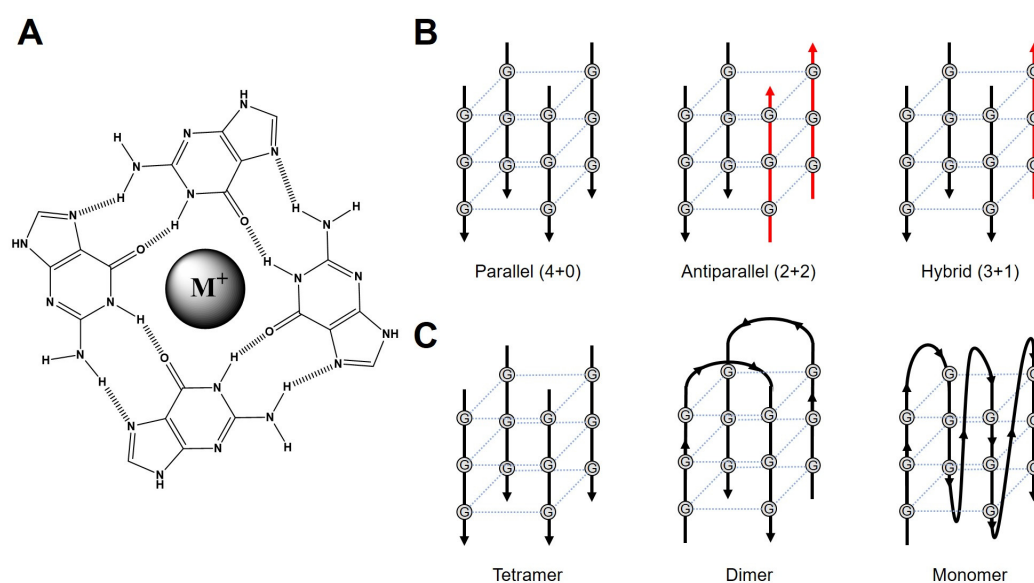

Figure 1. (**A**) G-tetrad formed by four guanines with Hoogsteen base-pairs and a monovalent cation. (**B**) G4 topologies depending on the orientations of the four G-tracts. (**C**) Examples of tetrameric, dimeric, and monomeric G4 structures.

G4 structures are conformationally heterogeneous because of the exchange of topology [21], G-register [22], and oligomeric states [23]. Depending on the relative orientations of the four G-tracts, G4 topologies are categorized as follows. The parallel is all four G-tracts are aligned in the same direction (4+0). The antiparallel indicates that two G-tracts are in the same direction, and the other two are in the opposite direction (2+2). The hybrid is that three G-tracts are in the same direction, and one is in the opposite (3+1) (Figure 1B). Loops connecting parallel G-tracts are usually short (1~2 nt) and called propellers. Loops connecting antiparallel tracts are longer (≥3 nt). All four G-tracts can be connected in sequence (intramolecular, monomeric) or can be separate (intermolecular, multimeric (dimeric, or tetrameric)) (Figure 1C). While extensive studies have been performed to reveal the structural aspects of the G4, the contributions of the dynamic properties of G4 are less acknowledged. However, it is becoming clear that the conformationally heterogeneous ensemble of G4s is essential for biological functions such as transcriptional repression in promoters [24].

There are excellent reviews that describe the details of NMR methods for studying the structures [25] and the dynamics of G4s [26]. In this article, we focus on recent dynamics studies of G4s alone and G4s in complex with proteins or other ligands.

2.1. G4 DNA

In the G-rich DNA sequence which forms unimolecular G4, four G-tracts are linked with three loops. When more than three G nucleotides are present within each tract, the different numbers of Gs in the four G-strands induce sequence-based dynamic polymorphism [6]. Different conditions of pH, cations, or ligands also can induce polymorphism. Depending on which G participates in the hydrogen-bonding networks of the G-tetrad, different isomers (G-register isomers) can be formed. Figure 2A shows that the G-register exchange can be coupled with topological transitions, as shown in the hTERT promoter sequence.
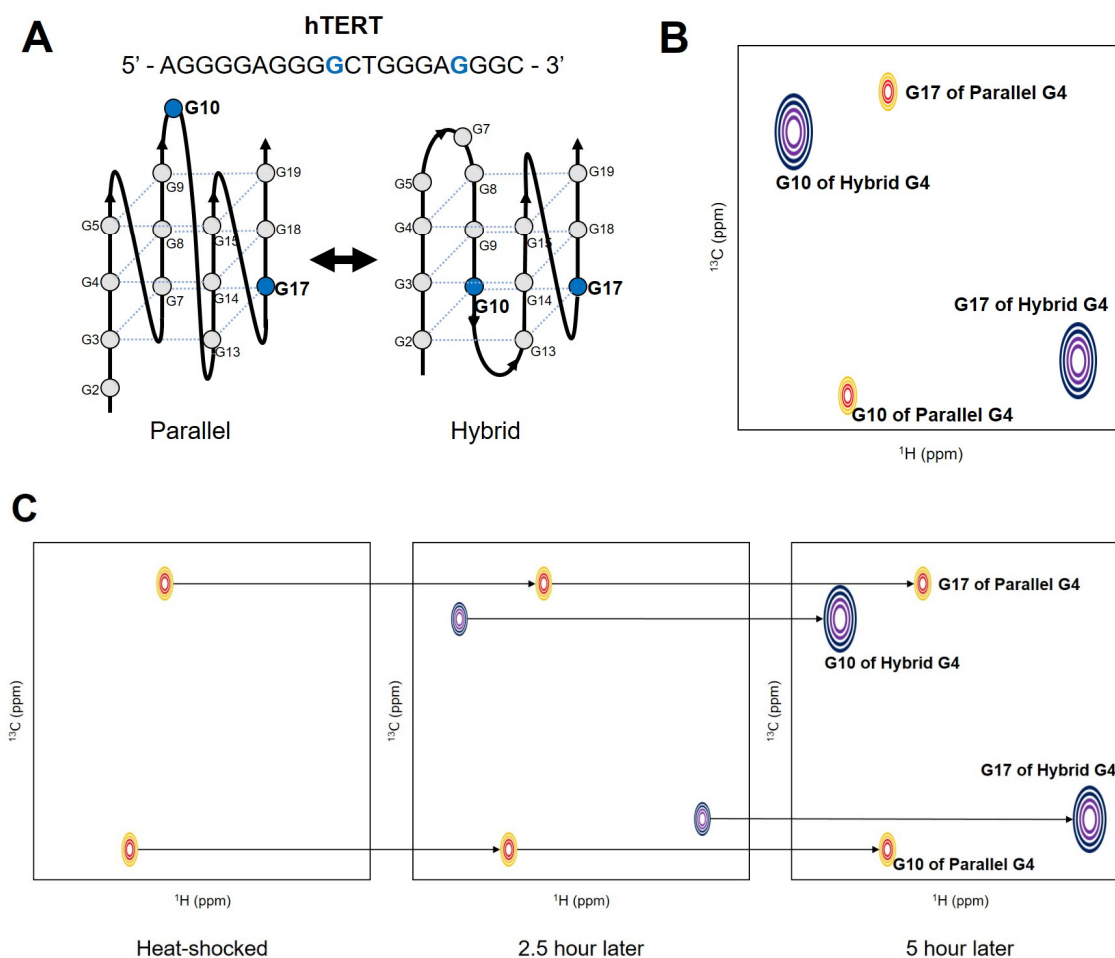

Figure 2. (**A**) G-register isomers of hTERT G4. The ^{13}C labeled G10 and G17 are shown in Blue. (**B**) Schematic drawing of ^{1}H-^{13}C HSQC spectra of hTERT G4 with ^{13}C labeled G10 and G17. (**C**) Schematic drawing of NMR monitoring of the re-equilibration process. Redrawn from Ref. [27] by permission of Oxford University Press on behalf of Nucleic Acids Research.

The base-pair opening occurs during structural fluctuations of DNA. To characterize the dynamic motions of base-pair opening, the imino protons are useful probes in NMR studies [28–33]. The exchange of imino protons provides quantitative information on the thermodynamics and kinetics of the open/closed state using a two-state model [30]. A recent study showed that NMR could be successfully applied to reveal the slow exchange process between G-register isomers of c-Myc and hTERT sequences [34]. By using a photolabile protecting group, a single conformation of each G4 was initially trapped, and the restoration of conformational heterogeneity after releasing the protecting group by irradiation was monitored with real-time ^{1}H NMR spectroscopy. The time dependence of the signal intensities of the imino protons from each isomer revealed that the re-equilibration occurs on a timescale of hours (the rate constants of exchange were measured as h^{-1}). To compare the real-time NMR

data with other kinetics experiments, the authors performed thermal hysteresis UV-Vis denaturation experiments to measure the folding/unfolding rate of the representative G4 isomers. K^+-induced folding measurements by time-resolved NMR spectroscopy were also employed. During isomer exchange, the c-Myc isomers maintain their partially folded intermediate structures, while the hTERT isomers follow an unfolding-refolding mechanism. The re-equilibration of conformational heterogeneity of the hTERT G4 was also monitored with atom-specific ^{13}C labeled dGs [27]. By incorporating 8-^{13}C-dG at the specific position, the authors observed two ^{1}H-^{13}C cross-peaks from one guanine, which indicates the existence of two distinct conformations in the slow chemical shift time regime. The re-equilibration process was monitored by measuring peak intensities in Band-selective Excitation Short-Transient (BEST)-TROSY HSQC [35] spectra after heat shock (Figure 2). These studies demonstrated that the inherent conformational heterogeneities of G4s can be successfully investigated with NMR spectroscopy. However, the ^{13}C/^{15}N labeling techniques and the use of photolabile protecting groups for the DNA sample preparation are not widespread yet. More advances and commercialization of these techniques are required to study the dynamics of G4s with heteronuclear NMR.

Besides the typical G4s with three G-tetrads, G4s formed with four G-tetrads also have conformational heterogeneity. The structural polymorphism of d($G_4C_2)_3G_4$ was recently revealed [36]. The d($G_4C_2)_3G_4$ is the minimal element of the d($G_4C_2)_n(G_2C_4)_n$ repeats in the non-coding region of the C9orf72 gene, known for potentially causing the neurodegenerative diseases amyotrophic lateral sclerosis and frontotemporal dementia [37,38]. Two major antiparallel structures were stabilized with 8Br-dG modification at a specific position, and the structures were investigated. Conventional 2D NMR techniques were applied to collect all restraints for the structural calculation. 1D ^{1}H NMR spectroscopy was used to identify the population of each form of G4 structure based on the imino and aromatic proton assignments at various temperatures. Interestingly, the population of each conformation was affected by the annealing speed and the pH. Their interconversion was very slow, so the peaks from both conformations were observed in the same spectra. The solution structure contained a pseudo-C-quartet formed by the Cs in the lateral loops on top of the four G-tetrads. This research showed that pH can affect G4 conformation by changing the protonation state of loop residues.

As we showed above, most NMR studies of G4 DNAs routinely monitor imino protons to demonstrate conformational heterogeneity and slow exchange. One noteworthy study used the ^{1}H 1D NMR of the imino protons quantitatively. Varizhuk et al. investigated the polymorphism between G4 monomers and polymers by integrating techniques such as atomic force microscopy, optical and electrophoretic analysis, and molecular modeling with NMR spectroscopy [39]. 2D diffusion ordered spectroscopy was used to estimate the molecular weight from the diffusion coefficients, and the data suggested that parallel G4s without overhangs forms oligomers. Also, after integrating the imino proton region of the ^{1}H NMR spectra, the total numbers of imino protons were found to be equal to the expected numbers in the oligomeric state. The research we mention here shows that the conformational differences between G4 structures and the exchange between them can be effectively studied with NMR spectroscopy, and we expect that advances in DNA labeling technology, chemical modification, and NMR techniques will contribute to this field.

2.2. G4-Protein Interaction

Many proteins are involved in biological processes such as telomere maintenance, replication, and transcription regulation by interacting with the exceptionally stable G4. The biochemical properties of G4 helicases that unfold G4 structures by ATP hydrolysis and their potential as anticancer targets are well documented in recent reviews [40,41]. Nonenzymatic proteins such as hnRNPA1, hPOT1, and high mobility group B1 (HMGB1) have also been revealed as G4 interactors [42–45]. NMR spectroscopy has been applied to characterize the binding surfaces, binding affinities, and the exchange timescales between the free and bound forms of the protein-G4 complexes.

Amato et al. investigated telomeric G4 recognition by HMGB1, which is involved in telomere maintenance in the nucleus [46]. The peak intensity changes in the ^{1}H-^{15}N HSQC spectra upon titrating

unlabeled DNA into ^{15}N-labeled proteins were analyzed. Decreases in peak intensity upon ligand addition are usually considered as evidence of interaction in the intermediate regime of the NMR timescale (Figure 3A). In combination with other biophysical and computational techniques such as surface plasmon resonance and docking simulations, the G4 binding surfaces and binding affinities were revealed. The third RGG motif (RGG3) in the translocated in liposarcoma (TLS)/fused in sarcoma (FUS) protein forms a tertiary complex with telomeric G4 DNA and telomeric repeat-containing RNA (TERRA). Interactions between RGG3 of TLS/FUS and telomeric G4 DNA and TERRA were monitored with NMR spectroscopy [47]. The chemical shift perturbations (CSPs) of the imino protons of the G4 and the amide (^{1}H-^{15}N) cross-peaks of the proteins were analyzed. Because of the poor spectral dispersion and the overlap of signals from many Arg/Gly residues, only partial assignments of the amide signals from other than Arg/Gly were available. However, the binding surfaces for two G4 nucleic acids are distinguishable based on the different CSP patterns of specific residues. This study showed that the atomic resolution of NMR spectroscopy could provide valuable information about the interaction when the complex structure cannot be obtained.
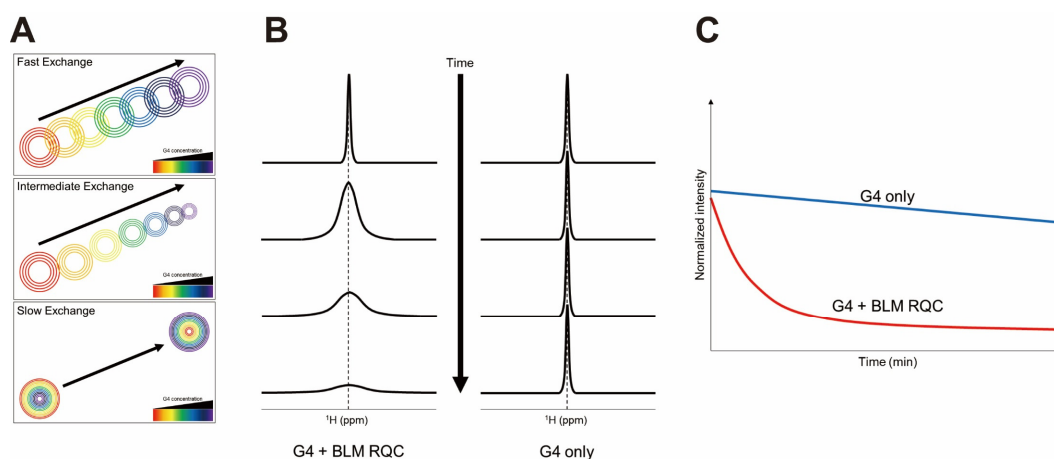

Figure 3. (**A**) CSP patterns with the G4-protein complexes in different exchange regimes on the NMR timescale. Schematic drawings of (**B**) a ^{1}H imino proton peak in the middle plane of the G4 in an H/D exchange experiment and (**C**) a decay profile of a specific imino proton.

Intrinsically disordered RGG motifs are found in several proteins. One of the RGG motifs in hnRNPA1 was shown to contribute to specific binding of telomeric G4 [48]. In this study, the CSP *vs* [DNA]/[Protein] ratio was used to determine the dissociation constants, which could not be determined with isothermal titration calorimetry because of the very weak binding affinity. By using various G4 structures as ligands, the authors showed that the RGG motif specifically recognizes the structured loop in the G4. The gradual changes in cross-peaks observed in this study indicate that the RGG motif and its G4 complex are in fast exchange in NMR timescale (Figure 3A). Also, the intensities of the imino protons of the G4 in 1D NMR decreased with increasing concentration of the RGG motif. It was consistent with the results of the G4 unfolding assay performed with circular dichroism (CD) spectroscopy.

The G4 unfolding mediated by helicases contains several steps that accompany structural rearrangements of both G4 and proteins [49,50]. The c-MYC G4 interaction with two human RecQ helicases (Werner syndrome protein (WRN) and Bloom syndrome protein (BLM)) was studied independently [51,52]. The RecQ C-terminal (RQC) domain of WRN was subjected to titration with non-G4 DNA or G4 DNA, and the residues which showed G4-specific responses were identified [51]. Interestingly, many amide peaks in the ^{1}H-^{15}N HSQC spectra disappeared upon addition of even small amounts of G4 DNA (< 0.05 molar equivalents). In this study, the G4-specific residues were not located in the duplex DNA binding surfaces identified by previous crystal structures [51,53].

In the case of BLM RQC, titration with up to 2 molar equivalents of DNA was performed, and CSPs could be observed, while only a few peaks disappeared upon addition of G4 [52]. Interestingly, the significantly perturbed residues were partially overlapped with the known duplex DNA binding surfaces [52,54]. Further investigation with Car-Purcell-Meiboom-Gill (CPMG) relaxation dispersion experiments showed that the BLM RQC-G4 interactions are in the intermediate regime on the NMR timescale. CPMG relaxation dispersion experiments have used to quantify micro − millisecond time scale dynamics of proteins by analyzing $R_{2,eff}$ on different CPMG frequencies [55]. The exchange rates, populations, and chemical shift differences between different states can be obtained. In this study, H/D exchange experiments were used to monitor G4 unfolding induced by BLM RQC. As expected, the imino protons in the middle plane were observed only after D_2O exchange, and the decay profiles were obtained per each guanine (Figure 3B,C). The results showed that the D_2O exchange rate is much faster in the presence of BLM RQC. This study exemplifies a quantitative way to evaluate G4 unfolding by proteins with NMR spectroscopy.

As we described above, most NMR studies of G4-protein interactions have used the amide cross-peaks of the protein and the imino protons of the G4 as the fingerprints. More probes, such as the aromatic ^{13}C-^1H cross-peaks of the G4, could complement the current tools. Also, more detailed dynamics investigations are expected to provide insights into G4 recognition by proteins.

2.3. G4-Ligand Interaction

G4 targeted ligands have recently emerged as a promising strategy for developing anticancer drugs. Because telomerase is highly expressed in many kinds of tumor cells, telomeric G4s have been considered as a potential target for ligands that bind to and stabilize the G4 for inhibition of telomerase [56,57]. G4s in oncogene promoters such as c-MYC, c-kit, and KRAS are also important in cancer biology. It is known that c-MYC transcription is upregulated in 80% of solid tumors, and it could be regulated by c-MYC targeted therapeutics [58,59]. There are several recent reviews of the design, synthesis, and therapeutic potential of G4 ligands [18,57,60].

1D ^1H NMR spectra of G4 imino protons have conventionally been used for monitoring G4-ligand interactions because they detect not only the binding but also more subtle structural conversions. A transition in the folding topology of Tel23 G4 (d[TAG$_3$(T$_2$AG$_3$)$_3$]) upon BMVC-8C3O (3,6-bis(1-methyl-4-vinylpyridiumiodide)-9-(1-methyl-piperidiniumiodide)-3,6,9-trioxaundecane) carbazole binding was monitored with time-resolved NMR spectroscopy [61]. Particularly, H/D exchange spectra showed that there was incomplete disruption of hydrogen bonding in the middle plane during the folding transition, indicating that the folding topology conversion is a slow process and not associated with the global unfolding events. Debnath et al. deduced that bis-triazolylcarbazole (BTC) interacts with and stabilizes a minor conformation of the c-MYC G4 by monitoring CSP of imino protons [62]. Together with single-molecule Förster resonance energy transfer (smFRET) data, the NMR data suggested that the G4 stabilization by BTC resulted from conformational selection and not an induced-fit process. The RNA polymerase I inhibitor BMH-21 was tested for c-MYC G4 (Pu22-T14/T23) binding [63]. The imino proton spectra of the G4 with increasing concentrations of the ligand clearly showed that the binding is an intermediate to slow exchange process, and the stoichiometry was also determined as 2:1 (drug: G4). Another anticancer agent, CX5461, was also titrated into c-MYC G4 (Pu22-T14/T23), and a similar change of the imino proton spectra was observed. The pi-pi stacking interaction between the G4 plane and small molecules is known as one of the most crucial binding determinants. However, the staking interaction is not quite specific because the planar base structures are all common regardless of different G4 structures.

There have been efforts to develop ligands that will bind specifically to each G4 structure. Thiazole peptides were developed and tested for specific c-MYC G4 binding [64]. These experiments also showed that the ligand binds to c-MYC's G4 in slow exchange on the NMR timescale. In this research, imino, anomeric, and aromatic signals of each nucleotide were monitored during the ligand titration, and CSPs were calculated. The data showed that both terminal regions are responsible for ligand

binding. The modified peptide (KR-12C) of human cathelicidin formed a complex with c-MYC G4 specifically, and the solution structure of the complex was determined with NMR spectroscopy [65]. The imino proton spectra of the c-MYC G4 with a different version of the peptide (KR-12A) showed that the signals from G4 were broadened and new signals from the hairpin Watson-Crick base-pairs appeared. This indicates that site-specific substitution affects the G4 binding and folding topology conversion of cathelicidin. The research discussed here used various biophysical techniques such as fluorescence spectroscopy, molecular docking simulations, and isothermal titration calorimetry together with NMR spectroscopy. To advance the design of more specific G4 binding ligands, these combined approaches are required. The dynamics of the G4-ligand complexes we have discussed are mostly in the slow regime in the NMR timescale. Combining with the structural heterogeneity of G4s themselves, these features warrant more detailed investigations.

3. I-Motif

The i-motif is one of the non-canonical DNA structures that can be built from strands complementary to those that form G4s. It has two parallel duplexes held together by hemiprotonated cytidine and protonated cytidine (C·C$^+$) base-pairs intercalated in an antiparallel orientation (Figure 4A) [66]. Because of the requirement of protonated bases, the formation of i-motifs is known to be more stable at acidic pH conditions [67]. The length of i-motifs is dependent on the number of C·C$^+$ base-pairs [68]. As there are large numbers of G4s in human promoters and telomeres [69], it is expected that i-motifs are also abundant in cells and that they, like G4s, are involved in the regulation of gene expression [70,71]. Previously, the biological relevance of the i-motif was suspected because of its poor stability at physiological pH and a lack of direct evidence that i-motifs actually existed in the intracellular environment. However, a recent study revealed the presence of an i-motif in vivo using NMR techniques [72]. Dzatko et al. found that the C-rich DNA sequence forms an i-motif in human HeLa cells, and the i-motif has better thermostability in vivo compared to in vitro, based on in-cell 1D ^{1}H NMR imino spectral data. Although there is a difference in the origin of the base-pair compared to G4s, the i-motif also has conformational heterogeneity depending on its topology [73] and oligomeric state [74]. Based on the position of the outermost C·C$^+$ base-pair, i-motif topologies can be classified into two types. When the C·C$^+$ base-pair is nearer to the 5′end of the DNA strand, it is called a 5′E topology. Conversely, if it is closer to the 3′end, it is a 3′E topology (Figure 4B) [75]. The number of DNA strands that make up the i-motif determines its oligomer state. I-motifs can be generated from one strand (intramolecular, monomer) or from more than one strand (multimer, dimeric or tetrameric) (Figure 4C).

There are reviews of the structural studies of the i-motif [75] and i-motif binding ligands [76]. As a complement to these, we summarize the NMR studies of the structural and dynamic properties of i-motifs.

3.1. I-Motif DNA

To form i-motifs, a C-rich DNA sequence must have multiple C·C$^+$ base-pairs. The C·C$^+$ base-pair is formed by hydrogen bonding in which the nitrogens in the 3-position of the two cytidines have a hemiprotonated N···H···N moiety [66]. Previously, it had been proposed that N···H···N bonding of the C·C$^+$ base-pair would follow the hypothesis of either 1) a symmetric single-well potential or 2) a delocalized proton oscillating between two wells with a double-well potential [77]. Lieblein et al. demonstrated which of the two hypotheses is correct through NMR spectroscopy [78]. The ^{1}H-^{15}N HMQC spectra showed that the ^{1}H chemical shifts of the two nitrogens forming the C·C$^+$ base-pair were identical, meaning that they share one proton. Also, it was found that the proton rapidly moves between two nitrogens through observation of the ^{1}J (NH) coupling constant. Thus, the C·C$^+$ base-pair, which is the basis of the i-motif, follows the pattern of hydrogen bonding characterized by a proton hopping between two nitrogens like hypothesis 2) (Figure 4A).
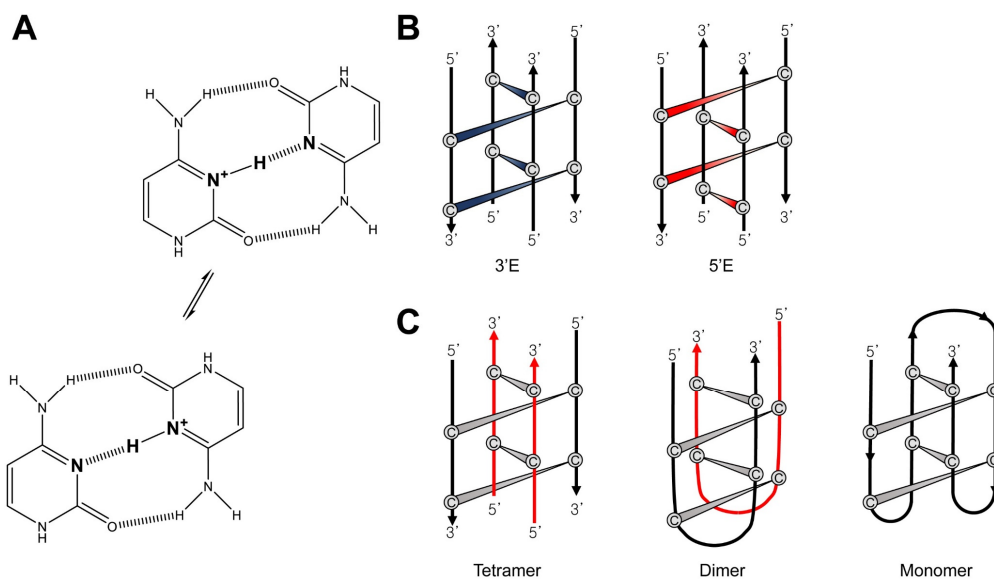
Figure 4. (**A**) Hemiprotonated C·C$^+$ base-pairs. (**B**) I-motif topologies (5'E and 3'E). (**C**) Oligomeric states of i-motifs.

The slow association rate and strong binding capacity of dimeric i-motif d(5mC$_2$TCACTC$_2$)$_2$ were characterized using NMR experiments [79]. Restraints for the structural calculation were obtained through conventional 2D NMR methods. At a slightly acidic pH (pH 6.8), d(5mC$_2$TCACTC$_2$)$_2$ had a dimeric topology. This was confirmed by the similarity of the imino proton spectra with that of the known d(5mC$_2$TCICTC$_2$)$_2$ [80]. Monitoring base-pair lifetimes through imino proton exchange experiments revealed that C·C$^+$ base-pairs contribute to the structural stability of the i-motif. In particular, it was found that the C·C$^+$ base-pair at the core of the i-motif has an extremely long lifetime. It is similar to the dimer topology lifetime, meaning that the innermost C·C$^+$ base-pair plays an important role in the maintenance of dimerization. In addition, analysis of dimer-monomer exchange kinetics revealed that the dimer stability has a pH dependency and the i-motif has an association rate four orders of magnitude slower than the formation rate of B-DNA duplexes.

Esmaili et al. investigated tetrameric structures of two versions of *Tetrahymena* telomeric DNA repeats, d(A$_2$C$_4$), and d(C$_4$A$_2$), by 2D NMR methods [81]. While the d(A$_2$C$_4$) tetramer had a single 5'E topology, it was found that the d(C$_4$A$_2$) tetramer shared two topologies. A series of TOCSY experiments was conducted to determine the ratio of the two topologies. It was found, by monitoring the intensities of the H6-H5 cross-peaks of the spectra, that the ratio of the 3'E topology increased compared to the 5'E topology as the temperature increased. However, the ratio of the two topologies did not depend on pH conditions or oligonucleotide concentrations. According to the base-pair opening kinetics analysis using proton exchange, the lifetime of the innermost C·C$^+$ base-pairs was significantly longer than that of the outer base-pairs. 1D NMR imino proton observation was conducted to identify whether the i-motif lost its tertiary structure due to the formation of G-C base-pairs when presented with G4 in solution. After d(A$_2$C$_4$)$_4$/d(C$_4$A$_2$)$_4$ + d(T$_2$G$_4$)$_4$ was stored for one month, broadening of the imino proton peaks of the i-motif and G4 did not occur. Rather, broadening of the adenine proton peaks and fast exchange rates of thymidine imino protons were observed, indicating a temporary interaction of terminal adenosines and thymidines.

The application of NMR to the vertebrate telomere sequence, d(C$_3$TA$_2$)$_3$C$_3$, enabled the study of conformational heterogeneities of the i-motif with pH [82]. d(C$_3$TA$_2$)$_3$C$_3$ was identified as having both 5'E and 3'E topologies through ^{1}H-^{15}N HMQC experiments. Using time-resolved NMR spectroscopy, i-motif folding and kinetic partitioning were observed with decreasing pH in proportion to time. Through monitoring the imino proton region by 1D NMR, it was found that a change in topology from 3'E to 5'E could occur depending on pH. In addition, fast exchange rates were shown for folding into

both topologies, but the exchange between the two topologies was very slow in the NMR time scale. This experiment showed that pH can influence the topology of the i-motif.

As mentioned above, heterogeneous structures and dynamics of the i-motif were routinely studied by monitoring chemical shift changes of imino protons under various conditions. It is expected that more non-canonical DNA structures and their properties will be revealed using NMR spectroscopy.

3.2. I-Motif Binding Ligands and Proteins

The discovery and design of i-motif binding ligands is not well developed due to the fact that its formation occurs in acidic pH in vitro [76]. Starting with TMPyP4 [83] and BisA [84], a few ligands that can bind to i-motifs have been described [76,85]. Among them, a study on ligands for the C-rich DNA sequence of the human B-cell lymphoma gene-2 (BCL2) promoter region is relatively well described [86]. Kendrick et al. discovered BCL2 C-rich sequence-specific binding ligands IMC-48 and IMC-76, which respectively induce formation of the i-motif and flexible hairpins, by FRET high-throughput screening assays (Figure 5A). After that, it was revealed by imino proton peak monitoring through 1D NMR that the topology heterogeneity of the BCL2 C-rich sequence changes with the addition of ligand. When IMC-48 was added, the imino proton peak corresponding to the i-motif was prominent (Figure 5B), but in the case of IMC-76 addition, the peaks of the flexible hairpin were observed (Figure 5C). Even in the state where one ligand is bound to DNA, a change of topology occurred by increasing the concentration of the other ligand. Actually, the shifting of the equilibrium in one direction by the ligands occurred in vivo, during their regulation of the expression of the *BCL2* gene.
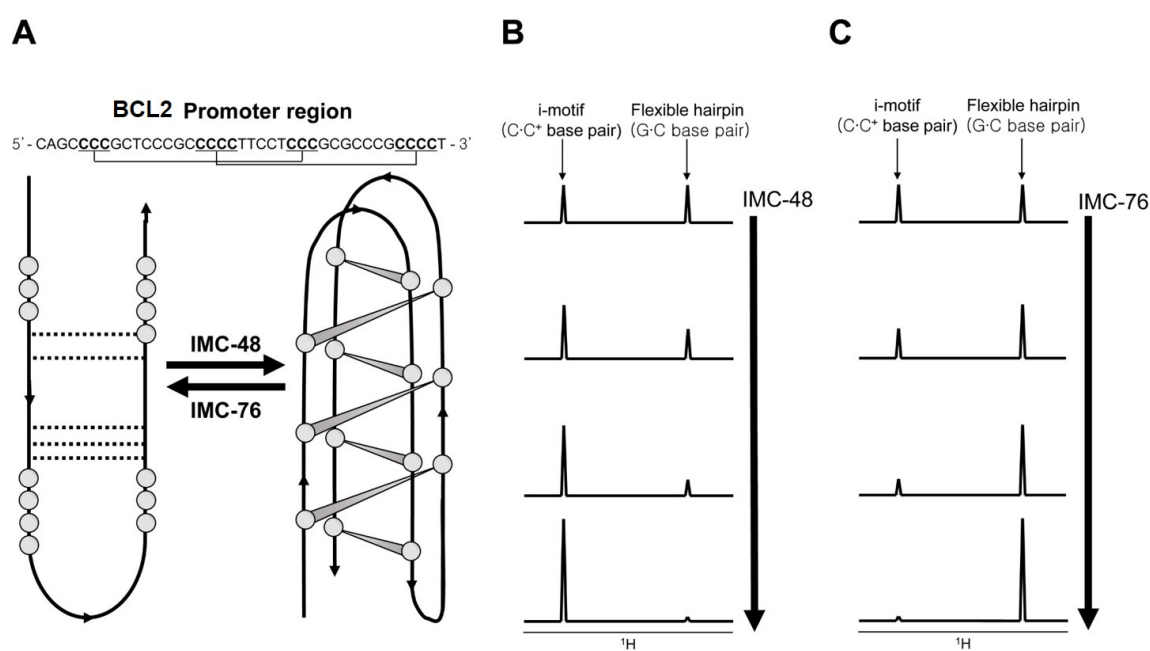

Figure 5. (**A**) The structural conversion from monomeric to dimeric BCL2 C-rich strand induced by ligands. Cytidines participating in C·C$^+$ base-pairs for i-motif formation are marked in bold. (**B,C**) Schematic drawing of ^{1}H imino proton peak of BCL2 C-rich strand changed by ligand titration.

The search for i-motif binding proteins has been very limited. In vitro biochemical assays revealed that the hnRNP K protein and the ASF/SF2 splicing factor bind to the telomeric C-rich strand, but no further structure-based studies have been conducted [87]. Currently, hnRNP LL is the only protein whose binding region for the i-motif has been identified [88]. The binding of the RRM1 and RRM2 domains (RRM12) to the P1 promoter of BCL2 contains a 39-mer C-rich sequence (Py39wt) that was evaluated using NMR experiments. ^{1}H-^{15}N HSQC spectra showed that the i-motif binding site of RRM12 is retained at physiological pH (pH 5.5–7.5). Interestingly, this binding was in the intermediate regime on the NMR timescale at low pH but changed to a slow regime when the pH increased. Further

^{31}P NMR and ^{1}H NMR experiments revealed that the conversion of the Py39wt i-motif to the duplex form due to the rise in pH causes a change in the RRM12 chemical exchange.

Structural and kinetic studies of i-motif binding ligands/proteins have been stagnant compared to studies on G4s. However, as it was revealed that the i-motif exists in vivo [54], it is expected that the research on the i-motif and its binders using NMR spectroscopy will be actively conducted in the future.

4. DNA Triplex

Major groove recognition of a homopurine and homopyrimidine DNA duplex by a third strand results in DNA triple helices (triplexes). Since the first discovery of a triplex in 1957, Dervan and coworkers showed the broad potential of triplex-forming oligonucleotides (TFOs) to inhibit sequence-specific DNA-protein interactions [89–91]. As described in Figure 6, the composition of the TFO, which uses Hoogsteen base-pairing, influences the orientation of a triplex. In the case of pyrimidine-rich (or purine-rich) oligonucleotides, the triplex forms a parallel structure (or antiparallel structure), so-called YRY (or RRY). Triplex DNA can be formed not only by the binding of a third strand to a DNA duplex, it can also be induced by homopurine-homopyrimidine mirror repeats pulled out of the upstream regions, so-called H-DNA [92,93]. The name of H-DNA originated from the fact that triplexes preferentially form under acidic conditions [6].
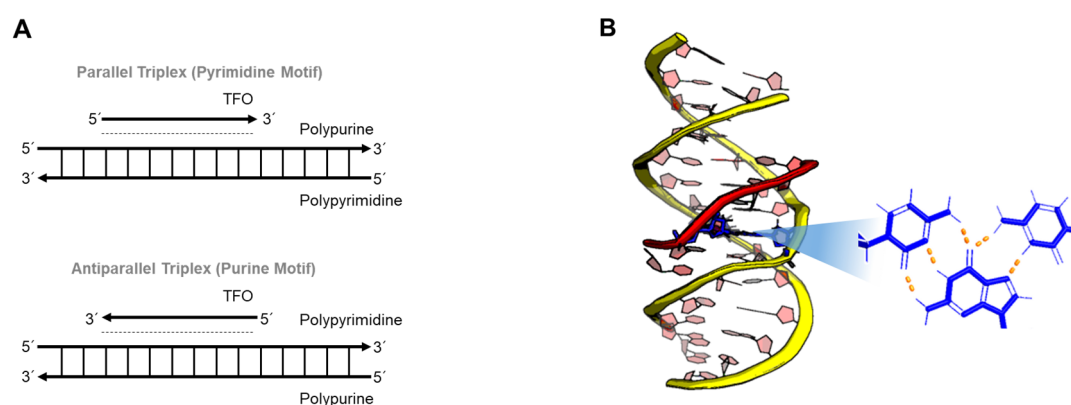

Figure 6. (**A**) Illustrations of parallel and antiparallel triple helices (TFO, triplex-forming oligonucleotide). (**B**) Side-view of a triple helix and base interactions at its midpoint (PDB id = 1BWG). The backbone of the TFO is shown in red. H-bonds between base-pairs are indicated by orange dashed lines.

Triplex structures have been an area of interest in biological research for several decades because of their biological roles in processes such as mutagenesis, lytic replication of the Epstein-Barr virus genome, and homologous recombination [94–96]. They also have potential applications in biotechnology and molecular medicine. To understand the biological roles of triplex DNA and the applications of TFOs, it is essential to explore triplex structure and dynamics, so that the molecular interactions between bases in solution are well defined.

The stability of DNA triplexes is strongly influenced by base sequence, length of the strands, and strand mismatches, as well as solution conditions, such as temperature, salt concentration, and pH [97–100]. How the third strand associates with or dissociates from a duplex has been studied using varying approaches, such as fluorescence spectroscopy and simulations [99–104]. For instance, base-stacking interactions in TAT triplex and AT duplex DNA were compared by fluorescence decay, and showed restricted motional dynamics in the TAT triplex compared to the AT duplex [105].

The exchange rate of protons using 1D NMR can be a useful technique to elucidate site-specific kinetics and dynamics of DNA triplexes. Russu and coworkers investigated the base-pair opening dynamics in DNA triplexes using imino proton exchange rates, and found that the stability of Hoogsteen base-pairs in a DNA triplex is comparable to that of Watson-Crick base-pairs in a DNA

duplex [106]. They also reported that a GTA triad has destabilizing effects on its neighboring triad, though it has similar stability to a TAT triad [107]. Later, the exchange rates of individual Watson-Crick and Hoogsteen imino protons were reported by the monitoring of water magnetization transfer and hydrogen/deuterium (H/D) exchange in the presence and absence of Mg^{2+} ions [108]. The H/D exchange measurement provided kinetic information on the slower structural opening reactions than kinetics determined by water magnetization transfer experiments. The results showed the energetic effects of ions at individual sites of DNA. They suggested that Mg^{2+} ions stabilize the closed conformations of base-pairs since Mg^{2+} ions are closely associated with the DNA. Recently, selective binding between triplex DNA and intercalators has been investigated, and the triplex DNA was shown to undergo covalent attachment based on CSPs in ^1H-^{15}N HSQC spectra and NOE data [109,110]. These studies elucidated sequence- and pH-dependent stability, showing a more stable TAT triplet at the 5′ junction under low pH, and slow exchange of chemical shifts between the coexisting species.

5. Z-DNA

One of the notable features of DNA is its ability to transition between right-handed B-DNA and left-handed Z-DNA, which has a zigzag backbone with syn- and anti- conformations [111–113]. The B–Z transition of DNA is known to happen in a sequence specific manner, occurring in regions enriched with purine-pyrimidine repeats, $(GC)_n$ [112]. It has been found that Z-DNA is present in vivo and plays important roles in biological functions, such as remodeling of chromatin structure, activation of transcription, and binding of proteins during viral infection [114–118]. To understand the kinetics and dynamics of Z-DNA, it has been extensively investigated using a broad range of methods, such as single-molecule techniques and NMR spectroscopy [119–123], while steady-state methods, such as X-ray crystallography, CD measurements, fluorescence spectroscopy, and 1D NMR results, have been used to determine the structural features of the B–Z transition [111,124–127]. In particular, hydrogen exchange measurements by NMR spectroscopy, which measure base-pair opening kinetics, have emerged as a promising technique to characterize the fluctuations of the inter- and intramolecular interactions of the bases and Z-DNA binding proteins (ZBPs) [33].

5.1. Z-DNA

Among the various non-canonical DNAs, Z-DNA has received attention due to its exceptional left-handed helical conformation, which shows a more elongated helical structure than B-DNA and different chirality [111–113]. The repeated units of $(CG)_n$ or $(GT)_n$ are the most common sequences forming Z-DNA, with the free energy cost for switching from B-form to Z-form ranging from 0.33 to 0.66 kcal/mol/bp [128,129]. The structural propensity to form Z-DNA within repeated sequences is as follows: d(CG) > d(TG/AC) > d(GGGG) > d(TATA). However, the formation of Z-DNA is also possible in sequences that lack pyrimidine-purine repeats or that contain AT base-pairs [130–133]. Based on numerous experimental and theoretical studies, it is widely known that the B–Z transition occurs under conditions of high ionic strength, negative supercoiling, and protein binding [134–137]. Tashiro et al. found that the Z-DNA conformation was favored at low temperature [138,139]. In addition, methylation of DNA, such as 5mCpG or 8mCpG methylation, greatly favors transitions from B-form to Z-form at physiological salt concentrations [120,140–142].

Based on the proton exchange data, the rate of base-pair opening of a Z-DNA duplex consisting of CG sequences was reported to be approximately five times smaller than B-DNA composed of same sequences, while B-DNA base-pairs are known to open on 1–10 microsecond timescales [31,143,144].

Figure 7A shows the 1D NMR imino spectra from a water magnetization transfer experiment for a 6 base-paire (6-bp) DNA duplex, $d(CG)_3$ (also called CG_6), with/without ZBPs, as described by Kang et al. [145]. The exchange rate k_{ex} for the CG_6 duplex is determined using the relative height of the imino resonances as a function of the delay time after water inversion. Figure 7B shows gradual

changes of the G2z resonance of d(CG)$_3$ and d(CG)$_3$-protein complexes. The relative peak intensities are fit to equation 1, which extracts k$_{ex}$ values based on the best fit (Figure 7C).

$$\frac{I(t)}{I_0} = 1 - 2\frac{k_{ex}}{(R_{1w} - R_{1a})}\left(e^{-R_{1a}t} - e^{-R_{1w}t}\right) \tag{1}$$
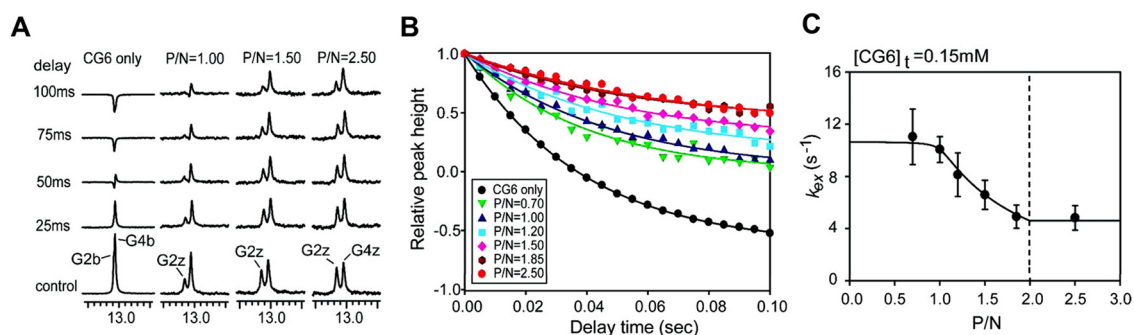

Figure 7. (**A**) One-dimensional imino proton spectra of the water magnetization transfer experiments. P, protein; N, nucleotide. (**B**) Relative peak intensities in the water magnetization transfer spectra for G2z imino resonances as a function of delay time. (**C**) Exchange rate constants of the G2z imino protons as a function of the hZα$_{ADAR1}$:d(CG)$_3$ molar ratio (P/N ratio). Reprinted with permission from Ref. [145]. Copyright 2009 American Chemical Society.

The exchange rates of the G2b and G4b imino protons of the d(CG)$_3$ duplex were determined as 21 and 12 s^{-1}, respectively. As the ratio of protein increased in the DNA-protein complexes, the imino protons of the Z-form are exchanged more slowly, indicating an increase in stability of the duplex. Dynamics and kinetic models of DNA-protein complexes will be discussed in the next section.

5.2. Z-DNA Complexes with ZBPs

ZBPs and antibodies can selectively bind to Z-DNA (or Z-RNA), inducing conformational transitions. Crucial biological functions of Z-DNA are induced by association with ZBPs, examples being the RNA editing enzyme ADAR1 [137,146,147], the DNA-dependent activator of IFN-regulatory factor (DAI, also known as DLM-1 or ZBP1) [148,149], the poxviral E3L protein [118,150], and RNA-dependent protein kinase (PKZ) [151]. These ZBP and DNA complexes adopt well-defined structures that have been elucidated by crystallographic and NMR studies [152]. Interestingly, despite the wide variety of biological functions, there are structural analogies between ZBPs. In each case, the residues in the α3 helix and β-hairpin (β2-loop-β3) of ZBPs mediated the intermolecular interactions for DNA recognition. Recent studies of the structures and biological functions of Z-DNA-ZBP complexes provide insights into DNA recognition and B–Z transition induced by ZBPs.

Recent NMR studies suggested an active mechanism of B–Z transition, which support two independent roles of ZBPs as follows [123,145,153–155]: (i) one ZBP molecule binds to 6-bp B-DNA, d(CG)$_3$, and facilitates the conversion from B-DNA to Z-DNA; (ii) the second ZBP forms the complete DNA–(ZBP)$_2$ complex. The Zα domain of human ADAR1 (hZα$_{ADAR1}$), which stabilizes the Z-DNA conformation, is able to bind to a 6-bp Z-DNA duplex with not only CG-repeat-rich sequences but also non-CG-repeat sequences [156]. The results of hydrogen exchange experiments of imino protons for free d(CG)$_3$ and hZα$_{ADAR1}$–d(CG)$_3$ complexes elucidated the overall slower exchange rates of imino proton in the Z-form versus the B-form [145]. A kinetic model demonstrated that active-mono B–Z conversion bridged the equilibrium between B-form complex (BP) and Z-form complex (ZP). Recently, Lee et al. carried out the structural dynamics analysis of hZα$_{ADAR1}$–d(CG)$_3$ complexes based on the global fitting methods of relaxation dispersion combined with CSP [155]. ^{15}N CPMG relaxation dispersion results showed that hZα$_{ADAR1}$ with Z-DNA undergoes a pseudo-three-state conformational exchange, which includes two independent B–Z transitions for the free and bound

states. The Zα domain of yatapoxvirus E3L (yabZα$_{E3L}$) exhibits efficient B–Z conformational changes of d(CG)$_3$ like the hZα$_{ADAR1}$–d(CG)$_3$ complex, while the Zβ domain of human DAI (hZβ$_{DAI}$) –d(CG)$_3$ complex undergoes lower-efficiency B–Z transitions [157].

Using the same methods as above, including water magnetization transfer, CSPs in ^{1}H-^{15}N HSQC spectra, and CPMG relaxation dispersion experiments, the B–Z transition mechanism and quantitative energy landscape of the Zα domain of PKZ from *Carassius auratus* (caZα$_{PKZ}$) have been investigated [154]. The caZα$_{PKZ}$ is able to transform d(CG)$_3$ to a Z-DNA conformation with lower activity rates than hZα$_{ADAR1}$ and yabZα$_{E3L}$. The caZα$_{PKZ}$–DNA complex, which has a Z-DNA conformation, induced significant CSPs for the α3, β1- β2, and β-hairpin regions at different salt concentrations, due to the reduced B-Z transition activity at high salt (see Figure 8). Taken together, despite the structural similarity of varying types of ZBPs, the complexes between different ZBPs and DNA have different kinetic and dynamic effects on B-Z transitions of DNA duplexes.
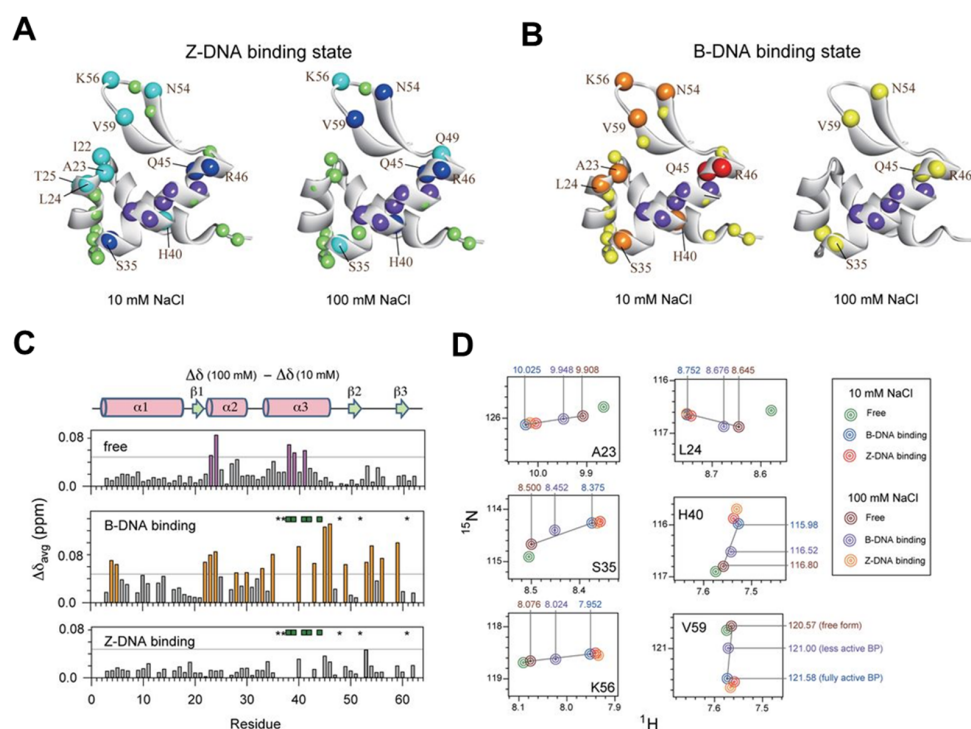

Figure 8. (**A,B**) DNA binding patterns of (A) Z-DNA and (B) B-DNA binding at 10 (left) or 100 mM NaCl (right). Chemical shifts are illustrated by different colors: red or blue, >0.18 ppm; orange or cyan, 0.12–0.18 ppm; and yellow or pale green, 0.08–0.12 ppm.The purple spheres indicate sites with disappeared or very weak cross-peaks during titration. (**C**) The average chemical shift differences between [NaCl] of 10 and 100 mM with free caZα$_{PKZ}$ (upper) and caZα$_{PKZ}$ – B-DNA (middle) and Z-DNA (lower) complex (**D**) The calculated ^{1}H-^{15}N HSQC cross-peaks of caZα$_{PKZ}$ – B-DNA or Z-DNA complexes in 10 or 100 mM NaCl solutions.

During B–Z transition, the most notable change is the shift in chirality between a right-handed and left-handed structure. Thus, as a common method, monitoring of the B–Z transition has been determined by inversion of the signs of CD spectra [124,158–160]. Recently, using a time course of signals recorded at 255 nm, the B–Z transition rate of the caZα$_{PKZ}$–DNA complex has been reported [161]. In this study, more significant changes of the time traces for caZα$_{PKZ}$–DNA than hZα$_{ADAR1}$–DNA suggested the crucial role of charge-charge interactions in the B-Z transition. Since fluorescence is strongly sensitive to the solvation environment, structural torsions, and the distance between fluorescence acceptor/donor pairs, fluorescence spectroscopy is one of the most useful methods for monitoring the change of DNA and DNA complexes [120,121,123–125]. Lee et al. investigated supercoiling-induced B–Z transitions at the single-molecule level using smFRET combined with

magnetic tweezers, and suggested that torsion plays an important role in the B-Z transition [122]. We anticipate that these spectroscopic methods will be a compatible partner with NMR to understand the biological functions induced by the dynamics of ZBP–Z-DNA complexes.

6. Conclusions

The non-canonical structures of DNA are essential for their diverse functions during various biological processes. These non-canonical structures can undergo conformational exchange among multiple structural states. Their dynamics data illustrates that each conformational state can play important roles in folding, stability, and biological function. Currently, NMR can be one of the most versatile tools for studying DNA dynamics and DNA-ligand interactions at the atomic level. The NMR techniques that allow monitoring of different time scale motions have been frequently employed. This review provides insight into how the dynamic features of the non-canonical structures revealed by NMR spectroscopy comprehensively can be understood. Further dynamics studies are expected to give a more detailed view of the dynamic landscape of the non-canonical DNA.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

G4	G-quadruplex
BEST	Band-selective Excitation Short-Transient
HMGB1	High Mobility Group B1
RGG3	The third RGG motif
TLS	Translocated in liposarcoma
FUS	Fused in sarcoma
TERRA	Telomeric repeat-containing RNA
CSP	The chemical shift perturbation
CD	Circular dichroism
WRN	Werner syndrome protein
BLM	Bloom syndrome protein
RQC	RecQ C-terminal
CPMG	Car-Purcell-Meiboom-Gill
BMVC-8C3O	3,6-bis(1-methyl-4-vinylpyridiumiodide)-9-(1-(1-methyl-piperidiniumiodide)-3,6,9-trioxaundecane
BTC	Bis-triazolylcarbazole
smFRET	single-molecule Förster resonance energy transfer
C·C$^+$	Cytidine and protonated cytidine base-pair
BCL2	Human B-cell lymphoma gene-2
RRM12	RRM1 and RRM2 domains
TFO	Triplex-forming oligonucleotide
ZBPs	Z-DNA binding proteins
DAI	DNA-dependent activator of IFN-regulatory factor
PKZ	ZBP-containing protein kinase
hZα_{ADAR1}	Zα domain of human ADAR1
yabZα_{E3L}	Zα domain of yatapoxvirus E3L
caZα_{PKZ}	Zα domain of PKZ from *Carassius auratus*

References

1. Gallie, D.R. A tale of two termini: A functional interaction between the termini of an mRNA is a prerequisite for efficient translation initiation. *Gene* **1998**, *216*, 1–11. [CrossRef]
2. Yagil, G. Paranemic structures of DNA and their role in DNA unwinding. *Crit. Rev. Biochem. Mol. Biol.* **1991**, *26*, 475–559. [CrossRef] [PubMed]
3. Blackburn, E.H. Switching and signaling at the telomere. *Cell* **2001**, *106*, 661–673. [CrossRef]
4. Aubert, G.; Lansdorp, P.M. Telomeres and aging. *Physiol. Rev.* **2008**, *88*, 557–579. [CrossRef] [PubMed]
5. Lacroix, L.; Mergny, J.-L.; Leroy, J.-L.; Hélène, C. Inability of RNA to form the i-motif: implications for triplex formation. *Biochemistry* **1996**, *35*, 8715–8722. [CrossRef] [PubMed]
6. Choi, J.; Majima, T. Conformational changes of non-B DNA. *Chem. Soc. Rev.* **2011**, *40*, 5893–5909. [CrossRef]
7. Wang, G.; Vasquez, K.M. Impact of alternative DNA structures on DNA damage, DNA repair, and genetic instability. *DNA Repair* **2014**, *19*, 143–151. [CrossRef]
8. Halder, S.; Bhattacharyya, D. RNA structure and dynamics: A base pairing perspective. *Prog. Biophys. Mol. Biol.* **2013**, *113*, 264–283. [CrossRef]
9. Heus, H.A.; Hilbers, C.W. Structures of non-canonical tandem base pairs in RNA helices: Review. *Nucleosides Nucleotides Nucleic Acids* **2003**, *22*, 559–571. [CrossRef]
10. Uhlenbeck, O.C.; Pardi, A.; Feigon, J. RNA structure comes of age. *Cell* **1997**, *90*, 833–840. [CrossRef]
11. Wöhnert, J.; Dingley, A.J.; Stoldt, M.; Görlach, M.; Grzesiek, S.; Brown, L.R. Direct identification of NH...N hydrogen bonds in non-canonical base pairs of RNA by NMR spectroscopy. *Nucleic Acids Res.* **1999**, *27*, 3104–3110.
12. Rachwal, P.A.; Brown, T.; Fox, K.R. Effect of G-tract length on the topology and stability of intramolecular DNA quadruplexes. *Biochemistry* **2007**, *46*, 3036–3044. [CrossRef] [PubMed]
13. Davis, J.T. G-quartets 40 years later: From 5'-GMP to molecular biology and supramolecular chemistry. *Angew. Chem. Int. Ed.* **2004**, *43*, 668–698. [CrossRef]
14. Wright, E.P.; Huppert, J.L.; Waller, Z.A.E. Identification of multiple genomic DNA sequences which form i-motif structures at neutral pH. *Nucleic Acids Res.* **2017**, *45*, 2951–2959. [CrossRef] [PubMed]
15. Fleming, A.M.; Ding, Y.; Rogers, R.A.; Zhu, J.; Zhu, J.; Burton, A.D.; Carlisle, C.B.; Burrows, C.J. 4n−1 is a "Sweet Spot" in DNA i-motif folding of 2'-deoxycytidine homopolymers. *J. Am. Chem. Soc.* **2017**, *139*, 4682–4689. [CrossRef] [PubMed]
16. Haider, S.; Parkinson, G.N.; Neidle, S. Crystal structure of the potassium form of an Oxytricha nova G-quadruplex. *J. Mol. Biol.* **2002**, *320*, 189–200. [CrossRef]
17. Smirnov, I.; Shafer, R.H. Effect of loop sequence and size on DNA aptamer stability. *Biochemistry* **2000**, *39*, 1462–1468. [CrossRef]
18. Asamitsu, S.; Obata, S.; Yu, Z.; Bando, T.; Sugiyama, H. Recent progress of targeted G-quadruplex-preferred ligands toward cancer therapy. *Molecules* **2019**, *24*, 429. [CrossRef]
19. De Magis, A.; Manzo, S.G.; Russo, M.; Marinello, J.; Morigi, R.; Sordet, O.; Capranico, G. DNA damage and genome instability by G-quadruplex ligands are mediated by R loops in human cancer cells. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 816. [CrossRef]
20. Spiegel, J.; Adhikari, S.; Balasubramanian, S. The structure and function of DNA G-quadruplexes. *Trends Chem.* **2020**, *2*, 123–136. [CrossRef]
21. Bessi, I.; Jonker, H.R.A.; Richter, C.; Schwalbe, H. Involvement of long-lived intermediate states in the complex folding pathway of the human telomeric G-quadruplex. *Angew. Chem. Int. Ed.* **2015**, *54*, 8444–8448. [CrossRef] [PubMed]
22. Harkness, R.W.V.; Mittermaier, A.K. G-register exchange dynamics in guanine quadruplexes. *Nucleic Acids Res.* **2016**, *44*, 3481–3494. [CrossRef] [PubMed]
23. Do, N.Q.; Phan, A.T. Monomer–dimer equilibrium for the 5'–5' stacking of propeller-type parallel-stranded G-quadruplexes: NMR structural study. *Chem. Eur. J.* **2012**, *18*, 14752–14759. [CrossRef]
24. Seenisamy, J.; Rezler, E.M.; Powell, T.J.; Tye, D.; Gokhale, V.; Joshi, C.S.; Siddiqui-Jain, A.; Hurley, L.H. The dynamic character of the G-quadruplex element in the c-MYC promoter and modification by TMPyP4. *J. Am. Chem. Soc.* **2004**, *126*, 8702–8709. [CrossRef] [PubMed]
25. Adrian, M.; Heddi, B.; Phan, A.T. NMR spectroscopy of G-quadruplexes. *Methods* **2012**, *57*, 11–24. [CrossRef] [PubMed]

26. Harkness, R.W.; Mittermaier, A.K. G-quadruplex dynamics. *Biochim. Biophys. Acta - Proteins Proteom* **2017**, *1865*, 1544–1554. [CrossRef]

27. Nußbaumer, F.; Juen, M.A.; Gasser, C.; Kremser, J.; Müller, T.; Tollinger, M.; Kreutz, C. Synthesis and incorporation of 13C-labeled DNA building blocks to probe structural dynamics of DNA by NMR. *Nucleic Acids Res.* **2017**, *45*, 9178–9192. [CrossRef]

28. Guéron, M.; Kochoyan, M.; Leroy, J.-L. A single mode of DNA base-pair opening drives imino proton exchange. *Nature* **1987**, *328*, 89–92. [CrossRef]

29. Guéron, M.; Leroy, J.L. Studies of base pair kinetics by NMR measurement of proton exchange. *Methods Enzymol.* **1995**, *261*, 383–413.

30. Russu, I.M. Probing site-specific energetics in proteins and nucleic acids by hydrogen exchange and nuclear magnetic resonance spectroscopy. *Methods Enzymol.* **2004**, *379*, 152–175.

31. Krueger, A.; Protozanova, E.; Frank-Kamenetskii, M.D. Sequence-dependent base pair opening in DNA double helix. *Biophys. J.* **2006**, *90*, 3091–3099. [CrossRef] [PubMed]

32. Nikolova, E.N.; Zhou, H.; Gottardo, F.L.; Alvey, H.S.; Kimsey, I.J.; Al-Hashimi, H.M. A historical account of Hoogsteen base-pairs in duplex DNA. *Biopolymers* **2013**, *99*, 955–968. [CrossRef] [PubMed]

33. Choi, S.-R.; Kim, N.-H.; Jin, H.-S.; Seo, Y.-J.; Lee, J.; Lee, J.-H. Base-pair opening dynamics of nucleic acids in relation to their biological function. *Comput. Struct. Biotechnol. J.* **2019**, *17*, 797–804. [CrossRef] [PubMed]

34. Grün, J.T.; Hennecker, C.; Klötzner, D.-P.; Harkness, R.W.; Bessi, I.; Heckel, A.; Mittermaier, A.K.; Schwalbe, H. Conformational dynamics of strand register shifts in DNA G-quadruplexes. *J. Am. Chem. Soc.* **2020**, *142*, 264–273. [CrossRef] [PubMed]

35. Lescop, E.; Kern, T.; Brutscher, B. Guidelines for the use of band-selective radiofrequency pulses in hetero-nuclear NMR: Example of longitudinal-relaxation-enhanced BEST-type ^{1}H–^{15}N correlation experiments. *J. Magn. Reson.* **2010**, *203*, 190–198. [CrossRef]

36. Brcic, J.; Plavec, J. NMR structure of a G-quadruplex formed by four d(G_4C_2) repeats: Insights into structural polymorphism. *Nucleic Acids Res.* **2018**, *46*, 11605–11617.

37. DeJesus-Hernandez, M.; Mackenzie, I.R.; Boeve, B.F.; Boxer, A.L.; Baker, M.; Rutherford, N.J.; Nicholson, A.M.; Finch, N.A.; Flynn, H.; Adamson, J.; et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* **2011**, *72*, 245–256. [CrossRef]

38. Renton, A.E.; Majounie, E.; Waite, A.; Simón-Sánchez, J.; Rollinson, S.; Gibbs, J.R.; Schymick, J.C.; Laaksovirta, H.; Van Swieten, J.C.; Myllykangas, L.; et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **2011**, *72*, 257–268. [CrossRef]

39. Varizhuk, A.M.; Protopopova, A.D.; Tsvetkov, V.B.; Barinov, N.A.; Podgorsky, V.V.; Tankevich, M.V.; Vlasenok, M.A.; Severov, V.V.; Smirnov, I.P.; Dubrovin, E.V.; et al. Polymorphism of G4 associates: From stacks to wires via interlocks. *Nucleic Acids Res.* **2018**, *46*, 8978–8992. [CrossRef]

40. Mendoza, O.; Bourdoncle, A.; Boulé, J.-B.; Brosh, R.M., Jr.; Mergny, J.-L. G-quadruplexes and helicases. *Nucleic Acids Res.* **2016**, *44*, 1989–2006. [CrossRef]

41. Sauer, M.; Paeschke, K. G-quadruplex unwinding helicases and their function in vivo. *Biochem. Soc. Trans.* **2017**, *45*, 1173–1182. [CrossRef] [PubMed]

42. Brázda, V.; Hároníková, L.; Liao, J.C.C.; Fojta, M. DNA and RNA quadruplex-binding proteins. *Int. J. Mol. Sci* **2014**, *15*, 17493–17517. [CrossRef] [PubMed]

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

44. Hwang, H.; Buncher, N.; Opresko, P.L.; Myong, S. POT1-TPP1 regulates telomeric overhang structural dynamics. *Structure* **2012**, *20*, 1872–1880. [CrossRef] [PubMed]

45. Amato, J.; Madanayake, T.W.; Iaccarino, N.; Novellino, E.; Randazzo, A.; Hurley, L.H.; Pagano, B. HMGB1 binds to the KRAS promoter G-quadruplex: A new player in oncogene transcriptional regulation? *Chem. Commun.* **2018**, *54*, 9442–9445. [CrossRef] [PubMed]

46. Amato, J.; Cerofolini, L.; Brancaccio, D.; Giuntini, S.; Iaccarino, N.; Zizza, P.; Iachettini, S.; Biroccio, A.; Novellino, E.; Rosato, A.; et al. Insights into telomeric G-quadruplex DNA recognition by HMGB1 protein. *Nucleic Acids Res.* **2019**, *47*, 9950–9966. [CrossRef]

47. Kondo, K.; Mashima, T.; Oyoshi, T.; Yagi, R.; Kurokawa, R.; Kobayashi, N.; Nagata, T.; Katahira, M. Plastic roles of phenylalanine and tyrosine residues of TLS/FUS in complex formation with the G-quadruplexes of telomeric DNA and TERRA. *Sci. Rep.* **2018**, *8*, 2864. [CrossRef]

48. Ghosh, M.; Singh, M. RGG-box in hnRNPA1 specifically recognizes the telomere G-quadruplex DNA and enhances the G-quadruplex unfolding ability of UP1 domain. *Nucleic Acids Res.* **2018**, *46*, 10246–10261. [CrossRef]

49. Chatterjee, S.; Zagelbaum, J.; Savitsky, P.; Sturzenegger, A.; Huttner, D.; Janscak, P.; Hickson, I.D.; Gileadi, O.; Rothenberg, E. Mechanistic insight into the interaction of BLM helicase with intra-strand G-quadruplex structures. *Nat. Commun.* **2014**, *5*, 5556. [CrossRef]

50. Wu, W.-Q.; Hou, X.-M.; Li, M.; Dou, S.-X.; Xi, X.-G. BLM unfolds G-quadruplexes in different structural environments through different mechanisms. *Nucleic Acids Res.* **2015**, *43*, 4614–4626. [CrossRef]

51. Ketkar, A.; Voehler, M.; Mukiza, T.; Eoff, R.L. Residues in the RecQ C-terminal domain of the human Werner syndrome helicase are involved in unwinding G-quadruplex DNA. *J. Biol. Chem.* **2017**, *292*, 3154–3163. [CrossRef] [PubMed]

52. Lee, S.; Lee, A.-R.; Ryu, K.-S.; Lee, J.-H.; Park, C.-J. NMR investigation of the interaction between the RecQ C-terminal domain of human Bloom syndrome protein and G-quadruplex DNA from the human c-Myc promoter. *J. Mol. Biol.* **2019**, *431*, 794–806. [CrossRef] [PubMed]

53. Kitano, K.; Kim, S.-Y.; Hakoshima, T. Structural basis for DNA strand separation by the unconventional winged-helix domain of RecQ helicase WRN. *Structure* **2010**, *18*, 177–187. [CrossRef] [PubMed]

54. Newman, J.A.; Savitsky, P.; Allerston, C.K.; Bizard, A.H.; Özer, Ö.; Sarlós, K.; Liu, Y.; Pardon, E.; Steyaert, J.; Hickson, I.D.; et al. Crystal structure of the Bloom's syndrome helicase indicates a role for the HRDC domain in conformational changes. *Nucleic Acids Res.* **2015**, *43*, 5221–5235. [CrossRef]

55. Palmer, A.G.; Kroenke, C.D.; Loria, J.P. Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods Enzymol.* **2001**, *339*, 204–238.

56. Neidle, S. Human telomeric G-quadruplex: The current status of telomeric G-quadruplexes as therapeutic targets in human cancer. *FEBS J.* **2010**, *277*, 1118–1125. [CrossRef]

57. Sun, Z.-Y.; Wang, X.-N.; Cheng, S.-Q.; Su, X.-X.; Ou, T.-M. Developing novel G-quadruplex ligands: From interaction with nucleic acids to interfering with nucleic acid–protein interaction. *Molecules* **2019**, *24*, 396. [CrossRef]

58. Siddiqui-Jain, A.; Grand, C.L.; Bearss, D.J.; Hurley, L.H. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11593–11598. [CrossRef]

59. Chen, B.-J.; Wu, Y.-L.; Tanaka, Y.; Zhang, W. Small molecules targeting c-Myc oncogene: Promising anti-cancer therapeutics. *Int. J. Biol. Sci.* **2014**, *10*, 1084–1096. [CrossRef]

60. Duarte, A.R.; Cadoni, E.; Ressurreição, A.S.; Moreira, R.; Paulo, A. Design of modular G-quadruplex ligands. *Chem. Med. Chem.* **2018**, *13*, 869–893. [CrossRef]

61. Wang, Z.-F.; Li, M.-H.; Chen, W.-W.; Hsu, S.-T.D.; Chang, T.-C. A novel transition pathway of ligand-induced topological conversion from hybrid forms to parallel forms of human telomeric G-quadruplexes. *Nucleic Acids Res.* **2016**, *44*, 3958–3968. [CrossRef] [PubMed]

62. Debnath, M.; Ghosh, S.; Panda, D.; Bessi, I.; Schwalbe, H.; Bhattacharyya, K.; Dash, J. Small molecule regulated dynamic structural changes of human G-quadruplexes. *Chem. Sci.* **2016**, *7*, 3279–3285. [CrossRef] [PubMed]

63. Musso, L.; Mazzini, S.; Rossini, A.; Castagnoli, L.; Scaglioni, L.; Artali, R.; Di Nicola, M.; Zunino, F.; Dallavalle, S. C-MYC G-quadruplex binding by the RNA polymerase I inhibitor BMH-21 and analogues revealed by a combined NMR and biochemical Approach. *BBA-GEN SUBJECTS.* **2018**, *1862*, 615–629. [CrossRef]

64. Dutta, D.; Debnath, M.; Müller, D.; Paul, R.; Das, T.; Bessi, I.; Schwalbe, H.; Dash, J. Cell penetrating thiazole peptides inhibit c-MYC expression via site-specific targeting of c-MYC G-quadruplex. *Nucleic Acids Res.* **2018**, *46*, 5355–5365. [CrossRef] [PubMed]

65. Sengupta, P.; Banerjee, N.; Roychowdhury, T.; Dutta, A.; Chattopadhyay, S.; Chatterjee, S. Site-specific amino acid substitution in dodecameric peptides determines the stability and unfolding of c-MYC quadruplex promoting apoptosis in cancer cells. *Nucleic Acids Res.* **2018**, *46*, 9932–9950. [CrossRef] [PubMed]

66. Gehring, K.; Leroy, J.-L.; Guéron, M. A tetrameric DNA structure with protonated cytosine-cytosine base pairs. *Nature.* **1993**, *363*, 561–565. [CrossRef]

67. Choi, J.; Kim, S.; Tachikawa, T.; Fujitsuka, M.; Majima, T. pH-induced intramolecular folding dynamics of i-motif DNA. *J. Am. Chem. Soc.* **2011**, *133*, 16146–16153. [CrossRef]

68. Mergny, J.-L.; Lacroix, L.; Han, X.; Leroy, J.-L.; Helene, C. Intramolecular folding of pyrimidine oligodeoxynucleotides into an i-DNA motif. *J. Am. Chem. Soc.* **1995**, *117*, 8887–8898. [CrossRef]

69. Chambers, V.S.; Marsico, G.; Boutell, J.M.; Di Antonio, M.; Smith, G.P.; Balasubramanian, S. High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol.* **2015**, *33*, 877–881. [CrossRef]

70. Sun, D.; Hurley, L.H. The Importance of Negative Superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: Implications for drug targeting and control of gene expression. *J. Med. Chem.* **2009**, *52*, 2863–2874. [CrossRef]

71. Brooks, T.A.; Kendrick, S.; Hurley, L. Making sense of G-quadruplex and i-motif functions in oncogene promoters. *FEBS J.* **2010**, *277*, 3459–3469. [CrossRef] [PubMed]

72. Dzatko, S.; Krafcikova, M.; Hänsel-Hertsch, R.; Fessl, T.; Fiala, R.; Loja, T.; Krafcik, D.; Mergny, J.-L.; Foldynova-Trantirkova, S.; Trantirek, L. Evaluation of the stability of DNA i-motifs in the nuclei of living mammalian cells. *Angew Chem. Int. Ed.* **2018**, *57*, 2165–2169. [CrossRef] [PubMed]

73. Phan, A.T.; Leroy, J.L. Intramolecular i-motif structures of telomeric DNA. *J. Biomol. Struct. Dyn.* **2000**, *17*, 245–251. [CrossRef] [PubMed]

74. Dong, Y.; Yang, Z.; Liu, D. DNA nanotechnology based on i-motif structures. *Acc. Chem. Res.* **2014**, *47*, 1853–1860. [CrossRef]

75. Abou Assi, H.; Garavís, M.; González, C.; Damha, M.J. i-Motif DNA: Structural features and significance to cell biology. *Nucleic Acids Res.* **2018**, *46*, 8038–8056. [CrossRef]

76. Day, H.A.; Pavlou, P.; Waller, Z.A.E. i-Motif DNA: Structure, stability and targeting with ligands. *Bioorg. Med. Chem.* **2014**, *22*, 4407–4418. [CrossRef]

77. Schah-Mohammedi, P.; Shenderovich, I.G.; Detering, C.; Limbach, H.-H.; Tolstoy, P.M.; Smirnov, S.N.; Denisov, G.S.; Golubev, N.S. Hydrogen/deuterium-isotope effects on NMR chemical shifts and symmetry of homoconjugated hydrogen-bonded ions in polar solution. *J. Am. Chem. Soc.* **2000**, *122*, 12878–12879. [CrossRef]

78. Lieblein, A.L.; Krämer, M.; Dreuw, A.; Fürtig, B.; Schwalbe, H. The nature of hydrogen bonds in cytidine···H+···cytidine DNA base pairs. *Angew. Chem. Int. Ed.* **2012**, *51*, 4067–4070. [CrossRef]

79. Canalia, M.; Leroy, J.L. Structure, internal motions and association–dissociation kinetics of the i-motif dimer of d(5mCCTCACTCC). *Nucleic Acids Res.* **2005**, *33*, 5471–5481. [CrossRef]

80. Leroy, J.L.; Gehring, K.; Kettani, A.; Gueron, M. Acid multimers of oligodeoxycytidine strands: Stoichiometry, base-pair characterization, and proton exchange properties. *Biochemistry* **1993**, *32*, 6019–6031. [CrossRef]

81. Esmaili, N.; Leroy, J.L. i-motif solution structure and dynamics of the d(AACCCC) and d(CCCCAA) tetrahymena telomeric repeats. *Nucleic Acids Res.* **2005**, *33*, 213–224. [CrossRef] [PubMed]

82. Lieblein, A.L.; Buck, J.; Schlepckow, K.; Fürtig, B.; Schwalbe, H. Time-resolved NMR spectroscopic studies of DNA i-motif folding reveal kinetic partitioning. *Angew. Chem. Int. Ed.* **2012**, *51*, 250–253. [CrossRef] [PubMed]

83. Fedoroff, O.Y.; Rangan, A.; Chemeris, V.V.; Hurley, L.H. Cationic porphyrins promote the formation of i-motif DNA and bind peripherally by a nonintercalative mechanism. *Biochemistry* **2000**, *39*, 15083–15090. [CrossRef] [PubMed]

84. Alberti, P.; Ren, J.; Teulade-Fichou, M.P.; Guittat, L.; Riou, J.-F.; Chaires, J.B.; Hélène, C.; Vigneron, J.-P.; Lehn, J.-M.; Mergny, J.-L. Interaction of an acridine dimer with DNA quadruplex structures. *J. Biomol. Struct. Dyn.* **2001**, *19*, 505–513. [CrossRef]

85. Sedghi Masoud, S.; Nagasawa, K. i-motif-binding ligands and their effects on the structure and biological functions of i-motif. *Chem. Pharm. Bull. (Tokyo)* **2018**, *66*, 1091–1103. [CrossRef]

86. Kendrick, S.; Kang, H.-J.; Alam, M.P.; Madathil, M.M.; Agrawal, P.; Gokhale, V.; Yang, D.; Hecht, S.M.; Hurley, L.H. The Dynamic Character of the BCL2 Promoter i-Motif Provides a Mechanism for Modulation of Gene Expression by Compounds That Bind Selectively to the Alternative DNA Hairpin Structure. *J. Am. Chem. Soc.* **2014**, *136*, 4161–4171. [CrossRef]

87. Lacroix, L.; Liénard, H.; Labourier, E.; Djavaheri-Mergny, M.; Lacoste, J.; Leffers, H.; Tazi, J.; Hélène, C.; Mergny, J.L. Identification of two human nuclear proteins that recognise the cytosine-rich strand of human telomeres in vitro. *Nucleic Acids Res.* **2000**, *28*, 1564–1575. [CrossRef]

88. Lannes, L.; Young, P.; Richter, C.; Morgner, N.; Schwalbe, H. Interaction of the N-terminal tandem domains of hnRNP LL with the BCL2 promoter i-motif DNA sequence. *ChemBioChem* **2017**, *18*, 2033–2044. [CrossRef]

89. Felsenfeld, G.; Davies, D.R.; Rich, A. Formation of a three-stranded polynucleotide molecule. *J. Am. Chem. Soc.* **1957**, *79*, 2023–2024. [CrossRef]

90. Moser, H.E.; Dervan, P.B. Sequence-specific cleavage of double helical DNA by triple helix formation. *Science* **1987**, *238*, 645–650. [CrossRef]

91. Beal, P.A.; Dervan, P.B. Second structural motif for recognition of DNA by oligonucleotide-directed triple-helix formation. *Science* **1991**, *251*, 1360–1363. [CrossRef] [PubMed]

92. Mirkin, S.M.; Lyamichev, V.I.; Drushlyak, K.N.; Dobrynin, V.N.; Filippov, S.A.; Frank-Kamenetskii, M.D. DNA H form requires a homopurine-homopyrimidine mirror repeat. *Nature* **1987**, *330*, 495–497. [CrossRef] [PubMed]

93. Mukherjee, A.; Vasquez, K.M. Triplex technology in studies of DNA damage, DNA repair, and mutagenesis. *Biochimie* **2011**, *93*, 1197–1208. [CrossRef] [PubMed]

94. Rooney, S.M.; Moore, P.D. Antiparallel, intramolecular triplex DNA stimulates homologous recombination in human cells. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 2141–2144. [CrossRef]

95. Wang, G.; Seidman, M.M.; Glazer, P.M. Mutagenesis in mammalian cells induced by triple helix formation and transcription-coupled repair. *Science* **1996**, *271*, 802–805. [CrossRef]

96. Portes-Sentis, S.; Sergeant, A.; Gruffat, H. A particular DNA structure is required for the function of a cis-acting component of the Epstein-Barr virus OriLyt origin of replication. *Nucleic Acids Res.* **1997**, *25*, 1347–1354. [CrossRef]

97. Plum, G.E.; Park, Y.W.; Singleton, S.F.; Dervan, P.B.; Breslauer, K.J. Thermodynamic characterization of the stability and the melting behavior of a DNA triplex: A spectroscopic and calorimetric study. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 9436–9440. [CrossRef]

98. Plum, G.E.; Pilch, D.S.; Singleton, S.F.; Breslauer, K.J. Nucleic acid hybridization: Triplex stability and energetics. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 319–350. [CrossRef]

99. Reither, S.; Jeltsch, A. Specificity of DNA triple helix formation analyzed by a FRET assay. *BMC Biochem.* **2002**, *3*, 27. [CrossRef]

100. Lee, J.E.; Kim, T.; Kim, S.Y.; Kim, S.W. Detection of triple helix DNA formation of guanine-rich oligonucleotide in sodium ion abundant buffer by cross-checking FRET scheme. *Chem. Phys. Lett.* **2010**, *490*, 230–233. [CrossRef]

101. Idili, A.; Vallée-Bélisle, A.; Ricci, F. Programmable pH-triggered DNA nanoswitches. *J. Am. Chem. Soc.* **2014**, *136*, 5836–5839. [CrossRef] [PubMed]

102. Shields, G.C.; Laughton, C.A.; Orozco, M. Molecular dynamics simulations of the d(T·A·T) triple helix. *J. Am. Chem. Soc.* **1997**, *119*, 7463–7469. [CrossRef]

103. Shields, G.C.; Laughton, C.A.; Orozco, M. Molecular dynamics simulation of a PNA·DNA·PNA triple helix in aqueous solution. *J. Am. Chem. Soc.* **1998**, *120*, 5895–5904. [CrossRef]

104. Esguerra, M.; Nilsson, L.; Villa, A. Triple helical DNA in a duplex context and base pair opening. *Nucleic Acids Res.* **2014**, *42*, 11329–11338. [CrossRef]

105. Ramreddy, T.; Kombrabail, M.; Krishnamoorthy, G.; Rao, B.J. Site-specific dynamics in TAT triplex DNA as revealed by time-domain fluorescence of 2-aminopurine. *J. Phys. Chem. B* **2009**, *113*, 6840–6846. [CrossRef]

106. Powell, S.W.; Jiang, L.; Russu, I.M. Proton exchange and base pair opening in a DNA triple helix. *Biochemistry* **2001**, *40*, 11065–11072. [CrossRef]

107. Jiang, L. Proton exchange and local stability in a DNA triple helix containing a G.TA triad. *Nucleic Acids Res.* **2001**, *29*, 4231–4237. [CrossRef]

108. Coman, D. Site-resolved stabilization of a DNA triple helix by magnesium ions. *Nucleic Acids Res.* **2004**, *32*, 878–883. [CrossRef]

109. Eick, A.; Riechert-Krause, F.; Weisz, K. Binding and NMR structural studies on indoloquinoline-oligonucleotide conjugates targeting duplex DNA. *Bioconjugate Chem.* **2012**, *23*, 1127–1137. [CrossRef]

110. Dickerhoff, J.; Riechert-Krause, F.; Seifert, J.; Weisz, K. Exploring multiple binding sites of an indoloquinoline in triple-helical DNA: A paradigm for DNA triplex-selective intercalators. *Biochimie* **2014**, *107*, 327–337. [CrossRef]

111. Ha, S.C.; Lowenhaupt, K.; Rich, A.; Kim, Y.-G.; Kim, K.K. Crystal structure of a junction between B-DNA and Z-DNA reveals two extruded bases. *Nature* **2005**, *437*, 1183–1186. [CrossRef] [PubMed]

112. Wang, A.H.; Quigley, G.J.; Kolpak, F.J.; Crawford, J.L.; Van Boom, J.H.; Van der Marel, G.; Rich, A. Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature* **1979**, *282*, 680–686. [CrossRef] [PubMed]

113. Pohl, F.M.; Jovin, T.M. Salt-induced co-operative conformational change of a synthetic DNA: Equilibrium and kinetic studies with poly (dG-dC). *J. Mol. Biol.* **1972**, *67*, 375–396. [CrossRef]

114. Wittig, B.; Wölfl, S.; Dorbic, T.; Vahrson, W.; Rich, A. Transcription of human c-myc in permeabilized nuclei is associated with formation of Z-DNA in three discrete regions of the gene. *EMBO J.* **1992**, *11*, 4653–4663. [CrossRef] [PubMed]

115. Liu, R.; Liu, H.; Chen, X.; Kirby, M.; Brown, P.O.; Zhao, K. Regulation of CSF1 promoter by the SWI/SNF-like BAF complex. *Cell* **2001**, *106*, 309–318. [CrossRef]

116. Maruyama, A.; Mimura, J.; Harada, N.; Itoh, K. Nrf2 activation is associated with Z-DNA formation in the human HO-1 promoter. *Nucleic Acids Res.* **2013**, *41*, 5223–5234. [CrossRef]

117. Takaoka, A.; Taniguchi, T. Cytosolic DNA recognition for triggering innate immune responses. *Adv. Drug Delivery Rev.* **2008**, *60*, 847–857. [CrossRef]

118. Kim, Y.-G.; Lowenhaupt, K.; Oh, D.-B.; Kim, K.K.; Rich, A. Evidence that vaccinia virulence factor E3L binds to Z-DNA in vivo: Implications for development of a therapy for poxvirus infection. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 1514–1518. [CrossRef]

119. Bothe, J.R.; Lowenhaupt, K.; Al-Hashimi, H.M. Sequence-specific B-DNA flexibility modulates Z-DNA formation. *J. Am. Chem. Soc.* **2011**, *133*, 2016–2018. [CrossRef]

120. Bae, S.; Kim, D.; Kim, K.K.; Kim, Y.-G.; Hohng, S. Intrinsic Z-DNA is stabilized by the conformational selection mechanism of Z-DNA-binding proteins. *J. Am. Chem. Soc.* **2011**, *133*, 668–671. [CrossRef]

121. Bae, S.; Kim, Y.; Kim, D.; Kim, K.K.; Kim, Y.-G.; Hohng, S. Energetics of Z-DNA binding protein-mediated helicity reversals in DNA, RNA, and DNA–RNA Duplexes. *J. Phys. Chem. B* **2013**, *117*, 13866–13871. [CrossRef] [PubMed]

122. Lee, M.; Kim, S.H.; Hong, S.C. Minute negative superhelicity is sufficient to induce the B-Z transition in the presence of low tension. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 4985–4990. [CrossRef] [PubMed]

123. Kim, S.H.; Lim, S.-H.; Lee, A.-R.; Kwon, D.H.; Song, H.K.; Lee, J.-H.; Cho, M.; Johner, A.; Lee, N.-K.; Hong, S.-C. Unveiling the pathway to Z-DNA in the protein-induced B–Z transition. *Nucleic Acids Res.* **2018**, *46*, 4129–4137. [CrossRef] [PubMed]

124. Bothe, J.R.; Lowenhaupt, K.; Al-Hashimi, H.M. Incorporation of CC steps into Z-DNA: Interplay between B–Z junction and Z-DNA helical formation. *Biochemistry* **2012**, *51*, 6871–6879. [CrossRef]

125. Dumat, B.; Larsen, A.F.; Wilhelmsson, L.M. Studying Z-DNA and B- to Z-DNA transitions using a cytosine analogue FRET-pair. *Nucleic Acids Res.* **2016**, *44*, e101Ce101. [CrossRef]

126. Xu, Y.; Zhang, Y.X.; Sugiyama, H.; Umano, T.; Osuga, H.; Tanaka, K. (P)-helicene displays chiral selection in binding to Z-DNA. *J. Am. Chem. Soc.* **2004**, *126*, 6566–6567. [CrossRef]

127. Zhang, Y.; Cui, Y.; An, R.; Liang, X.; Li, Q.; Wang, H.; Wang, H.; Fan, Y.; Dong, P.; Li, J.; et al. Topologically constrained formation of stable Z-DNA from normal sequence under physiological conditions. *J. Am. Chem. Soc.* **2019**, *141*, 7758–7764. [CrossRef]

128. Kladde, M.P.; Kohwi, Y.; Kohwi-Shigematsu, T.; Gorski, J. The non-B-DNA structure of d(CA/TG)n differs from that of Z-DNA. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 1898–1902. [CrossRef]

129. Peck, L.J.; Wang, J.C. Energetics of B-to-Z transition in DNA. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 6206–6210. [CrossRef]

130. Ellison, M.J.; Feigon, J.; Kelleher, R.J.; Wang, A.H.; Habener, J.F.; Rich, A. An assessment of the Z-DNA forming potential of alternating dA-dT stretches in supercoiled plasmids. *Biochemistry* **1986**, *25*, 3648–3655. [CrossRef]

131. McLean, M.J.; Blaho, J.A.; Kilpatrick, M.W.; Wells, R.D. Consecutive A X T pairs can adopt a left-handed DNA structure. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 5884–5888. [CrossRef] [PubMed]

132. Coll, M.; Fita, I.; Lloveras, J.; Subirana, J.A.; Bardella, F.; Huynh-Dinh, T.; Igolen, J. Structure of d(CACGTG), a Z-DNA hexamer containing AT base pairs. *Nucleic Acids Res.* **1988**, *16*, 8695–8705. [CrossRef] [PubMed]

133. Eichman, B.F.; Schroth, G.P.; Basham, B.E.; Ho, P.S. The intrinsic structure and stability of out-of-alternation base pairs in Z-DNA. *Nucleic Acids Res.* **1999**, *27*, 543–550. [CrossRef] [PubMed]

134. Abascal, J.L.F.; Domercq, M.; Gil Montoro, J.C. Computer simulation of the ionic atmosphere around Z-DNA. *J. Phys. Chem. B* **2006**, *110*, 25080–25090. [CrossRef]

135. Herbert, A.; Rich, A. The biology of left-handed Z-DNA. *J. Biol. Chem.* **1996**, *271*, 11595–11598. [CrossRef]

136. Kwon, J.-A.; Rich, A. Biological function of the vaccinia virus Z-DNA-binding protein E3L: Gene transactivation and antiapoptotic activity in HeLa cells. *Proc. Natl. Acas. Sci. USA* **2005**, *102*, 12759–12764. [CrossRef]

137. Rich, A.; Zhang, S. Timeline: Z-DNA: The long road to biological function. *Nat. Rev. Genet.* **2003**, *4*, 566–572. [CrossRef]

138. Tashiro, R.; Sugiyama, H. A Nanothermometer based on the different π stackings of B- and Z-DNA. *Angew Chem. Int. Ed.* **2003**, *42*, 6018–6020. [CrossRef]

139. Tashiro, R.; Sugiyama, H. Biomolecule-based switching devices that respond inversely to thermal stimuli. *J. Am. Chem. Soc.* **2005**, *127*, 2094–2097. [CrossRef]

140. Behe, M.; Felsenfeld, G. Effects of methylation on a synthetic polynucleotide: The B–Z transition in poly(dG-m5dC) poly(dG-m5dC). *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 1619–1623. [CrossRef]

141. Temiz, N.A.; Donohue, D.E.; Bacolla, A.; Luke, B.T.; Collins, J.R. The role of methylation in the intrinsic dynamics of B- and Z-DNA. *PLoS ONE* **2012**, *7*, e35558. [CrossRef] [PubMed]

142. Xu, Y.; Ikeda, R.; Sugiyama, H. 8-Methylguanosine: A powerful Z-DNA stabilizer. *J. Am. Chem. Soc.* **2003**, *125*, 13519–13524. [CrossRef] [PubMed]

143. Kochoyan, M.; Leroy, J.L.; Guéron, M. Proton exchange and base-pair lifetimes in a deoxy-duplex containing a purine-pyrimidine step and in the duplex of inverse sequence. *J. Mol. Biol.* **1987**, *196*, 599–609. [CrossRef]

144. Kochoyan, M.; Leroy, J.L.; Guéron, M. Processes of base-pair opening and proton exchange in Z-DNA. *Biochemistry* **1990**, *29*, 4799–4805. [CrossRef]

145. Kang, Y.-M.; Bang, J.; Lee, E.-H.; Ahn, H.-C.; Seo, Y.-J.; Kim, K.K.; Kim, Y.-G.; Choi, B.-S.; Lee, J.-H. NMR Spectroscopic elucidation of the B−Z transition of a DNA double helix induced by the Zα domain of human ADAR1. *J. Am. Chem. Soc.* **2009**, *131*, 11485–11491. [CrossRef]

146. Herbert, A.; Lowenhaupt, K.; Spitzner, J.; Rich, A. Chicken double-stranded RNA adenosine deaminase has apparent specificity for Z-DNA. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 7550–7554. [CrossRef]

147. Schwartz, T.; Rould, M.A.; Lowenhaupt, K.; Herbert, A.; Rich, A. Crystal structure of the Zα domain of the human editing enzyme ADAR1 bound to left-handed Z-DNA. *Science* **1999**, *284*, 1841–1845. [CrossRef]

148. Wang, Z.; Choi, M.K.; Ban, T.; Yanai, H.; Negishi, H.; Lu, Y.; Tamura, T.; Takaoka, A.; Nishikura, K.; Taniguchi, T. Regulation of innate immune responses by DAI (DLM-1/ZBP1) and other DNA-sensing molecules. *Proc. Natl. Acad. Sci. USA.* **2008**, *105*, 5477–5482. [CrossRef]

149. Yanai, H.; Savitsky, D.; Tamura, T.; Taniguchi, T. Regulation of the cytosolic DNA-sensing system in innate immunity: A current view. *Curr. Opin. Immunol.* **2009**, *21*, 17–22. [CrossRef]

150. Kim, Y.-G.; Muralinath, M.; Brandt, T.; Pearcy, M.; Hauns, K.; Lowenhaupt, K.; Jacobs, B.L.; Rich, A. A role for Z-DNA binding in vaccinia virus pathogenesis. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 6974–6979. [CrossRef]

151. Rothenburg, S.; Deigendesch, N.; Dittmar, K.; Koch-Nolte, F.; Haag, F.; Lowenhaupt, K.; Rich, A. A PKR-like eukaryotic initiation factor 2α kinase from zebrafish contains Z-DNA binding domains instead of dsRNA binding domains. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 1602–1607. [CrossRef] [PubMed]

152. Lee, A.-R.; Kim, N.-H.; Seo, Y.-J.; Choi, S.-R.; Lee, J.-H. Thermodynamic model for B-Z transition of DNA induced by Z-DNA binding proteins. *Molecules* **2018**, *23*, 2748. [CrossRef] [PubMed]

153. Lee, E.-H.; Seo, Y.-J.; Ahn, H.-C.; Kang, Y.-M.; Kim, H.-E.; Lee, Y.-M.; Choi, B.-S.; Lee, J.-H. NMR study of hydrogen exchange during the B-Z transition of a DNA duplex induced by the Zα domains of yatapoxvirus E3L. *FEBS Letters* **2010**, *584*, 4453–4457. [CrossRef] [PubMed]

154. Lee, A.-R.; Park, C.-J.; Cheong, H.-K.; Ryu, K.-S.; Park, J.-W.; Kwon, M.-Y.; Lee, J.; Kim, K.K.; Choi, B.-S.; Lee, J.-H. Solution structure of the Z-DNA binding domain of PKR-like protein kinase from *Carassius auratus* and quantitative analyses of the intermediate complex during B–Z transition. *Nucleic Acids Res.* **2016**, *44*, 2936–2948. [CrossRef]

155. Lee, A.-R.; Hwang, J.; Hur, J.H.; Ryu, K.-S.; Kim, K.K.; Choi, B.-S.; Kim, N.-K.; Lee, J.-H. NMR dynamics study reveals the Zα domain of human ADAR1 associates with and dissociates from Z-RNA more slowly than Z-DNA. *ACS Chem. Biol.* **2019**, *14*, 245–255. [CrossRef]

156. Ha, S.C.; Choi, J.; Hwang, H.-Y.; Rich, A.; Kim, Y.-G.; Kim, K.K. The structures of non-CG-repeat Z-DNAs co-crystallized with the Z-DNA-binding domain, hZα$_{ADAR1}$. *Nucleic Acids Res.* **2008**, *37*, 629–637. [CrossRef]

157. Kim, H.-E.; Ahn, H.-C.; Lee, Y.-M.; Lee, E.-H.; Seo, Y.-J.; Kim, Y.-G.; Kim, K.K.; Choi, B.-S.; Lee, J.-H. The Zβ domain of human DAI binds to Z-DNA via a novel B-Z transition pathway. *FEBS Lett.* **2011**, *585*, 772–778. [CrossRef]

158. Sheardy, R.D.; Suh, D.; Kurzinsky, R.; Doktycz, M.J.; Benight, A.S.; Chaires, J.B. Sequence dependence of the free energy of B-Z junction formation in deoxyoligonucleotides. *J. Mol. Biol.* **1993**, *231*, 475–488. [CrossRef]

159. Sheardy, R.D.; Levine, N.; Marotta, S.; Suh, D.; Chaires, J.B. A thermodynamic investigation of the melting of B-Z junction forming DNA oligomers. *Biochemistry* **1994**, *33*, 1385–1391. [CrossRef]

160. González, V.M.; Fuertes, M.A.; Pérez, J.M. Kinetics of the salt-induced B-to Z-DNA transition. *Biophys. J.* **1998**, *27*, 417–423. [CrossRef]

161. Kim, D.; Hur, J.; Park, K.; Bae, S.; Shin, D.; Ha, S.C.; Hwang, H.-Y.; Hohng, S.; Lee, J.-H.; Lee, S.; et al. Distinct Z-DNA binding mode of a PKR-like protein kinase containing a Z-DNA binding domain (PKZ). *Nucleic Acids Res.* **2014**, *42*, 5937–5948. [CrossRef] [PubMed]