*Published online 11 December 2008* *Nucleic Acids Research, 2009, Vol. 37, No. 2* ***629–637***
*doi:10.1093/nar/gkn976*

# The structures of non-CG-repeat Z-DNAs co-crystallized with the Z-DNA-binding domain, $hZ\alpha_{ADAR1}$

**Sung Chul Ha[1], Jongkeun Choi[1], Hye-Yeon Hwang[1], Alexander Rich[2], Yang-Gyun Kim[3,*] and Kyeong Kyu Kim[1,*]**

[1]Department of Molecular Cell Biology, Samsung Biomedical Research Institute, Sungkyunkwan University School of Medicine, Suwon 440-746, Korea, [2]Department of Biology, Massachusetts Institute of technology, Cambridge, Massachusetts 02139, USA and [3]Department of Chemistry, Sungkyunkwan University, Suwon 440-746, Korea

Received October 17, 2008; Revised and Accepted November 19, 2008

## ABSTRACT

**The Z-DNA conformation preferentially occurs at alternating purine-pyrimidine repeats, and is specifically recognized by Zα domains identified in several Z-DNA-binding proteins. The binding of Zα to foreign or chromosomal DNA in various sequence contexts is known to influence various biological functions, including the DNA-mediated innate immune response and transcriptional modulation of gene expression. For these reasons, understanding its binding mode and the conformational diversity of Zα bound Z-DNAs is of considerable importance. However, structural studies of Zα bound Z-DNA have been mostly limited to standard CG-repeat DNAs. Here, we have solved the crystal structures of three representative non-CG repeat DNAs, $d(CACGTG)_2$, $d(CGTACG)_2$ and $d(CGGCCG)_2$ complexed to $hZ\alpha_{ADAR1}$ and compared those structures with that of $hZ\alpha_{ADAR1}/d(CGCGCG)_2$ and the Zα-free Z-DNAs. $hZ\alpha_{ADAR1}$ bound to each of the three Z-DNAs showed a well conserved binding mode with very limited structural deviation irrespective of the DNA sequence, although varying numbers of residues were in contact with Z-DNA. Z-DNAs display less structural alterations in the Zα-bound state than in their free form, thereby suggesting that conformational diversities of Z-DNAs are restrained by the binding pocket of Zα. These data suggest that Z-DNAs are recognized by Zα through common conformational features regardless of the sequence and structural alterations.**

## INTRODUCTION

DNA can adopt various secondary conformations other than the classical right-handed B-DNA structure under certain specific physiological conditions (1). Left-handed Z-DNA has been studied in detail by numerous methods over the last two decades since its first crystal structure was solved (2). Z-DNA conformations in many different sequences have been a challenge to crystallize. Nonetheless, several structures have been determined by X-ray crystallographic study (3). Many characteristic features of the Z-DNA structure have been identified from the accumulated structural data. The overall shapes of most Z-DNA structures share common features that are similar to those seen in the first crystallized Z-DNA structure of $d(CGCGCG)_2$ (2). Z-DNA has a zig-zag sugar phosphate backbone and is longer and thinner than B-DNA. Its nucleotides form in a dinucleotide repeat in which they alternate with *syn* and *anti* conformations. Z-DNA is favored in alternating pyrimidine-purine (APP) sequences (4). The alternating CG-repeat sequence is the most favorable energetically for Z-DNA formation (5). Networks of water molecules hydrate Z-DNA, forming hydrogen bonds with atoms in both the backbone and base. However, some Z-DNAs form with different structural features, especially those with sequences without APP or including A-T base pairs (6–8). When A-T base pairs are introduced, the overall Z-DNA structure

*To whom correspondence should be addressed. Tel: +82 31 299 6136; Fax: +82 31 299 6159; Email: kkim@med.skku.ac.kr
Correspondence may also be addressed to Yang-Gyun Kim, Tel: +82 31 299 4563; Fax: +82 31 299 4575; Email: ygkimmit@skku.edu

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID codes of $hZ\alpha_{ADAR1}$ in complex with $d(CACGTG)_2$, $d(CGTACG)_2$ and $d(CGGCCG)_2$ are 3F21,3F22 and 3F23, respectively)

© 2008 The Author(s)
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

becomes partially distorted due to disruption of the hydration spine (6). In addition, the non-APP base pairs have been found to be highly buckled when compared with other base pairs in Z-DNA (7,8). Up to now, however, there have been only a limited number of structural studies of Z-DNAs containing non-APP or A-T base pairs. Most studies have been carried out in the presence of excess cations or using dsDNAs modified by methylation or bromation. Thus, study of Z-DNA under low salt conditions or in the presence of Z-DNA-binding domains (Zα) remains to be explored, and this may provide insight into the effect that sequence variability has on Z-DNA conformation under physiological conditions.

The protein Zα domain was first identified from human ADAR1 (double-stranded RNA adenosine deaminase) and subsequently found in other proteins (DLM1, E3L and PKZ). These domains provide a unique opportunity to explore Z-DNA and its various roles in biological systems (9–13). The Zα domains are highly specific for the Z-conformation of nucleic acids, including dsDNA, dsRNA as well as DNA–RNA hybrids. They have binding affinities in the low-nanomolar range. In the crystal structures of double-stranded $d(CGCGCG)_2$ complexed with $hZ\alpha_{ADAR1}$, $mZ\alpha_{DLM1}$, $yabZ\alpha_{E3L}$ or $hZ\beta_{DLM1}$ (10,11,14–16), a close resemblance was found between the bound and unbound states of the Z-DNA structure of $d(CGCGCG)_2$ (2). It is now known that Z-DNA is found in the genome as an active transcription modifier that functions by modulating chromatin structure (17,18). The Z-DNA conformation is not limited only to APP sequences, but can appear in many other nucleotide sequences. Recent functional studies of the Z-DNA-binding domains in vaccinia E3L showed that it acts as a transcription modulator of several apoptosis-related genes (19). More recently, the Z-DNA-binding protein DLM1 has been found to act in the innate immune system as a cytosolic receptor for dsDNA, recognizing foreign pathogenic DNA (20). These results support the idea that diverse sequences of Z-DNA can be recognized by Z-DNA-binding domains and that their binding is essential for cellular processes. To understand the binding mode of Zα to Z-DNAs in various sequence contexts, it is necessary to investigate structural features of Z-DNAs bound to Zα and to compare their structures with Z-DNAs stabilized by base-modification and/or positively charged ions. In this regard, we undertook a study to solve Z-DNA structures with non-CG-repeat sequences stabilized by the same Zα domain. Here we report the co-crystal structures of three non-CG-repeat Z-DNAs containing either A-T base pairs or non-APP sequence bound to $hZ\alpha_{ADAR1}$. These complexes reveal how the structural diversity of Z-DNA caused by non-CG-repeat sequences is recognized and stabilized by the Z-DNA-binding domain.

## MATERIALS AND METHODS

### Expression and purification

$hZ\alpha_{ADAR1}$ (residues 133–209) from human ADAR1 was expressed and purified as described previously (21). In brief, $hZ\alpha_{ADAR1}$ was purified through sequential chromatographic steps involving HiTrap metal affinity column (GE Healthcare, Piscataway, NJ), thrombin digestion for the removal of N-terminal his-tag and a Resource S column (GE Healthcare, Piscataway, NJ). The purity and concentration of $hZ\alpha_{ADAR1}$ were estimated by SDS–PAGE and the Bradford method, respectively. DNAs were purchased (IDT, Coralville, IA) and purified as described previously (22).

### Crystallization

The DNA used for crystallization all had 6nt, plus a 5′ T overhang which acts as a stabilizing capping residue. For crystallization, $hZ\alpha_{ADAR1}$ was mixed with dsDNA [(d(TCGCCCG)::d(TCGGGCG), $d(TCACGTG)_2$, $d(TCGTACG)_2$ or $d(TCGGCCG)_2$)] at a 0.66 mM : 0.33 mM molar ratio in 5 mM HEPES-NaOH pH 7.5 containing 20 mM NaCl, and incubated at 303 K for at least 2 h. All crystallization experiments were performed using the hanging drop vapor diffusion method with a VDX plate at 295 K. Among the four dsDNA used in crystallization, only d(TCGCCCG)::d(TCGGGCG) was successfully co-crystallized with $hZ\alpha_{ADAR1}$ when ammonium sulfate was used as precipitant. For crystallization of the other three complexes, crystals of $hZ\alpha_{ADAR1}$/d(TCGCCCG)::d(TCGGGCG) were used as seeds, and initial crystals were again used as seeds for further crystallization. Diffraction quality crystals were obtained within a month using 2.2 M ammonium sulfate and 10% glycerol in the reservoir solution.

### Data collection and structure determination

Preliminary X-ray diffraction analyses were performed at beamline BL6A of PAL (Pohang, Korea). X-ray diffraction data of frozen crystals were collected at 100 K with a MAR CCD165 detector at the BL41-XU beamline of Spring-8 (Harima, Japan). Crystals were frozen either in liquid nitrogen directly or by using paraton as a cryoprotectant. Diffraction data were processed and integrated using HKL2000 (23). The unit cell parameters, space group and other data collection statistics are summarized in Supplementary Table 1. The crystal structure of $hZ\alpha_{ADAR1}$/$d(TCGCGCG)_2$ (PDB ID 1QBJ) was used for the initial model of the other complex structures. Refinement and model building were performed by CNS (24) and O (25), respectively. The refinement statistics are summarized in Supplementary Table 1. The crystal structure of $hZ\alpha_{ADAR1}$/d(TCGCCCG)::d(TCGGGCG) was not refined because of ambiguity in base assignment. All figures were drawn using Molscript, Raster3D and Pymol (26,27, http://www.pymol.org). The structural superposition was performed using the LSQKAB CCP4 program (28).

## RESULTS AND DISCUSSION

### Structure determination

The crystal structures of $hZ\alpha_{ADAR1}$ complexed with three non-CG-repeat dsDNAs, $d(TCACGTG)_2$, $d(TCGTACG)_2$

and d(TCGGCCG)$_2$ were determined at resolutions of 2.2, 2.5 and 2.7 Å, respectively (Supplementary Table 1). The simulated annealed omit maps contoured at 3σ near the altered sequences confirmed that current structural information represents each non-CG-repeat DNA (Supplementary Figure 1). In all three structures, the deoxythymidine overhangs at the 5′ end of the oligonucleotides were not modeled due to their weak electron densities, and therefore, are not mentioned in current study. There are three hZα$_{ADAR1}$ domains, assigned as chains A, B and C, and three DNA strands, chains D, E and F, in one asymmetric unit. Chains D and E form the Z-DNA duplex, and chain F forms a duplex with chain F in another asymmetric unit that is related by crystallographic 2-fold symmetry. Each DNA strand is bound to one hZα$_{ADAR1}$ domain. The protein/DNA complex made up of the chain C and chain F pair was used for structural analyses unless specified otherwise because its average temperature factor is the lowest among the three complexes in the same asymmetric unit.

### Overall structures

Overall, the hZα$_{ADAR1}$ domains bound to the three different Z-DNAs used in this study have almost identical structures to that of hZα$_{ADAR1}$ bound to d(CGCGCG)$_2$ (14). hZα$_{ADAR1}$ has an α/β topology containing a three-helix bundle flanked on one side by a twisted antiparallel β sheet. The root mean square deviations (RMSDs) between hZα$_{ADAR1}$ bound to d(CGCGCG)$_2$ and hZα$_{ADAR1}$ bound to d(CACGTG)$_2$, d(CGTACG)$_2$ or d(CGGCCG)$_2$ are 0.52, 0.19 and 0.19 Å, respectively, when calculated using the 64 Cα atoms of hZα$_{ADAR1}$ (Figure 1A). The three Z-DNAs bound to hZα$_{ADAR1}$ are all in the Z-conformation with alternating *anti*-and *syn*-glycosidic bonds regardless of their sequence (Figure 1B). With the exception of purines G3 and pyrimidines C4 of d(CGGCCG)$_2$ that adopt the *anti* and *syn* conformations, respectively, all other purines and pyrimidines are in the *syn* and *anti* conformations, respectively (Figure 1B). The calculated double-stranded DNA parameters are within the range of a Z conformation of DNA (Supplementary Table 2). For Z-DNAs complexed with hZα$_{ADAR1}$, the RMSD values between d(CGCGCG)$_2$ and other DNAs, d(CACGTG)$_2$, d(CGTACG)$_2$ and d(CGGCCG)$_2$ were 0.28, 0.33 and 0.56 Å, respectively, when calculated with 63 DNA backbone atoms of chain F for each DNA structure (Figure 1B and Supplementary Table 3).

### Interactions between hZα$_{ADAR1}$ and non-CG-repeat Z-DNAs

The structural analyses of hZα$_{ADAR1}$, mZα$_{DLM1}$ and yabZα$_{E3L}$ bound to the Z conformation of d(CGCGCG)$_2$ revealed well-conserved interactions with the Zα domains (10,11,14). Zα domains with a winged helix–turn–helix motif bind to Z-DNA in a conformation-specific manner. The residues in the recognition helix (α3) and in the wing play critical roles in binding Z-DNA through direct or water-mediated hydrogen bonds and van der Waals interactions (Figure 2). On Z-DNA, one continuous surface composed of the sugar-phosphate backbone of Z-DNA is mostly involved in protein contact.

Similarly, in the crystal structures of hZα$_{ADAR1}$ bound to the three non-CG-repeat dsDNAs the residues located in α3 and the wing mostly make contact with the Z-DNA backbones either directly or through water-mediated interactions (Figure 2). When the DNA-binding surfaces of the hZα$_{ADAR1}$ domains bound to the four Z-DNAs are compared, their curvatures and surface charge distributions are nearly identical (Figure 3). These results, together with the limited structural alterations found in hZα$_{ADAR1}$ domains, strongly demonstrate that the overall binding mode of hZα$_{ADAR1}$ to Z-DNA are well conserved regardless of the sequence of the bound Z-DNA.

The number and type of hZα$_{ADAR1}$ residues contacting bases on each DNA vary, mainly because some of the important residues that were identified as the DNA-contacting residues in hZα$_{ADAR1}$/d(CGCGCG)$_2$ (14) are disordered (Figure 2). However, it does not appears that variations in protein/DNA contacts are related to the sequence of the Z-DNA, since the number of contacting residues is not the same, even among the three hZα$_{ADAR1}$ domains in one asymmetric unit of each Zα$_{ADAR1}$/DNA crystal. For example, in most structures, the roles of Arg174 and Thr191 are neither well defined nor involved in DNA contact. In hZα$_{ADAR1}$/d(CGCGCG)$_2$, Arg174 was shown to form a direct hydrogen bond to a phosphate atom and a water-mediated hydrogen bond to the ribose ring (14). However, in seven of the nine hZα$_{ADAR1}$ domains used in this study, the amine groups of Arg174 are not modeled due to their weak electron densities (Figure 2). In the case of Thr191, no DNA interaction has been found. Similarly, Lys169 and Lys170 are only in contact with DNA in some structures (Figure 2). In contrast, Asn173, Tyr177, Pro192, Pro193 and Trp195 of hZα$_{ADAR1}$ contribute to the recognition of Z-DNA in all cases. Asn173 and Tyr177, located in the recognition helix (α3), recognize the phosphate backbone of Z-DNA through direct or water-mediated hydrogen bonds as observed in hZα$_{ADAR1}$/d(CGCGCG)$_2$ (14). Likewise, Pro192 and Pro193 in the wing interact with Z-DNA through van der Waals interactions. In some cases, Trp195 makes a water-mediated hydrogen bond to a phosphate, but its major role seems to be supporting Tyr177 via the hydrophobic edge-to-face interaction which stabilizes the interactions between Tyr177 and the bases in the N4 position (Figure 2). Moreover, some coordinated waters that are well defined and mediate hydrogen bonds between protein and DNA in the crystal structure of hZα$_{ADAR1}$/d(CGCGCG)$_2$ are not found in some structures.

The decrease in protein/DNA interactions in the three complex structures compared to hZα$_{ADAR1}$/d(CGCGCG)$_2$ can be explained in part by the decrease in diffraction resolution. However, it seems that protein–DNA interactions are not all affected by diffraction resolution since well-defined interactions are consistently observed in all three structures despite their resolution differences. For example, the electron density of the well-defined hydrogen bonds between Tyr177 and the phosphate group of G3 are very clear even in the case of hZα$_{ADAR1}$/d(CGGCCG)$_2$

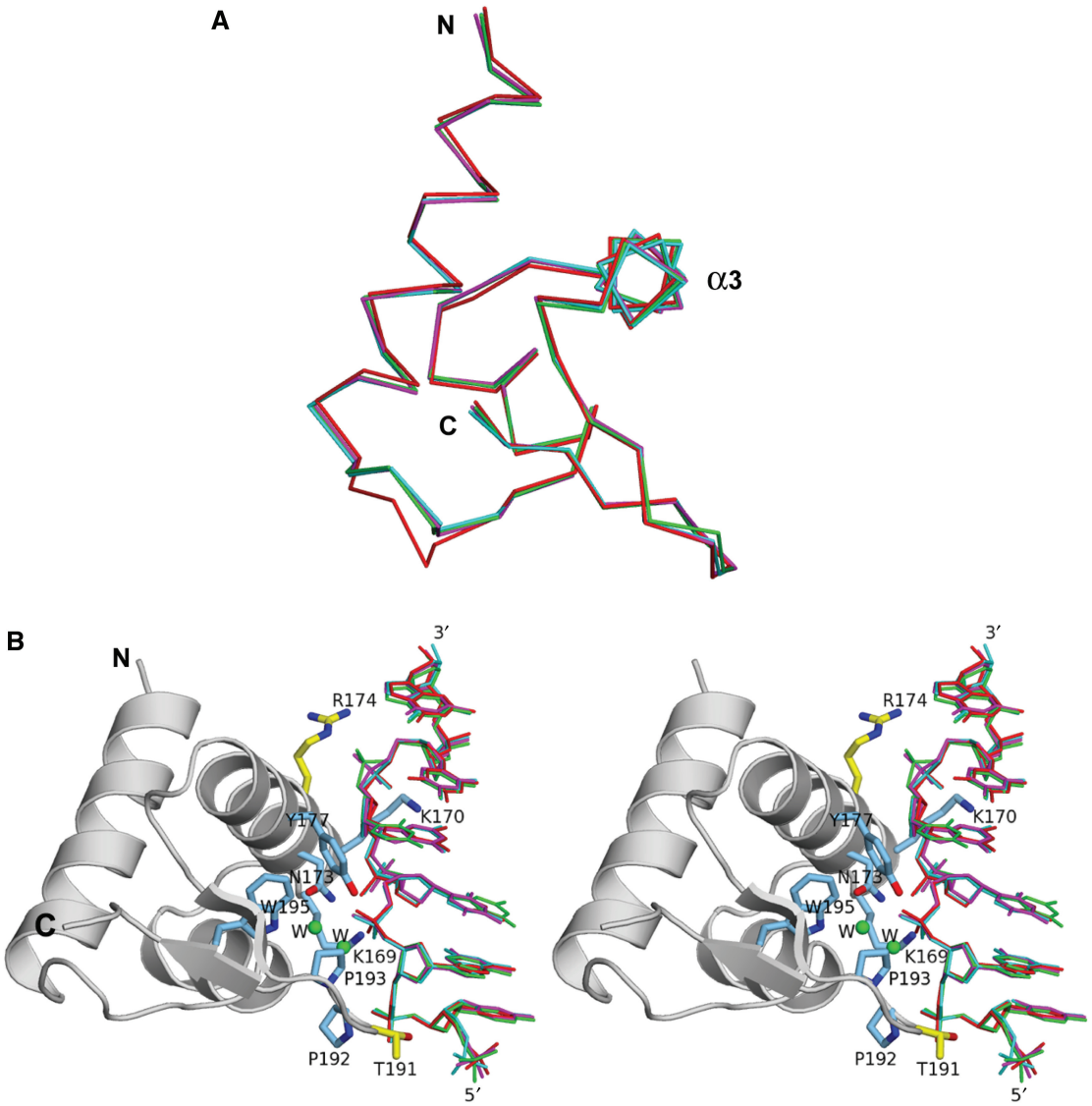


**Figure 1.** Structural comparison of the hZα$_{ADAR1}$ domain and the Z-DNAs in several complexes. **(A)** Overlapping Cα traces of the hZα$_{ADAR1}$ domains when complexed to d(CGCGCG)$_2$ [cyan], (NDB ID PH0001), d(CACGTG) $_2$ [red], d(CGTACG)$_2$ [magenta] and d(CGGCCG)$_2$ [green]. Sixty-four Cα atoms of chain C in each structure were used for the superposition. The N- and C-termini and α3 are labeled. **(B)** Four Z-DNA strands, CACGTG (red), CGTACG (magenta), CGGCCG (green) and CGCGCG (cyan), bound to hZα$_{ADAR1}$ were compared by superimposing four Zα/Z-DNA complexes using 64 Cα atoms of hZα$_{ADAR1}$ domains. In order to show the relative orientation of the Z-DNA and Zα, chain C of the hZα$_{ADAR1}$/d(CGCGCG)$_2$ complex was drawn in a ribbon diagram. The amino acid residues and water molecules involved in DNA binding are drawn as sky blue stick models and green balls, respectively. Arg174 and Thr191 are marked by yellow stick models to differentiate them from other core resides. These residues are mostly disordered in the current study, but are well defined and bind to DNA in the crystal structure of the hZα$_{ADAR1}$/d(CGCGCG)$_2$ complex.

whose structure was determined at 2.7 Å resolution (Supplementary Figure 2). Conversely, it is suspected that some of the unseen protein/DNA interactions from these three Z-DNA structures are neither strong nor essential for Z-DNA recognition. From these results, it can be hypothesized that some of the residues previously identified as Z-DNA binders are not indispensable for Z-DNA binding but that they have auxiliary roles in DNA binding and probably produce tighter binding. However, we cannot rule out the possibility that crystal packing forces may have affected or destabilized some of the interactions between hZα$_{ADAR1}$ and Z-DNA.

In the Zα/d(CGCGCG)$_2$ structure as well as in the current crystal structure of hZα$_{ADAR1}$/d(CATGCG)$_2$, the tyrosine residue of the recognition helix (α3) has a unique role in recognizing the C-8 carbon of the *syn* deoxyguanosine at the forth position (G4) via a CH–π interaction (10,11,14; Figure 2A). In the case of Zα/d(CGTACG)$_2$, the deoxyadenosine (A4) also adopts the *syn* conformation, which was very similar to that observed in previous studies (Figure 2B). More interesting is the observation that hZα$_{ADAR1}$ stabilizes the *syn* conformation of deoxycytidine at the forth position (C4) of d(CGGCCG)$_2$ (Figure 2C). Generally, the *syn* conformation is not

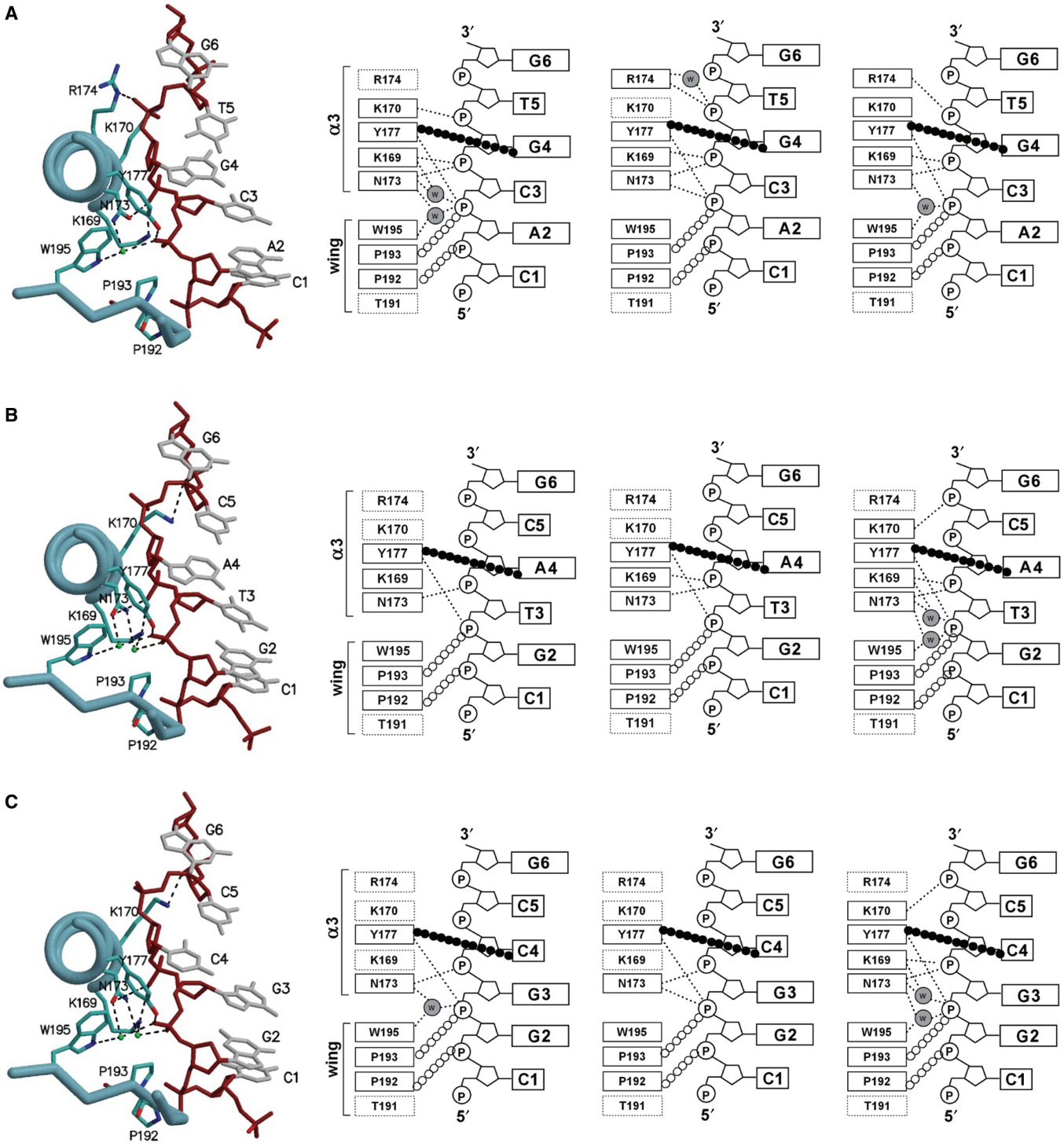


**Figure 2.** The protein–DNA interactions in three different Zα-non-alternating CG ZDNA complexes. **(A)** $hZ\alpha_{ADAR1}/d(CACGTG)_2$, **(B)** $hZ\alpha_{ADAR1}/d(CGTACG)_2$ and **(C)** $hZ\alpha_{ADAR1}/d(CGGCCG)_2$. One representative Zα/DNA complex (chains C and F) of the three complexes in one asymmetric unit is drawn as a tubular ribbon diagram (left). The three complexes in the asymmetric unit are then shown as schematic diagrams (left to right). These are chains A and D, chains B and E, and chains C and F, respectively. In the schematic diagrams, the amino acids identified as the DNA-contacting residues in the structure of the $hZ\alpha_{ADAR1}/d(CGCGCG)_2$ complex are marked by boxes. Disordered residues are indicated by dotted boxes. Hydrogen bonds are represented by dotted lines and van der Waals contacts by open circles. The CH–π interaction between the conserved Tyr residue and *syn* deoxynucleoside is indicated by filled circles. Waters are shown as gray circles. In the tubular ribbon diagrams, the same residues used in the schematic diagrams are drawn as stick models and labeled. The DNA backbones and labeled bases are shown as red and gray stick models, respectively. Water molecules are shown as green spheres. Hydrogen bonds are drawn as dashed lines.

favored for pyrimidine nucleotides unless they are modified (3,4). However in the crystal structure of $hZ\alpha_{ADAR1}$/d(CGGCCG)$_2$, the *syn* conformation of deoxycytidine is stabilized by the CH–π interaction between the C-5 and C-6 carbons of C4 and the π orbital of the tyrosine ring (Figure 2C). These data show that the binding mode of $hZ\alpha_{ADAR1}$ observed for d(CGCGCG)$_2$ is also a template for non-CG-repeat Z-DNAs containing non-APP repeat sequences or A-T base pairs. As with the earlier structure, there are no sequence-specific interactions between Zα and Z-DNA. These results reinforce the idea that the binding of Zα to Z-DNA is sequence-independent but conformation specific.

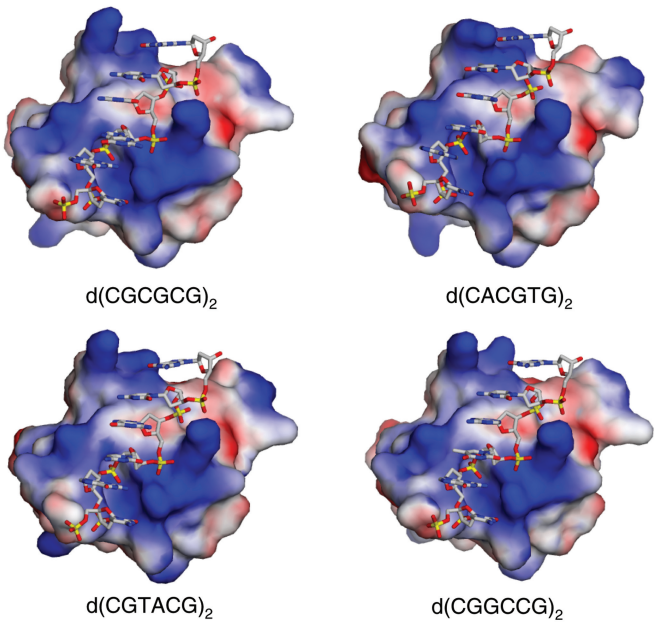


**Figure 3.** Surface charge distributions of the $hZ\alpha_{ADAR1}$ domains complexed with various Z-DNAs, as viewed along the DNA binding cleft. DNA-binding surfaces of the $hZ\alpha_{ADAR1}$ domains bound to d(CGCGCG)$_2$ (upper left), d(CACGTG)$_2$ (upper right), d(CGTACG)$_2$ (bottom left) and d(CGGCCG)$_2$ (bottom right) are drawn, and their surface charge distributions are displayed. Arg174 and Thr174 were not used for the surface charge calculations since they are not well defined in most structures. The red and blue areas represent the negatively and positively charged surfaces, respectively. DNA backbones are shown in stick models with phosphate atoms in yellow, oxygen in red, nitrogen in blue and carbon in gray.

### Structural variations between free and $hZ\alpha_{ADAR1}$-bound Z-DNAs

The crystal structures of d(CGCGCG)$_2$, d(CACGTG)$_2$, d($^{m5}$CGTA$^{m5}$CG)$_2$ and d($^{m5}$CGGC$^{m5}$CG)$_2$ were used to represent Z-DNAs free of protein binding (Supplementary Table 3; 2,8,29,30), and they have been compared with the $hZ\alpha_{ADAR1}$-bound Z-DNA structure. In order to stabilize and crystallize Z-DNA with non-APP or A-T base pairs, base modification is used or it is necessary to add cations such as metal ions or polyamines. The 5 position of cytosine was methylated for the crystallization of Z-form d(CGTACG)$_2$ and d(CGGCCG)$_2$ (8,29), and spermine was added for the crystallization of Z-form d(CGCGCG)$_2$ and d(CACGTA)$_2$ (2,30). When the protein-free and $hZ\alpha_{ADAR1}$-bound Z-DNAs were compared by superimposing each DNA pair using 63 DNA backbone atoms, RMSDs of d(CGCGCG)$_2$, d(CACGTG)$_2$, d(CGTACG)$_2$ and d(CGGCCG)$_2$ pairs were 0.89, 0.75, 0.57 and 0.61 Å, respectively (Figure 4 and Supplementary Table 3). Regardless of their sequence, the main structural differences between the free and bound Z-DNAs were found in the helical rise: when bound to $hZ\alpha_{ADAR1}$ it increased in the N3pN4 step, but decreased in the N2pN3 and N4pN5 steps (Supplementary Table 2). However, the distance changes in the N1pN2 and N5pN6 steps occurred regardless of Zα binding (Supplementary Table 2).

Structural variations between protein-free and $hZ\alpha_{ADAR1}$-bound Z-DNAs were compared directly by plotting the distance between two corresponding atoms in Z-DNAs along the DNA backbone atoms (Figure 5A). In this manner, the structures of non-CG-repeat Z-DNAs and CG-repeat Z-DNA were also analyzed (Figure 5B–D). Structural deviations are plotted for all comparisons of protein-free and Zα-bound Z-DNA sequences. The largest structural variation between protein-free and $hZ\alpha_{ADAR1}$-bound Z-DNAs was detected in the two phosphates at the N2pN3 and N4pN5 phosphodiester steps, respectively. These results were expected because the $Z_I$ conformation of the GpC

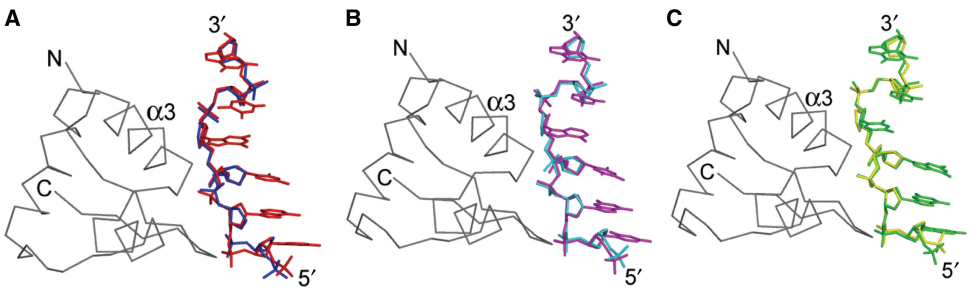


**Figure 4.** Structural comparisons of free- and $hZ\alpha_{ADAR1}$-bound Z-DNAs. Superposition of the free d(CACGTG)$_2$ (blue) with Zα-bound d(CACGTG)$_2$ (red) **(A)**, the free d($^{m5}$CGTA$^{m5}$CG)$_2$ (cyan) with Zα-bound d(CGTACG)$_2$ (magenta) **(B)** and the free d($^{m5}$CGGC$^{m5}$CG)$_2$ (yellow) with Zα-bound d(CGGCCG)$_2$ (green) **(C)**. For the structural overlap, 63 DNA backbone atoms of chain F of each $Z\alpha_{ADAR1}$-bound Z-DNA and chain A of each free Z-DNA were used. The Cα trace of $hZ\alpha_{ADAR1}$ in each complex is drawn in gray. The N- and C- termini, recognition helix (α3) of $hZ\alpha_{ADAR1}$ and the 5′ and 3′ ends of the DNA are labeled.

phosphodiester step is preferred when Z-DNA forms a complex with $hZ\alpha_{ADAR1}$ due to the specific interaction between the phosphate groups of DNA and the charged residues of Zα, whereas protein-free Z-DNA can have two alternative conformations, $Z_I$ and $Z_{II}$ (14,31). The difference between two corresponding atoms of the compared Z-DNAs is >2.5 Å when the protein-free Z-DNA is in the $Z_{II}$ conformation, and it is <0.5 Å when the protein-free Z-DNA is in the $Z_I$ conformation (Figure 5A). It is interesting that the current structures revealed that $hZ\alpha_{ADAR1}$-bound Z-DNAs do not absolutely have the $Z_I$ conformation. Because ionic interactions between the phosphates at the N5 position and the charged residues, Lys170 and Arg174, are absent or weak in $d(CGTACG)_2$ and $d(CGGCCG)_2$, the phosphodiester conformation of the N4pN5 step is not a typical $Z_I$ conformation (Figures 1B, 5C and D). As a result, in both Z-DNAs, the phosphates at the N5 position still show structural deviation from that of $d(CGCGCG)_2$ in $hZ\alpha_{ADAR1}$-bound structures (sky blue lines in Figure 5C and D). Specifically, the N4pN5 phosphodiester step of $d(CGGCCG)_2$ adopts an intermediate conformation between $Z_I$ and $Z_{II}$ in the Zα/DNA complex.

It is well known that there is no significant sequence-dependent conformational alteration in the Z-DNA backbone structure, although base-packing parameters vary in a sequence-dependent manner (2,3,8,30–32). However, when corresponding atoms of each DNA pair are compared, differences near all of the phosphate groups are notable although the extent varies depending on their positions (Figure 5B–D). However, the differences near the phosphate groups of each DNA pair are reduced when the DNAs are bound to $hZ\alpha_{ADAR1}$ (Supplementary Table 3 and Figure 5) except in the case of N3pN4 in comparing $d(CGCGCG)_2$ and $d(CGGCCG)_2$. The reduction in structural alteration of the phosphate groups of Zα-bound DNA is probably due to structural restraint enforced by Zα binding. These results strongly support the idea that the preformed binding pocket of Zα functions as a mold in recruiting various Z-DNAs, and it freezes the backbone conformations of Z-DNAs. As a result, structural variations of Z-DNAs that are caused mostly by conformational heterogeneity near the phosphate groups are reduced upon their binding into the Zα cleft.

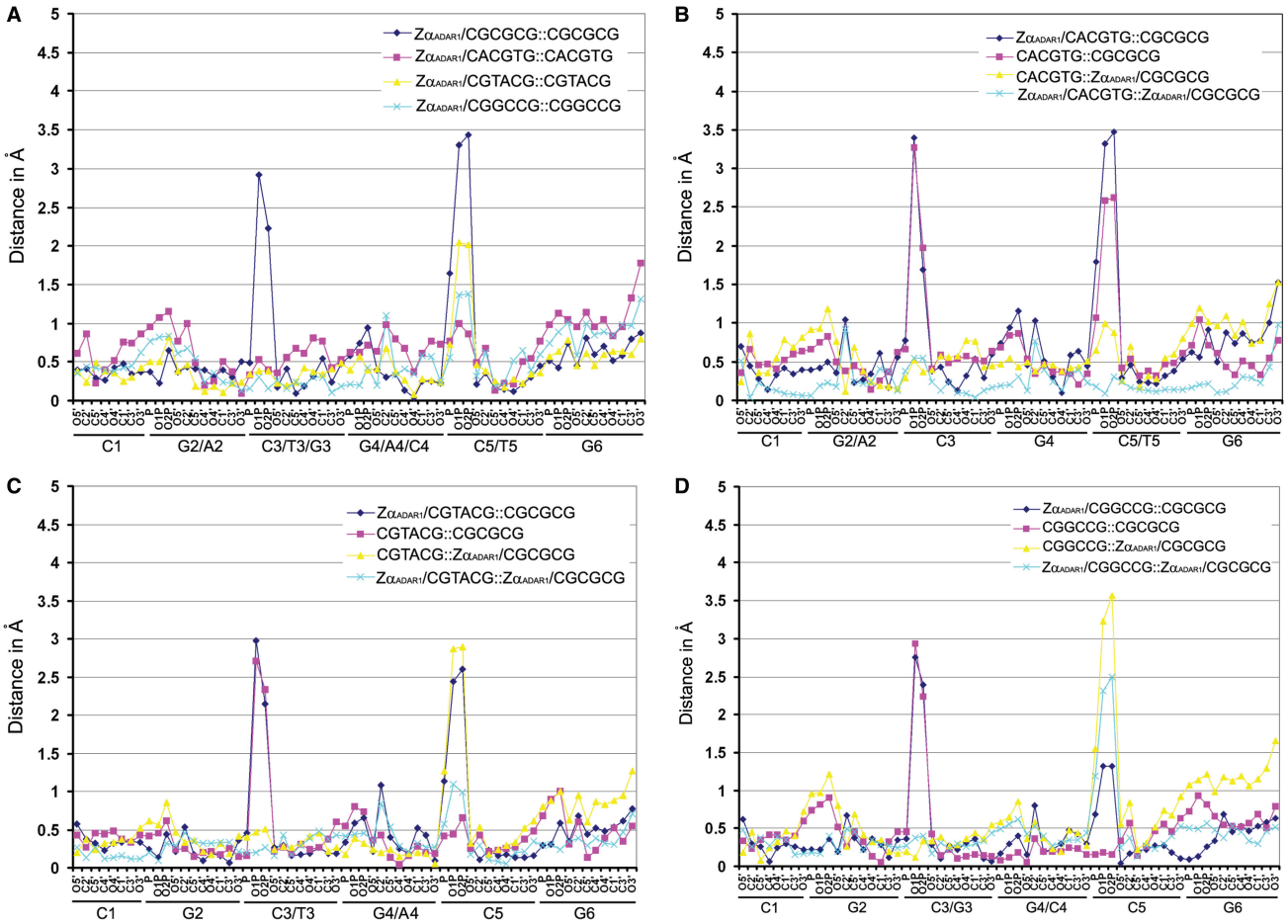


**Figure 5.** Structural deviation of each DNA backbone atom, comparing free Z-DNA and $Z\alpha_{ADAR1}$-bound Z-DNA in both non-CG-repeat and CG-repeat Z-DNAs. The distances in angstrom between two backbone atoms in the same position of each paired DNA were plotted against the DNA backbone atoms. The structures were compared in the same way as shown in Figure 4. (**A**) The backbone atoms of Z-DNAs with CGCGCG, CACGTG, CGTACG and CGGCCG sequences in free- and $hZ\alpha_{ADAR1}$-bound forms were compared. (**B**) CACGTG, (**C**) CGTACG and (**D**) CGGCCG DNAs in both free and $hZ\alpha_{ADAR1}$-bound forms were compared with free- and $hZ\alpha_{ADAR1}$-bound CG-repeat Z-DNA (CGCGCG).

Z-DNAs containing non-APP sequences or A-T base pair(s) have intrinsic structural instability and an increased solvent-exposed surface since base pairs are buckled out of the base-pair plane and protrude into the major groove (3,8,30,31). These structural features are still observed in $hZ\alpha_{ADAR1}$-bound Z-DNAs (Supplementary Table 2). For example, $d(CGTACG)_2$ and $d(CGGCCG)_2$ reveal a huge buckle and decreased stacking interactions, whether bound to Zα or not (Figure 4 and Supplementary Table 2). Therefore, it is thought that the base packing pattern and structural instability of non-CG-repeat Z-DNAs are always maintained, whereas the phosphate backbone conformations are restrained by the Z-DNA-binding pocket of $hZ\alpha_{ADAR1}$. Overall, additional structural instability of the Z conformation caused by non-CG-repeat sequences did not significantly affect the maintenance of the Z confirmation of DNAs in their complexes with Zα. These results suggest that more energy is gained upon Zα binding than is needed to overcome the structural instability of non-CG-repeat Z-DNAs.

## CONCLUSIONS

Our structural data strongly support the idea that $hZ\alpha_{ADAR1}$ recognizes Z-DNA in a conformation-specific manner, but not in a sequence-specific manner. The structures of $hZ\alpha_{ADAR1}$ complexed with three different non-CG-repeat double-stranded Z-DNAs (with non-APP sequences or A-T base pairs) reveal that $hZ\alpha_{ADAR1}$ binds and stabilizes the Z conformation of DNA via a similar binding mode to that of the $Z\alpha/d(CGCGCG)_2$ complex, regardless of sequence context. Most notably, the CH–π interaction was observed between Tyr177 and the cytosine in the *syn* conformation. While structures of phosphate backbones in free and $hZ\alpha_{ADAR1}$-bound states display some variation near each phosphate group, $hZ\alpha_{ADAR1}$ does not exert a profound effect on the Z-DNA base pair parameters induced by incorporation of non-APP and A-T base pairs. Irrespective of the heterogeneity in sequence and structure of non-CG-repeat Z-DNAs, Zα recognizes the DNA backbone atoms and imposes structural restraint through the core residues located on the DNA-binding surface. It has been suggested that different Z-DNA-binding proteins can bind to chromosomal or foreign DNAs and take part in essential biological processes. In this context, the Zα domain is expected to recognize and stabilize stretches of DNA in the Z-DNA conformation. Our results provide a molecular basis for understanding how the Zα proteins recognize and stabilize Z-DNAs in various sequence contexts.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## FUNDING

The Korea Science and Engineering Foundation (KOSEF) through the National Research Laboratory Program funded by the Ministry of Education, Science and Technology (NRL-2006-02287). Funding for open access charges: the National Laboratory Program grant of the Korean government (MEST).

*Conflict of interest statement*. None declared.

## REFERENCES

1. Rich,A. (1993) DNA comes in many forms. *Gene*, **135**, 99–109.
2. Wang,A.H., Quigley,G.J., Kolpak,F.J., Crawford,J.L., van Boom,J.H., van der Marel,G. and Rich,A. (1979) Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature*, **282**, 680–686.
3. Ho,P.S. and Mooers,B.H.M. (1997) Z-DNA crystallography. *Biopolymers*, **44**, 65–90.
4. Rich,A., Nordheim,A. and Wang,A.H. (1984) The chemistry and biology of left-handed Z-DNA. *Annu. Rev. Biochem.*, **53**, 791–846.
5. Ho,P.S. (1994) The non-B-DNA structure of d(CA/TG)n does not differ from that of Z-DNA. *Proc. Natl Acad. Sci. USA*, **91**, 9549–9553.
6. Wang,A.H., Hakoshima,T., van der Marel,G., van Boom,J.H. and Rich,A. (1984) AT base pairs are less stable than GC base pairs in Z-DNA: the crystal structure of d(m5CGTAm5CG). *Cell*, **37**, 321–331.
7. Schroth,G.P., Kagawa,T.F. and Ho,P.S. (1993) Structure and thermodynamics of nonalternating C.G base pairs in Z-DNA: the 1.3-A crystal structure of the asymmetric hexanucleotide $d(^{m5}CGGG^{m5}CG).d(^{m5}CGCC^{m5}CG)$. *Biochemistry*, **32**, 13381–13392.
8. Eichman,B.F., Schroth,G.P., Basham,B.E. and Ho,P.S. (1999) The intrinsic structure and stability of out-of-alternation base pairs in Z-DNA. *Nucleic Acids Res.*, **27**, 543–550.
9. Herbert,A., Alfken,J., Kim,Y.G., Mian,I.S., Nishikura,K. and Rich,A. (1997) A Z-DNA binding domain present in the human editing enzyme, double-stranded RNA adenosine deaminase. *Proc. Natl Acad. Sci. USA*, **94**, 8421–8426.
10. Schwartz,T., Behlke,J., Lowenhaupt,K., Heinemann,U. and Rich,A. (2001) Structure of the DLM-1-Z-DNA complex reveals a conserved family of Z-DNA-binding proteins. *Nat. Struct. Biol.*, **8**, 761–765.
11. Ha,S.C., Lokanath,N.K., Quyen,D.V., Wu,C.A., Lowenhaupt,K., Rich,A., Kim,Y.G. and Kim,K.K. (2004) A poxvirus protein forms a complex with left-handed Z-DNA: crystal structure of a Yatapoxvirus Zalpha bound to DNA. *Proc. Natl Acad. Sci. USA*, **101**, 14367–14372.
12. Hu,C.Y., Zhang,Y.B., Huang,G.P., Zhang,Q.Y. and Gui,J.F. (2004) Molecular cloning and characterisation of a fish PKR-like gene from cultured CAB cells induced by UV-inactivated virus. *Fish Shellfish Immunol.*, **17**, 353–366.
13. Rothenburg,S., Deigendesch,N., Dittmar,K., Koch-Nolte,F., Haag,F., Lowenhaupt,K and Rich,A. (2005) A PKR-like eukaryotic initiation factor 2alpha kinase from zebrafish contains Z-DNA binding domains instead of dsRNA binding domains. *Proc. Natl Acad. Sci. USA*, **102**, 1602–1607.
14. Schwartz,T., Rould,M.A., Lowenhaupt,K., Herbert,A. and Rich,A. (1999) Crystal structure of the Zalpha domain of the human editing enzyme ADAR1 bound to left-handed Z-DNA. *Science*, **284**, 1841–1845.
15. Ha,S.C., Lowenhaupt,K., Rich,A., Kim,Y.G. and Kim,K.K. (2005) Crystal structure of a junction between B-DNA and Z-DNA reveals two extruded bases. *Nature*, **437**, 1183–1186.
16. Ha,S.C., Kim,D., Hwang,H.Y., Rich,A., Kim,Y.G. and Kim,K.K. (2009) The crystal structure of the second Z-DNA binding domain of human DAI (ZBP1) in complex with Z-DNA reveals an unusual binding mode to Z-DNA. *Proc. Natl Acad. Sci. USA*, doi:10.1073/pnas.0810463106.
17. Liu,H., Mulholland,N., Fu,H. and Zhao,K. (2006) Cooperative activity of BRG1 and Z-DNA formation in chromatin remodeling. *Mol. Cell. Biol.*, **26**, 2550–2559.
18. Wang,G. and Vasquez,K.M. (2007) Z-DNA, an active element in the genome. *Front. Biosci.*, **12**, 4424–4438.
19. Kwon,J.A. and Rich,A. (2005) Biological function of the vaccinia virus Z-DNA-binding protein E3L: gene transactivation and

antiapoptotic activity in HeLa cells. *Proc Natl Acad. Sci. USA*, **102**, 12759–12764.
20. Takaoka,A., Wang,Z., Choi,M.K., Yanai,H., Negishi,H., Ban,T., Lu,Y., Miyagishi,M., Kodama,T., Honda,K. *et al.* (2007) DAI (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response. *Nature*, **448**, 501–505.
21. Schwartz,T., Lowenhaupt,K., Kim,Y.G., Li,L., Brown,B.A. II, Herbert,A. and Rich,A. (1999) Proteolytic dissection of Zab, the Z-DNA-binding domain of human ADAR1. *J. Biol. Chem.*, **274**, 2899–2906.
22. Schwartz,T., Shafer,K., Lowenhaupt,K., Hanlon,E., Herbert,A. and Rich,A. (1999) Crystallization and preliminary studies of the DNA-binding domain Zα from ADAR1 complexed to left-handed DNA. *Acta Crystallogr. D Biol. Crystallogr.*, **55**, 1362–1364.
23. Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307–326.
24. Brunger,A.T., Adams,P.D., Clore,G.M., DeLano,W.L., Gros,P., Grosse-Kunstleve,R.W., Jiang,J.S., Kuszewski,J., Nilges,M., Pannu,N.S. *et al.* (1998) Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.*, **54**, 905–921.
25. Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A*, **47**, 110–119.
26. Merritt,E.A. (1994) A program for photorealistic molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **50**, 869–873.
27. Kraulis,P. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.
28. Collaborative Computational Project, Number 4. (1994) The CCP4 suite: Programs for protein crystallography. *Acta Cryst.*, **D50**, 760–763.
29. Wang,A.H., Gessner,R.V., van der Marel,G.A., van Boom,J.H. and Rich,A. (1985) Crystal structure of Z-DNA without an alternating purine-pyrimidine sequence. *Proc. Natl Acad. Sci. USA*, **82**, 3611–3615.
30. Narayana,N., Shamala,N., Ganesh,K.N. and Viswamitra,M.A. (2006) Interaction between the Z-type DNA duplex and 1,3-propanediamine: crystal structure of $d(CACGTG)_2$ at 1.2 Å resolution. *Biochemistry*, **45**, 1200–1211.
31. Wang,A.J., Quigley,G.J., Kolpak,F.J., van der Marel,G., van Boom,J.H. and Rich,A. (1981) Left-handed double helical DNA: variations in the backbone conformation. *Science*, **211**, 171–176.
32. Sadasivan,C. and Gautham,N. (1995) Sequence-dependent micro-heterogeneity of Z-DNA: the crystal and molecular structures of d(CACGCG):d(CGCGTG) and d(CGCACG):d(CGTGCG). *J. Mol. Biol.*, **248**, 918–930.