Recognition of non-CpG repeats in Alu and ribosomal RNAs by the Z-RNA binding domain of ADAR1 induces A-Z junctions

Parker J. Nichols, Shaun Bevers, Morkos Henen, Jeffrey S. Kieft, Quentin Vicens & Beat Vögeli

Adenosine-to-inosine (A-to-I) editing of eukaryotic cellular RNAs is essential for protection against auto-immune disorders. Editing is carried out by ADAR1, whose innate immune response-specific cytoplasmic isoform possesses a Z-DNA binding domain (Zα) of unknown function. Zα also binds to CpG repeats in RNA, which are a hallmark of Z-RNA formation. Unexpectedly, Zα has been predicted — and in some cases even shown — to bind to specific regions within mRNA and rRNA devoid of such repeats. Here, we use NMR, circular dichroism, and other biophysical approaches to demonstrate and characterize the binding of Zα to mRNA and rRNA fragments. Our results reveal a broad range of RNA sequences that bind to Zα and adopt Z-RNA conformations. Binding is accompanied by destabilization of neighboring A-form regions which is similar in character to what has been observed for B-Z-DNA junctions. The binding of Zα to non-CpG sequences is specific, cooperative and occurs with an affinity in the low micromolar range. This work allows us to propose a model for how Zα could influence the RNA binding specificity of ADAR1.
Distinguishing between self and non-self RNA is a hallmark of the innate immune system. In humans, self RNAs are edited by an adenosine deaminase that acts on RNA (ADAR1), which converts some adenosines to inosines. Because inosines tend to disrupt double-stranded regions, such RNAs are not recognized by PKR, MDA5, and RIG-I, which are enzymes that trigger an immune response through binding to foreign and therefore unedited double-stranded regions.

ADAR1 is constitutively expressed in most cells as a stable p110 isoform localized in the nucleus. Upon invasion by a pathogen, the cell launches an interferon (IFN) response, resulting in the expression of a longer p150 isoform, which contributes to resisting the infection by editing self RNAs in the cytoplasm. ADAR1p150 catalyzes A-to-I editing within Alu repeats within the human genome. Alu elements recently have been shown to be one of the primary activators of IFN through a network of interactions involving other amino acids (notably Lys169, Asn173, and Trp195). This activation indicates that many dinucleotide sequences can adopt a Z-conformation, as observed in RNA for an even broader range of sequences (Supplementary Fig. 1). Although non-CpG sequences within DNA bind to Zα, the resulting complexes are less stable compared to CpG repeats. Intriguingly, the adoption of Z-conformations within the context of larger dsDNA stretches requires the formation of B-Z junctions, where the B-DNA regions flanking the Z-forming sequence become destabilized by Zα binding, and nucleotides adjacent to the Z-DNA sequence flip out in order to create continuous base-stacking between the B- and Z-DNA.

In such conformations, four Zα molecules are bound to an eight base-pair stretch, indicating that Zα can bind to a variable number of base pairs depending on the context. While the formation of equivalent A-Z junctions in RNA has been shown to be possible by fluorescence studies, our knowledge of the sequence preferences and contexts for the formation of such conformations upon Zα binding is limited. Within RNA, some of the sequence combinations (especially UpG and UpA repeats) have been predicted to shift to Z-RNA within the heavily edited Alu foldback structures. In further support of a broad recognition of RNA sequences by Zα, pull-down assays revealed that Zα binds to ribosomal stem-loops sometimes devoid of CpG repeats. This broader RNA sequence specificity is in support of the Zα-mediated increase in ADAR1’s specificity and activity proposed by Rich and could also explain the recently described surge in mouse anti-viral response upon sensing endogenous retroviral elements by the Zα domain of Z-DNA-binding protein.

We hypothesize that the >10,000 editing sites in the human transcriptome are associated with widespread Z-RNA formation as CpG as well as non-CpG sequences within Alu foldbacks. To begin to address this possibility, we applied NMR, circular dichroism (CD), isothermal titration calorimetry (ITC), and analytical ultracentrifugation (AUC) to characterize the binding of Zα to RNA fragments with various frequencies of CpG repeats. We demonstrate binding of Zα at specific sites on a fragment from an Alu foldback rich in UpG steps (AluSu1J)=AluSu11J) and on hairpins from bacterial and human ribosomal RNAs that contain various types of YpR repeats (Y = pyrimidine, R = purine)’.

Our NMR results show that binding results in the formation of A-Z junctions in these RNAs and that RNAs with regions of low helical stability adjacent to a Z-forming sequence. The crystal structure of Zα in complex with a (CpG)₃ RNA duplex showed that binding specificity is achieved through a critical tyrosine residue (Tyr177), which makes a side chain interaction with the syn purine in the CpG Z-step. Tyr177 is positioned to interact with Z-RNA through a network of interactions involving other amino acids (notably Lys169, Asn173, and Trp195), which cause a 60% reduction in A-to-I editing levels when they are mutated. This reduction in editing leads to the Aicardi-Goutières syndrome, a neurodevelopmental disorder characterized by hyperactive immune responses. Zα is clearly important for the integrity of the editing process, but how does that relate to its ability to bind Z-DNA and/or Z-RNA?

Although most studies of Zα were done with CpG repeats, Zα can also recognize TpA, Cpa, CpG, and TpG in DNA. This observation indicates that many dinucleotide sequences can adopt a Z-conformation, as observed in RNA for an even broader range of sequences (Supplementary Fig. 1). Although non-CpG sequences within DNA bind to Zα, the resulting complexes are less stable compared to CpG repeats. Intriguingly, the adoption of Z-conformations within the context of larger dsDNA stretches requires the formation of B-Z junctions, where the B-DNA regions flanking the Z-forming sequence become destabilized by Zα binding, and nucleotides adjacent to the Z-DNA sequence flip out in order to create continuous base-stacking between the B- and Z-DNA.

In such conformations, four Zα molecules are bound to an eight base-pair stretch, indicating that Zα can bind to a variable number of base pairs depending on the context. While the formation of equivalent A-Z junctions in RNA has been shown to be possible by fluorescence studies, our knowledge of the sequence preferences and contexts for the formation of such conformations upon Zα binding is limited. Within RNA, some of the sequence combinations (especially UpG and UpA repeats) have been predicted to shift to Z-RNA within the heavily edited Alu foldback structures. In further support of a broad recognition of RNA sequences by Zα, pull-down assays revealed that Zα binds to ribosomal stem-loops sometimes devoid of CpG repeats. This broader RNA sequence specificity is in support of the Zα-mediated increase in ADAR1’s specificity and activity proposed by Rich and could also explain the recently described surge in mouse anti-viral response upon sensing endogenous retroviral elements by the Zα domain of Z-DNA-binding protein.

We hypothesize that the >10,000 editing sites in the human transcriptome are associated with widespread Z-RNA formation as CpG as well as non-CpG sequences within Alu foldbacks. To begin to address this possibility, we applied NMR, circular dichroism (CD), isothermal titration calorimetry (ITC), and analytical ultracentrifugation (AUC) to characterize the binding of Zα to RNA fragments with various frequencies of CpG repeats. We demonstrate binding of Zα at specific sites on a fragment from an Alu foldback rich in UpG steps (AluSu1J)=AluSu11J) and on hairpins from bacterial and human ribosomal RNAs that contain various types of YpR repeats (Y = pyrimidine, R = purine).

Our NMR results show that binding results in the formation of A-Z junctions in these RNAs and that RNAs with regions of low helical stability adjacent to a Z-forming sequence.
are better Z-RNA adopters. For the Alu fragment and for the E. coli h43 hairpin in particular, we further determine that two Za molecules molecules four to five base pairs, and that binding is cooperative, with $K_D$ values of 1–9 µM and 40−110 nM. In sum, when binding to AluSx1Jo, Za adopts a conformation that has characteristics of both a B-Z DNA junction and the (CpG)$_3$ RNA repeat. Overall, our work offers a rationale for earlier observations that ADAR1 binds in multiple steps to double-stranded RNA regions often flanked by regions rich in non-Watson−Crick pairs $^{39,40}$. This study also provides a framework for better understanding widespread and pervasive editing in cells.

Results

Za induces A-Z junctions within an AluSx1Jo foldback fragment and ribosomal hairpin h43. The core model systems of our study are an element from the AluSx1Jo foldback within the Cathepsin S (CTSS) gene on human chromosome 1, as well as hairpin 43 from the small subunit ribosome comprising hairpin h43 (Fig. 2a). This region of AluSx1Jo had been predicted to adopt a Z-conformation $^{32}$, and h43 had been reported to bind to Za in vitro $^{37}$. The AluSx1Jo region comprises predominantly UpG repeats, but also one CpG, one CpA, and one UpA dinucleotide steps. Hairpin h43 contains a symmetrical (or ‘canonical’) CpG step together with several UpG and CpA steps. Based on current knowledge of Za recognition, it is not immediately apparent how these two RNAs could adopt the Z-RNA conformation required for Za binding.

Using chemically synthesized RNA fragments for AluSx1Jo, h43, and the control sequences (CpG)$_3$ and two A-Z junctions, we first carried out CD experiments, which have traditionally been used to monitor transitions to the Z-conformation in DNA and RNA $^{22,41}$. Upon titrating increasing amounts of Za to the control (CpG)$_3$ fragment (from 12:1 to 1:6 RNA:Za), we observed the characteristic decrease in the A-form peak at 266 nm, the appearance of a peak at 285 nm, and the flip from negative to positive at 295 nm (Fig. 2b, full titration shown in Supplementary Fig. 2). Together, these three changes were indicative of a conversion to the Z-conformation $^{32}$, and are unlikely to be due to a conformational change in Za upon binding RNA as our $^{15}$N-HSQC titration indicates binding only results in minor rearrangements to the RNA binding residues of Za (Supplemental Figs. S13 and S14). However, our selected test RNAs had diverse sequence context and likely did not fully adopt the Z-conformation. Performing the same experiment with RNAs which form A-Z junctions (A-Z 1 is the RNA version of a previously characterized B-Z junction $^{31}$, and A-Z 2 has been shown by fluorescence studies to adopt an A-Z junction $^{35}$),

Fig. 2 Za induces a partial Z-conformation in AluSx1Jo and h43. a Location of AluSx1Jo and h43 on secondary structure diagrams of the CTSS gene (Chromosome 1) from H. sapiens and the small subunit ribosomal RNA from E. coli. For h43, the first three G=C base pairs were engineered for added stability. In this and subsequent figures, CpG and YpR steps are shown in dark (CpG) and light (non-CpG YpR) shades of purple/orange colors. b CD spectra of the AluSx1Jo and h43 RNAs in absence of protein (black) and at the 1:6 RNA:Za ratio (red), at which binding is saturated. Controls: 6 M NaClO$_4$ (dotted line),ionic condition that promotes Z-RNA formation of CpG repeats $^{57}$, A-Z junction $^{35}$ (second to right-most panel), which is an RNA that has a (CpG)$_3$ sequence followed by an A-RNA forming sequence, and (CpG)$_2$ (right-most panel), which is fully converted to Z-RNA at a 1:2 RNA:Za ratio. c Same as b, but with Za$_{77A}$ instead of wild-type Za. d E$_Z$ scores quantifying the extent of Z-conformation for the following fragments: (ApU)$_6$ (supplemental Fig. 5). An error of 0.1 was determined to be appropriate for the calculated E$_Z$ scores by taking into account the difference between the (CpG)$_3$ and (CpG)$_6$ RNAs (which theoretically both have an E$_Z$ score of 1) and the difference in the E$_Z$ score between repeat measurements on h43 E.coli. All other E$_Z$ scores were determined from one CD measurement.

NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-021-21039-0 | www.nature.com/naturecommunications
we observed instead a reduction of the peak at 266 nm and smaller-in-magnitude increases to the molar ellipticities values at 285 and 295 nm (Fig. 2b, Supplementary Fig. S2). Therefore, the CD spectra of A-Z junctions is population-weighted depending on the amount of Z-conformation adopted by a particular RNA.

We also carried out CD experiments with a Zα construct where Tyr177 was mutated to alanine (ZαTyr177Ala)—resulting in impaired Z-RNA binding capabilities37,42—which led to smaller absolute magnitudes of the shifts at these three wavelengths for the (CpG)3 RNA and resulted in a growth of the molar ellipticity at 266 nm for the A-Z junction (Fig. 2c). This growth likely occurred because ZαTyr177Ala can still bind to RNA, without causing the conformation switch to the Z-conformation (discussed later and shown by NMR in Supplementary Fig. S9).

A change in the CD spectra similar to that for the control A-Z junction was observed for AluSx1Jo and h43 as the Zα concentration was increased, indicative of the adoption of A-Z junctions in these RNAs (Fig. 2b). Measuring CD on stem-loops as in h43 also resulted in a decreased magnitude compared to measurement with only stems, because the loop restricts the conformationally accessible space. We verified this effect by comparing the CD curves for the AluSx1Jo duplex (Fig. 2b) and an engineered stem-loop version (Supplementary Fig. 3). Z-RNA adoption by h43 and AluSx1Jo was further supported by our negative control using ZαTyr177Ala which showed again an increase in the molar ellipticity at 266 nm and an entirely different behavior over the course of the titration (Fig. 2c and Supplemental Fig. S2).

To parametrize the extent of Z-RNA formation in AluSx1Jo and h43, we derived a Z-conformation score from the CD spectra. Specifically, we calculated the extent of Z-RNA formation (EZ) as the growth in molar ellipticity at 285 and 295 nm and decrease in molar ellipticity at 266 nm of the fully saturated RNA, normalized to the molar ellipticity at 266 nm of the free form (calculated using Eq. 1 in “Methods” section). When the RNA is fully converted to a Z-conformation, as with (CpG)3 and (CpG)6α, EZ ∼ 1 (Fig. 2d). Conversely, if the RNA remains in the A-form, EZ ∼ 0 (for a (ApuU)6 G16C10 duplex, as well as for cUUCGg and cGAAAg hairpins; Fig. 2d). A-Z 1 and A-Z 2 had values of 0.2 and 0.19, respectively, again, reflecting the fact that these RNA do not fully adopt the Z-conformation (Fig. 2d). For AluSx1Jo and h43, EZ ∼ 0.3 and ~0.2, respectively, confirming further that these RNAs are very similar to the control A-Z junctions (Fig. 2d). The smaller changes in the CD spectra detected when using ZαTyr177Ala instead of Zα (Fig. 2b, c) resulted in a decrease in the EZ score that was threefold to fourfold larger for AluSx1Jo, h43, A-Z 1, and A-Z 2 than for the (CpG)3 control (Fig. 2e, calculating using Eq. 2 in “Methods” section). The EZ score therefore represents a suitable standard for evaluating the extent of Z-RNA formation for any sequence.

Zα saturates the double-stranded region of AluSx1Jo, h43, and other ribosomal hairpins. We tested additional RNA fragments for their propensity to locally adopt a Z-RNA conformation upon addition of Zα. Using CD, we characterized the binding of Zα to chemically synthesized fragments corresponding to helices h41, h30, H25, and H66 (large subunit, LSU) from E. coli ribosomes, and h41, h30, and H66 from H. sapiens ribosomes, which had all been shown to bind to Zα in ribosome pull-down experiments37. We calculated the EZ score of each tested fragment (Supplementary Figs. 3 and 4a; Supplementary Table 1).

For these RNAs, the fully saturated spectra more closely resembled A-Z 1 and A-Z 2 than they did (CpG)3. EZ scores ranged from ~0.1 to 0.47 (Supplementary Table 1), indicating that all tested fragments adopted A-Z junctions, but to various extents. Most fragments had EZ values in the 0.1–0.2 range, unless they formed extended duplexes at the high concentrations required by the CD measurements, as was the case for H66 from H. sapiens (0.3) and for h41 from E. coli (0.47), which we monitored by AUC (see next paragraph). These duplexes were nonetheless interesting, as their predicted secondary structure (Supplementary Fig. 4a) suggested that in addition to recognizing Ypz steps—which of these fragments had a high concentration—Zα may favor regions rich in non-canonical base pairs. In support of that possibility, the EZ score correlated well to the number of predicted non-Watson–Crick pairs (R = 0.78) and to RNA length (R = 0.75) (Supplementary Fig. 4b).

To determine the stoichiometric ratio of RNA and Zα within a complex, we conducted AUC experiments with Zα:RNA at a ratio of 6:1. We tested the following fragments: (CpG)3, cUUCGg hairpin, h43 (E. coli), H25 (E. coli), H66 (H. sapiens), H66 (E. coli), h41 (H. sapiens and E. coli), and the AluSx1Jo RNA. Specifically, we compared the fitted molecular weights from the AUC measurements with the theoretical molecular weights from increasing stoichiometric ratios of Zα:RNA. For each fragment except the negative cUUCGg control (major peak (78%) at 7.8 kDa, corresponding to unbound Zα), all potential Zα binding sites according to RNA length were occupied (Supplementary Fig. 5). For example, h43 bound to four Zα molecules (Supplementary Fig. 5), the maximum theoretically possible as its stem comprises 10 base pairs and each Zα spans ~4–5 base pairs25,31. However, while 96% of the complexes with h43 had a 4:1 ratio of Zα:RNA, the percentage of the saturated complex varied for the other fragments within 54–87% (Supplementary Fig. 5). In agreement with the CD data, this varying saturation further suggests variable A-Z junction adoption within these RNAs.

We also measured h43 at 1:2 and 1:4 stoichiometric ratios of RNA:Zα, which gave molecular weights of 21.3 kDa (at 2:1 [RNA]:[Zα]) and 27.9 kDa (at 4:1 [RNA]:[Zα]), indicating 1:2, 1:3 h43:Zα complexes, respectively (Supplemental Fig. S6). This suggests that while there are four binding sites for Zα on h43, the second pair of sites are less favorable than the first, as at a 1:4 stoichiometric ratio of h43:Zα, there is a 1:3 complex and the 1:4 complex only becomes stable by adding additional Zα proteins (at a 1:6 molar ratio of h43:Zα). This result highlights that while Zα is capable of binding a wide variety of RNA sequences, it still retains a sequence specificity that is dependent upon the ability of a particular sequence to adopt the Z-conformation.

A-Z junction adoption in AluSx1Jo, H66, h43, and H25 involves destabilization of adjacent A-RNA. The observations that Zα binds to a variety of Alu and ribosomal sequences and that binding is correlated to the number of non-canonical base pairs raise the question of its sequence specificity. To investigate the context for Z-RNA formation at the nucleotide level upon Zα binding, we used 1D 1H-NMR because signatures of Z-DNA/RNA formation determined from previous work on (CpG)n RNA/DNA repeats23,43 and B-Z junctions31,32 can be used as a reference for A-Z junction formation. Specifically, it was shown that Z-RNA formation within the (CpG)3 repeat mostly follows a two-step process. First, one Zα binds to one side of the repeat, which begins to convert the RNA to the Z-form. This event promotes binding of a second Zα molecule, which stabilizes the Z-conformation23,43–45. The conversion from the A-form to the Z-form in (CpG)3 DNA/RNA causes the imino proton of the guanine base within CpG repeats contacted by Zα to experience slow exchange between the chemical shift positions of the A- and Z-forms23,43. For B-Z DNA junctions, on the other hand, the
adoption of the Z-DNA in the Z-forming sequence is inhibited by neighboring B-DNA\textsuperscript{31,32}. Therefore, these regions must be destabilized by initial Z binding before the DNA can adopt the Z-conformation, which causes the imino protons in such regions to disappear into the noise\textsuperscript{31,32}. Following this initial destabilization, the Z-forming sequence is shifted to the Z-conformation and can either resemble the (CpG)\textsubscript{3} case with slow exchange of the syn purines from the B-Z junction, or chemical shift perturbations (CSPs) can be observed instead, depending on the sequence\textsuperscript{31,32}.

We followed Z-RNA formation within AluSx\textsubscript{1}Jo, h43, H66, H25, and A-Z 1 with increasing amounts of Z\textsubscript{α} (imino region of spectra in Fig. 3; 2D \textsuperscript{1}H-\textsuperscript{1}H NOESY assignments and full 1D imino titrations in Supplementary Figs. S7 and S8). Following the addition of Z\textsubscript{α} up to a 1:1 ratio, the imino proton resonances of U\textsubscript{14}, U\textsubscript{22}, U\textsubscript{16}, U\textsubscript{18}, G\textsubscript{17}, and G\textsubscript{9} of AluSx\textsubscript{1}Jo disappeared into the noise in a manner similar to what had been observed for the neighboring B-DNA in B-Z junctions\textsuperscript{31} and the A-Z 1 control, indicating those base pairs were destabilized (Fig. 3a, e). In addition, the imino protons in the Z-forming regions began to shift towards their bound positions concurrently with the disappearance of the imino protons in the destabilized regions, suggesting that the Z-formation and the destabilization steps were directly coupled. We observed the same behavior for the other three RNA constructs tested (Fig. 3b–d). We confirmed both effects were dependent upon Z-RNA formation as the reduction in the intensity of the same imino peaks when using Z\textsubscript{α} Tyr\textsubscript{177}Ala was stunted relative to the wild-type, and the peaks never fully disappeared into the noise (Supplementary Fig. 9). In addition, the CSPs went in different directions than with wild-type Z\textsubscript{α} (Supplementary Fig. 9).

At higher stoichiometric ratios of RNA:Z\textsubscript{α} (1:1, 1:2, 1:3, 1:4), we also observed CSPs and line broadening (indicating the increase in complex size) in the base pairs that were not
destabilized below 1:1 RNA:Zα. These results indicated loading of additional Za proteins onto the different RNAs, in agreement with our CD and AUC results (Fig. 2 and Supplementary Figs. S2–S5). Plotting of the CSPs onto the secondary structures of the tested RNAs allowed us to determine the Za binding sites (Fig. 3, green boxes). Similarly to what was observed for B-Z DNA junctions31 and A-Z 1, Za binds to specific regions within AluSx1Jo, H66 (extended duplex), and h43, usually having a high concentration of G–C base pairs and one or more (YpR) steps. Binding also coincides with destabilization of the adjacent base pairs (Fig. 3). This is consistent with the formation of A-Z junctions within these RNAs and is supported by our CD results, which showed only partial Z-formation.

For H25, Za appeared to target the lower portion of the stem which contained only two (YpR) steps and two G.A base pairs (Fig. 3d). However, the exact binding site could not be pinpointed as for the other fragments due to a lack of observable imino proton resonances for the G.A pairs at the temperature used for the titration. In addition, we cannot confirm whether the imino peak of G4 disappears into the noise due to Za-dependent destabilization or whether it simply coalesces with G28 (Supplementary Fig. S8). The observation that the base-paired U27 peak does not disappear supports the latter conclusion, suggesting that the binding site comprises nine base pairs (four Za binding sites) and would be in agreement with our AUC data (Supplementary Fig. S5). We did not observe any slow exchange between the A- and Z-form of any imino peaks in our RNA constructs as was the case for the (CpgG) DNA junctions31, with the exception of U11 and G21 of H25 (although we could not confirm that these were due to Z-RNA formation). This could be for a number of reasons, the most likely being that the formation of Z-RNA in RNAs with some sequence contexts is more dynamic and thus the conformations are exchanging on a faster timescale.

While the observed disappearance of imino peaks could be caused by line broadening due to intermediate exchange, the striking similarity to titrations done on the B-Z DNA junction31 suggests that this is not the case. In addition, we did not observe the re-emergence of these peaks at higher concentrations of Za which we would expect as the complex becomes more stable in the case of intermediate exchange.

Zα binds cooperatively to AluSx1Jo and h43 in the ~500 nM–2 μM range. Having determined that Za induces novel A-Z junctions in our RNA constructs, we sought to characterize how Za interacts with such elements. To characterize the binding of Za to AluSx1Jo and h43, we used ITC and NMR measurements. For AluSx1Jo titration into Za, these experiments showed an initial endothermic binding event up to a stoichiometric ratio of 2:1 of Za:AluSx1Jo, followed by a second exothermic event which continued until saturation (Fig. 5a). The first binding event likely reflected a combination between binding, destabilization of the neighboring A-form regions, and Z-RNA formation (an endothermic cooperative process) at the saturating levels of Za early in the titration. The second event captured the reshuffling of the populations from the 2:1 to 1:1 Za:AluSx1Jo complex as the RNA began to saturate Za. This model was supported by NMR measurements of the tumbling times (rotational correlation times, τcorr, calculated from Eq. 3 and the ratio of R2/R1 as described in “Methods” section) of Za with different AluSx1Jo concentrations. These data indicated that the complex size was maximum at 2:1 Za:AluSx1Jo, with a value of 12.6 ns, and decreased to 10.6 ns at 1:1 and finally 9.2 ns at 1:2 Za:AluSx1Jo (Fig. 5b, residue-specific relaxation rates shown in Supplementary Fig. 10). These tumbling times were consistent with a 2:1 Za:AluSx1Jo complex being formed at a 4:1 and 2:1 protein:RNA molar ratio, an intermediate complex between 2:1 and 1:1 at a 1:1 molar ratio, and a 1:1 complex at 1:2 protein:RNA. These results demonstrated cooperative binding, in agreement with the observed line broadening of the RNA imino peaks upon increasing concentrations of Za (Fig. 3). The maximum of two bound Za proteins was confirmed by AUC carried out at a 1:6 molar ratio of AluSx1Jo:Zα which showed that the majority of the RNA (85%) was in a complex of molecular weight 21.3 kDa (Fig. 5c) compared to the theoretical complex size of 22.5 kDa.

The ITC profile was best fit to a two-site binding model with a Kd of 1.14 ± 8.85 μM when the ratio of Za:AluSx1Jo is ≤2 and 37.6 ± 103.8 nM when ≥2 (Table 1), although the large fitting error and other ITC parameters attests to a more complex binding behavior. Upon globally fitting NMR CSPs (calculated using Eq. 6 from “Methods” section) for five binding-site residues (His159, Lys170, Glu171, Asn173, and Thr191) from a 15N-HSQC titration of Za and the AluSx1Jo RNA to a two-site binding model (Fig. 5d, individual per-residue fits shown in Supplementary Fig. 11), we obtained a Kd of 2.04 ± 1.73 μM. Together, these two independent measurements revealed that the affinity of Za for the AluSx1Jo RNA is in the mid-nanomolar to low-micromolar range.
ITC measurements of Zα titrated into h43 showed exothermic binding heat with injections up to a Zα:RNA molar ratio of 2:1, followed by decreasing binding heat up to a ratio of 4:1 (Fig. 5e). This suggested that two Zα molecules were required before h43 could convert to the Z-conformation, with up to four Zα domains bound in the final state, in agreement with the AUC data (Fig. 5f and Supplemental Fig. S6). Doing the reverse ITC experiment (injecting h43 into Zα) showed a similar profile of multiple exothermic binding events (Supplemental Fig. S12). The thermogram could be fit to a multisite model, from which K_d values were extrapolated as 538 ± 314 nM for the first site and 512 ± 1720 nM for the second. Together, our binding analysis revealed that Zα binds specifically and cooperatively to the AluSx1Jo and h43 E.coli RNAs with affinities similar to those of Zα binding to perfect (CpG)_n repeats (K_d of 241.5 ± 1300 nM for r(CpG)_3, Fig. 5g, Table 1; and K_d of 9 nM for (CpG)_6, according to bio-layer interferometry measurements in ref. 23), although with more complexity.

**Binding of Zα to AluSx1Jo resembles binding to (CpG)_3 but also to a B-Z DNA junction.** Previous NMR studies showed that Zα interacts with (CpG)_3 RNA and DNA repeats differently than B-Z DNA junctions 31,32. For (CpG)_n repeats, significant CSPs...
<0.2 ppm) to the backbone residues of Zα were observed for β1–α2 and β2–loop–β3 regions of Zα,32,43, which reflect the binding of two Zαs and conversion to the Z-conformation. For B-Z DNA junctions, Zα experiences less extreme CSPs (<0.15 ppm) and the disappearance of residues 173–177 due to intermediate exchange, which was proposed to reflect an “initial contact conformation” where Zα binds to the Z-adopting sequence and destabilizes the neighboring B-form regions32.

We employed NMR on 15N-labelled Zα to investigate how similar the binding site of AluSx1Jo on Zα is to that of (CpG)n repeats and to that of a B-Z DNA junction. The CSPs (calculated using Eq. 5 from “Methods” section) were non-linear and coupled with line broadening (Fig. 6a, full 15N-HSQC titration shown in Supplementary Fig. 13). These results indicated complex, multisite binding, consistent with our previous ITC, CD, AUC, and NMR results (Figs. 2–5). In addition, there were distinct differences in the peak positions between 2:1 and 1:1/1:2 Zα:AluSx1Jo for some of the binding interface residues (Fig. 6b), indicating differences in the backbone conformation of Zα between the state (one versus two Zα bound). This, in combination with our tumbling time analysis (Fig. 5b), suggests that at 2:1 Zα:RNA, two Zα molecules are bound to AluSx1Jo, which adopts a fully formed A-Z junction. At 1:1/1:2 Zα:RNA, however, Zα interacts with AluSx1Jo in the “initial contact conformation”, as was seen for B-Z DNA junctions32.

The largest CSPs measured from our 15N-HSQC titration of AluSx1Jo into Zα were mainly concentrated to the Z-RNA binding surface of Zα (amino acids 170–180, 190–194, green

---

**Table 1 Parameters extracted from isothermal titration calorimetry experiments.**

| Interaction (cell/syringe) | N (cell/syringe) | Kd (nM) | ΔH (kcal mol⁻¹) | TΔS (kcal mol⁻¹) | ΔG (kcal mol⁻¹) |
|---------------------------|------------------|---------|----------------|-----------------|----------------|
| Zα/(CpG)n               | 0.4 ± 0.0        | 241.5 ± 1300.0 | −6.0 ± 0.1    | 3.0             | −9.0           |
| h43/Zα                   | 0.4 ± 0.1/2.5 ± 0.0 | 538.0 ± 314.0/512.0 ± 1720.0 | −0.6 ± 0.2/−2.3 ± 0.1 | 7.9/6.3 | −8.6/−8.6 |
| Zα/AluSx1Jo              | 0.2 ± 0.0/0.3 ± 0.0 | 37.6 ± 103.8/1140.3 ± 8849.6 | −0.1 ± 0.2/5.9 ± 0.6 | 10.0/14.0 | −10.1/−8.1 |

---

Fig. 5 Zα binds to AluSx1Jo and h43 in the nanomolar to low-micromolar range. a ITC thermogram of AluSx1Jo titrated into Zα and fit to a two-site binding model. b Average (of the corresponding residue-specific values) longitudinal (R₁), and rotating-frame 15N NMR relaxation rates (R₁ρ), and the extracted effective overall correlation times (τcorr) are plotted versus the molar ratio of Zα:AluSx1Jo in the experiment (residue-specific R₁, R₁ρ, and τcorr are shown in Supplementary Fig. 9). Measured values are from one set of relaxation rate experiments. c Sedimentation coefficient distribution as obtained from analytical ultracentrifugation (AUC) with a 1:6 molar ratio of AluSx1Jo:Zα. The inset shows the raw data from the AUC run with the window position on the y-axis and individual scans over time going from left to right. d Global Kd fit of chemical shift perturbations (CSPs) of binding-site residues from the 15N-HSQC titration of AluSx1Jo into Zα, assuming a two-site binding model. Values are determined from one 15N-HSQC titration. e Isothermal calorimetry (ITC) indicating the multiple Zα domain binding events for E. coli h43. f Sedimentation coefficient distribution as obtained from analytical ultracentrifugation (AUC) with a 1:6 molar ratio of h43:Zα. g ITC thermogram and fit from titrating the (CpG)₃ RNA into Zα. All ITC parameters are given in Table 1. ITC thermograms are representatives of two measurements and AUC data for AluSx1Jo and h43 are determined from one measurement each.
color in Fig. 6c, Supplementary Fig. 14). The affected amino acids were similar to those observed when bound to the (CpG)₃ RNA (red, Fig. 6c) and the (CpG)₃ DNA (cyan, Fig. 6c). However, their CSP magnitude was more similar to that observed for amino acids in the complex with a B-Z DNA junction (purple, Fig. 6c).

In contrast to B-Z junctions (where residues 173–177 disappeared), only Tyr177 experiences intermediate exchange and line broadening beyond detection when bound to AluSx1Jo (Fig. 6a). This was confirmed by Chemical Exchange Saturation Transfer (CEST) measurements at 10:1 Zα:RNA, which showed that Tyr177 exhibited exchange on the micro–millisecond timescale with potentially multiple minor states, while Asn173, Arg174, Val175, and Ser178 showed little to no exchange (Supplementary Fig. 15).

Hence, Zα interacts with AluSx1Jo in a manner that appears to have characteristics of both a B-Z DNA junction and the (CpG)₃ RNA repeat.

Our analysis also revealed significant CSPs for Glu141 and His159. Perturbations in Glu141 were unexpected as the amino acid lies near the back side of Zα, adjacent to the binding helix. Glu141 may make specific interactions with A-Z junctions as this residue did not shift when bound to a B-Z DNA junction (purple, Fig. 6c).

Notably, His159 had previously been shown to undergo chemical shift changes upon binding to (CpG)₃ DNA but not RNA, possibly due to differences in the hydrogen bonding networks (23). Our observation of a significant CSP for His159, combined with CSPs at residues 190–194 that do not occur when bound to (CpG)₃ DNA, suggests that Zα may invoke a hybrid binding mechanism between Z-DNA and Z-RNA.

Fig. 6 The binding of Zα to AluSx1Jo resembles both binding to the (CpG)₃ RNA and a B-Z DNA junction. a Representative peaks and their perturbations from the Zα ¹⁵N-HSQC upon addition of AluSx1Jo. b Differences in peak positions between free Zα, 2:1, 1:1, and 1:2 Zα:AluSx1Jo. Red arrows help show the change in peak positions. c Residues undergoing chemical shift perturbation (CSP) and the CSP magnitudes (Supplementary Fig. 14a) from the ¹⁵N-HSQC titration are indicated on a surface plot of Zα (left, plotted on PDB: 2GXB25) and compared to CSPs from titrations of a B-Z DNA junction (purple, Fig. 6c), the (CpG)₃ RNA (cyan, Fig. 6c), and the (CpG)₃ DNA (right, PDB: 3F21 (41)).

The binding of Zα to AluSx1Jo is more dynamic than for (CpG)₃ RNA. Because we did not observe slow exchange between the A-form and Z-form peaks in our RNAs (only CSPs and line broadening), we wondered whether the dynamics of Z-RNA formation might be faster in our constructs relative to the (CpG)₃ RNA. To answer this question, we acquired Carr-Purcell-Meiboom-Gill (CPMG) measurements at 2:1 Zα:AluSx1Jo (two Zα residues symmetrically bound around a single Z-step) and compared them to previous CPMG measurements on the (CpG)₃ DNA and RNA also measured at a 2:1 ratio of Zα:RNA (41). CPMG profiles were globally fit (using Eq. 4 from “Methods” section) to give an exchange rate (kₑₓ) between the bound and unbound states of 634 s⁻¹ and major-state population of 96.2% (Fig. 7, Supplementary Fig. 16, Supplementary Table 2).

Global fits from solely the binding-site helix (residues 169–180, 191) gave a kₑₓ of 511 s⁻¹ and kₑₓ for binding-site residues alone of 762 s⁻¹, and were therefore roughly 200-fold higher than for the (CpG)₃ RNA (23). Thus, Z-RNA formation in our RNA constructs is more dynamic than for the canonical (CpG)₃ repeat, and this may explain why we did not observe slow exchange.
exchange between the A- and Z-form peaks in our imino titrations (Fig. 3).

Discussion

We have characterized various RNA fragments that bind to the Za domain of ADAR1, even though they do not exclusively comprise CpG repeats. We have found that these RNAs likely adopt A-Z junctions, evidenced by the similarity of our imino titrations with Za to those observed for B-Z DNA junctions. NMR characterization of the binding of Za to AluSx1Jo suggests that the mechanism for A-Z junction formation upon Za binding lies somewhere between that of binding to the canonical (CpG)₃ RNA and that of B-Z junction formation. In addition, our NMR results suggest that the initial interaction of Za with such RNAs involves the proposed “initial contact formation” where Za begins to destabilize the neighboring A-form regions, followed by the binding of an additional Za molecule and adoption of the Z-conformation (Fig. 8).

Our findings suggest that a wider variety of sequence contexts than previously assumed are able to adopt a Z-conformation, as had been shown for Z-DNA. This observation is in line with the adoption of Z-RNA-like steps by other sequences than CpG within a large variety of non-coding RNA. This work also offers experimental evidence to support previous proposals that sequences, such as UpG, adopt a Z-RNA conformation. Whether the Z-RNA structures adopted by these diverse sequences are similar or represent various types of Z-conformations (as proposed from FRET measurements for CpG repeats) will require structural determination of these complexes.

Even though the nucleotide sequences supporting A-Z junction arrangements are diverse, Z-RNA formation is still subject to a degree of specificity. We observed that Za targets specific regions with dsRNA that are adjacent to regions which are predicted to be less thermodynamically stable. It has been shown that B-DNA next to a Z-forming sequence in B-Z junctions can present a significant thermodynamic barrier for Z-DNA formation. It follows from our analysis that the stability of A-form regions adjacent to a Z-forming sequence in RNA would pose a similar barrier, and that Za may favor binding to regions which can be converted to the Z-conformation more easily (dsRNA regions adjacent to regions with low stability). The sequence of the Z-adopting region is likely also important, as for Za converting B- to Z-DNA. The number of hydrogen bonds at the protein-DNA interface is maximal for CpG but minimal for TpA. We similarly notice in our study that Alu adopts the Z-form at a lower protein concentration than H66 (extended duplex) > H43 > H25. From this analysis we conclude that (CpG)ₙ sequences are better Z-adopting sequences compared to other combinations, but that Gs in any configuration at the binding site (i.e., within UpG, ApG or GpG) tend to be preferred. Therefore, sequence specificity is likely driven by both the Z-forming propensity of the binding site and the stability of the adjacent regions (Fig. 8). We attempted to correlate the calculated Eₛ scores to other factors in our tested RNA fragments, such as the number and identity of dinucleotide steps within the RNA. However, the only significant correlations to come out of this analysis were to RNA length and the number of non-canonical base pairs. This highlights the complexity in the factors which contribute to the ability of a particular RNA to adopt the Z-conformation, of which our simple analysis was unable to uncover.

This work supports the hypothesis that Z-RNA formation is a general feature of RNA. We know that Z-DNA conformations are readily adopted during replication and transcription due to
unwinding of the two DNA strands. So we could expect Z-RNA formation to take place upon the many folding/unfolding events associated with the life of an RNA. Many crystal structures of non-coding RNAs reveal Z-RNA-like steps, suggesting they participate in non-standard RNA structure and dynamic events. Our NMR data indicate that Z-RNA formation within the ribosomal and Alu stems is more dynamic than within duplexes made exclusively of CpG repeats. A dynamic binding of Zα could be a desirable feature for ADAR1, as being bound to only certain RNAs for an extended period of time would likely not be conducive to editing >10,000 sites within cellular RNAs.

If Z-RNA formation was a broad mechanism for subtle regulation of RNA expression and/or folding, it would certainly make sense that its sequence requirements would be somewhat loose. A loose RNA sequence requirement for Z-RNA formation would not be paradoxical to recognition by Zα if the sequence requirements for the recognition process were also loose. In fact, double-stranded RNA binding domains from ADAR2 and other proteins display somewhat relaxed sequence requirements. Loose sequence requirements for Zα recognition may actually represent another illustration of the inherent “messy” nature of biology that would be at play during A-to-I editing.

Widespread Z-RNA formation would influence the editing activity of ADAR1, which in turn would affect the proportion of edited versus unedited double-stranded RNA. This mechanism could participate in helping the cell to modulate the innate immune IFN response. Our demonstration that Zα binds to non-CpG repeats at sites rich in non-Watson–Crick pairs and regions of lower stability is in agreement with earlier proposals that ADAR enzymes recognize their substrates in two separate events. As such, the role of Zα could be to attract or deliver ADAR1 to a particular structure adjacent to sites that would thus eventually be recognized by the double-stranded RNA binding domains and the deaminase domain of ADAR1. Zα could increase the lifetime of ADAR1 on RNAs that already have been significantly edited, through binding to or near the mismatches produced by A-to-I editing.

Whether Zα binds to ribosomal stem-loops in vivo remains unclear. It is possible that results from the pull-down assays were biased towards ribosomal RNA sequences, as they are most abundant in RNA extracts. On the other hand, Zα binding to ribosomal RNA could be relevant to ribosome assembly processes, or to the fate of ribosomes in the stress granules, where ADAR1 and other Z-binding proteins are also located. Studying Z-RNA formation in rRNA stem-loops helped us nonetheless with characterizing Zα binding to non-CpG repeat sequences and to identify similar regions in Alu elements.

Overall, pinpointing Z-RNA formation, occurrence, stability, and recognition by other macromolecules is crucial to getting the full picture of gene regulation. Once we better understand the structural underpinnings of Z-RNA formation, we will be able to determine how widespread such conformations are and what their role is within the cell.

**Methods**

**Plasmid construction, protein purification, and experimental buffers.** The N-terminal Zα domain of *Homo sapiens* ADAR1 in the pet-28a(+) plasmid (N-terminal 6x His-tag and thrombin cleavage site between His-tag and the Zα sequence) was a gift from Drs. Peter Dröge and Alekos Athanasiadis. Zα was expressed and purified similarly to. Briefly, the plasmid was transformed and expressed in BL21(DE3) E. coli. The cultures were grown in Luria Broth and induced with 1 M IPTG at an OD600 of 0.6 and allowed to express Zα for 4 h at 37 °C, then centrifuged to collect the cell pellets. Pellets were resuspended in lysis buffer (50 mM Tris–HCl (pH 8.0), 300 mM NaCl, 10 mM Imidazole, 5 mM β-Mercaptoethanol (BME)) and sonicated. Lysate was centrifuged and the supernatant was applied to a His-trap column, washed with 40 mL of lysis buffer, 80 mL of wash buffer (50 mM Tris–HCl (pH 8.0), 1 M NaCl, 10 mM Imidazole, 5 mM BME), and eluted in 20 mL of elution buffer (50 mM Tris–HCl (pH 8.0), 300 mM NaCl, 500 mM Imidazole, 1 mM BME). The eluted fraction was then dialyzed into 0.5 mM DTT, and incubated with bovine thrombin (Millipore-Sigma, Burlington, MA) at room temperature overnight. The resulting samples were then reapplied to a His-trap column and the flow-through was collected. The flow-throughs were...
then concentrated to ~2 mL and applied to a Superdex 75 Gel Filtration Column (120 mL, GE Healthcare) and peak corresponding to pure Za was collected and concentrated using an Amicon 3 kDa cutoff centrifugal filter (Millipore-Sigma, Burlington, MA). Za(Try177Ala), mutant was ordered from GenScript and prepared in the same way as Za. Proteins were dialyzed into 20 mM potassium phosphate (pH 6.4), 25 mM NaCl, 0.5 mM EDTA and 1 mM DTT. Proteins were concentrated using Amicon 3 kDa cutoff centrifugal filters. For CD and AUC, proteins were diluted into respective buffers from concentrated stocks. The buffers were as follows: 20 mM potassium phosphate (pH 6.4), 25 mM NaCl, 0.5 mM EDTA and 1 mM DTT; 20 mM potassium phosphate (pH 7.0), 25 mM NaCl, 0.5 mM EDTA, and 1 mM DTT for AUC.

RNA preparation. RNA constructs were purchased from Horizon Discovery (Boulder, CO) with HPLC purification. RNAs for NMR titrations were dialyzed into 20 mM potassium phosphate (pH 6.4), 25 mM NaCl, 0.5 mM EDTA, and concentrated to 300 µM using Amicon 3 kDa cutoff centrifugal filters. For CD and AUC, proteins were diluted into respective buffers from concentrated stocks. The buffers were as follows: 20 mM potassium phosphate (pH 6.4), 25 mM NaCl, 0.5 mM EDTA, and 1 mM DTT; 20 mM potassium phosphate (pH 7.0), 25 mM NaCl, 0.5 mM EDTA, and 1 mM DTT for AUC.

Circular dichroism. All CD measurements were collected in 1 nm steps from 320 to 220 nm using a JASCO J-810 CD spectrometer (run using Spectra Manager version 2 (JASCO)) in a 0.1 cm quartz cuvette in 20 mM potassium phosphate (pH 6.4), 25 mM NaCl, 0.5 mM EDTA, and 1 mM DTT at 25 °C. Two scans were acquired for each sample (average). RNAs were heat denatured at 95 °C for 1 min followed by rapid cooling on ice and titration samples were prepared by incubating 50 µM of the RNA constructs with the specified amount of Za or Za(Try177Ala) at 42 °C for 30 min and then bringing the samples down to 25 °C over a period of 20 min. Control experiments were run to ensure that the absorbance of Za and Za(Try177Ala) alone at the same concentrations was minimal in the 250–320 nm range (the range which reports on RNA secondary structure) (Supplementary Fig. 2). This indicated that changes in the CD spectra were due to conformational changes in the nucleic acid and not from the superposition of the RNA and Za spectra. We tested our RNAs with 6 M sodium perchlorate which was shown to fully induce the Z-conformation in the (CpG)n, RNAs57. We refer to either the RNA:Protein or the Protein:RNA ratio, with the first molecule mentioned being the one that is being titrated into, for CD, ITC, and NMR measurements.

Extent of Z-conformation calculation. To experimentally quantify the extent of Z-conformation present in a RNA construct, we introduced a Z-conformation score derived from CD spectra (EZ). EZ is based on the CD intensities at wavelengths 285 and 295 nm (Int285 and Int295), both of which have been shown to be characteristic of Z-conformation22,41, and the decrease in intensity at 266 nm (Int266), which reports on the reduction of the A-form. We assumed that the growth of the intensities at 285 and 295 nm and decrease at 266 nm are proportional to the presence of Z-conformation with an offset accounting for intensities measured when no Z-conformation is present. We calibrated the following equation from CD intensities measured from (CpG)n, where the intensities reach plateaus at 100% Z-conformation:

\[ E_Z = \frac{(1.800 \times \text{decay}^{266} + 0.718 \times \text{growth}^{285} + 1.109 \times \text{growth}^{295})}{3} \]

where \( \text{decay} = \left( \frac{\text{Int}^{266}_{\text{free}} - \text{Int}^{266}_{\text{bound}}}{\text{Int}^{266}_{\text{free}}} \right) \)

\( \text{growth}^{285} = \left( \frac{\text{Int}^{285}_{\text{bound}} - \text{Int}^{285}_{\text{free}}}{\text{Int}^{285}_{\text{free}}} \right) \)

\( \text{growth}^{295} = \left( \frac{\text{Int}^{295}_{\text{bound}} - \text{Int}^{295}_{\text{free}}}{\text{Int}^{295}_{\text{free}}} \right) \)

and the prefactors were chosen so that the EZ score of the (CpG)n RNA would be equal to one. EZ ideally takes values between 0 (no Z-conformation) and 1 (each nucleotide 100% in Z-conformation). Values between 0 and 1 indicate that the molecule is partially in Z-conformation. This may arise from differences in the various nucleotides such that a specific EZ score can be achieved by different Z patterns. For example, if only a subset of the nucleotides adopts Z-conformation at 100%, EZ will be lower than 1.

An error of 0.1 was determined to be appropriate for the calculated EZ scores by taking into account the difference between the (CpG)n and (CpG)RNA (which theoretically both have an EZ score of 1) and the difference in the EZ score between repeat measurements on h43 E.coli.

The fractional decrease in the EZ score (Fig. 2e) was calculated by the following equation:

\[ \Delta E_Z = \frac{E_Z^{\text{Zanp}} - E_Z^{\text{Zanp'/Zan}}}{{E_Z^{\text{Zanp'}} + E_Z^{\text{Zanp'}/Zan}}} \]

where \( E_Z^{\text{Zanp'}} \) and \( E_Z^{\text{Zanp'/Zan}} \) are the EZ scores determined from the CD measurements with the wild type Za and mutant Za (Try177Ala), respectively.

Isothermal titration calorimetry. RNA constructs for ITC and the Za protein were dialyzed overnight into 20 mM potassium phosphate (pH 7.0), 25 mM NaCl, 0.5 mM EDTA, and 1 mM DTT to match buffer conditions (in the same beaker with the RNA (to avoid any differences) and for ITC, proteins were dialyzed in the same beaker with the RNA (to avoid any differences). Values between 0 and 1 indicate that the molecule is 100% in Z-conformation. Values between 0 and 1 indicate that the molecule is 100% in Z-conformation. While the EZ score can be slightly underestimated by negative contribution to the molar ellipticity at ~290 nm at the higher concentrations of Za (Supplementary Fig. 2), this contribution is relatively small and is most canceled out by the fact that we weigh the EZ score according to the (CpG)n control (which has the same concentration of Za as the other constructs).

Analytical ultracentrifugation. For the interactions between Za and the different RNAs (Extended Data Fig. 5), the concentrations of Za and the RNA tested were 2 and 12 µM, respectively, corresponding to a 1:6 ratio of RNA:protein, except for h43 which was measured at three different concentrations of Za (4, 8, and 12 µM). Samples were loaded into a cell composed of standard 12 mm EPON pericopes with quartz windows and sedimented at 50,000 RPM at 25 °C using an XL-1 (Beckman Coulter) AUC instrument (run using ProteomeLab version 6.0 (Beckman)). UV absorbance was monitored at 260 nm for 16 h. Data were analyzed with SEDFIT (Version 14.7 g, NIH)39 using a specific volume which was normalized to the weight-average of RNA and protein78. The complex stoichiometry was chosen according to which theoretical weight was the closest to the measured weight. Error in the measured molecular weight by AUC can be caused by deviation of the predicted specific volume or viscosity of the sample from the actual values. In addition, if one of the binding sites is weaker than the others resulting in decreased site occupancy, this may cause a deviation in the observed molecular weight from the predicted complex size.
32 scans. The spectral widths were 20 × 20 ppm for both 1H dimensions. The 2D 1H-1H NOESY recorded for H25 was acquired with a mixing time of 200 ms, 1396 × 400 complex points (162 of the points were collected following a 40% NUS sampling schedule), a 1.3 s recycle delay, and 32 scans. The spectral widths were 24.5 × 24.5 ppm for both 1H dimensions. 1D spectra from the Bruker spectrometer were processed using TopSpin 2.4.263. The 15N-HSQC spectra for the titration were recorded using a Varian 900 MHz spectrometer. All NMR measurements were carried out on the Varian 900 MHz spectrometer. The 15N-HSQC spectra for the titration were measured with 1048 (1H) × 64 (15N) complex points, a 1.3 s recycle delay, and 32 scans. The spectral widths were 16 × 35 ppm for the 1H and 15N dimensions, respectively. During the R1 relaxation experiments were measured with 1048 (1H) × 64 (15N) complex points, a 1.3 s recycle delay, and 32 scans, and relaxation delays of 0, 10, 20, 30, 40, 60, 80, 100, 120, and 160 ms. The spectral widths were 16 × 35 ppm for the 1H and 15N dimensions, respectively. After the R1 relaxation time, a 15N spin-lock field of 1500–1500 Hz strength was applied. The transverse relaxation rate R2 was calculated from R1 and R2, using the following equation:

\[ R_2 = R_{tp} + (R_{tp} - R_1) \times \tan(\theta) \] (3)

where θ is the resonance offset, \( \gamma_B/2\pi \) is the gyromagnetic ratio of the 15N spin, and \( \gamma_B/2\pi \) is the magnetic field of the 15N-HSQC experiment. The CPMG experiment at 2:1 Z.1 Jo into 15N-labelled Z.1 Jo was acquired with a mixing time of 200 ms, 1048 (1H) × 60 (15N) complex points, a 1.3 recycle delay, and 32 scans. The spectral widths were 16 × 35 ppm for the 1H and 15N dimensions, respectively. The R2 relaxation experiments were measured with 1048 (1H) × 64 (15N) complex points, a 1.3 s recycle delay, 32 scans, and relaxation delays of 0, 10, 20, 30, 40, 60, 80, 100, 120, and 160 ms. The spectral widths were 16 × 35 ppm for the 1H and 15N dimensions, respectively. During the R1 relaxation time, a 15N spin-lock field of 1500–1500 Hz strength was applied. The transverse relaxation rate R2 was calculated from R1 and R2, using the following equation:

\[ R_{2eff} = R_2 + p_r \times \left(1 - \frac{1}{\tau_{relax}}\right) \times \left(\Delta\nu\right)_R \times \left(1 - \left(4 \times \gamma_D/\gamma_N \right) \times \left(\frac{\Delta\nu}{\gamma_N}\right)\right) \] (4)

where \( \Delta\nu_R \) and \( p_r \) are the R2 relaxation rate and the population of state A, \( \Delta\nu \) is the difference in the chemical shift between states A and B, and \( \Delta\nu \) is the rate of exchange between states A and B, and \( \Delta\nu \) is the effective field strength of the refocusing pulse train.

The 15N-CEST experiment at 2:1 Z.1 Jo into 15N-labelled Z.1 Jo was recorded on the Varian 900 MHz spectrometer with 1024 (1H) × 64 (15N) complex points, a 1 s recycle delay, and 16 scans. The spectral widths were 16 × 35 ppm for the 1H and 15N dimensions, respectively. A weak 15N spin-lock field of 5 Hz strength was applied during a 400 ms relaxation time. A total of 66 datasets were acquired corresponding to a chemical shift range of 102–132 ppm with steps of 0.5 ppm. From 115 to 116 ppm, the steps were decreased to 0.25 ppm to acquire additional points around Tyr177. In all experiments, 1H decoupling was achieved using a 90,240,90, composite pulse.

\( K_0 \) fitting of NMR titration points. CSPs from the 15N-HSQC titration of AluSx1Jo into 15N-labelled Z.1 Jo were acquired corresponding to a chemical shift range of 102–132 ppm with steps of 0.5 ppm. From 115 to 116 ppm, the steps were decreased to 0.25 ppm to acquire additional points around Tyr177. In all experiments, 1H decoupling was achieved using a 90,240,90, composite pulse.

\[ K_0 = \frac{[L]}{[L] + [P]} - \left(\frac{[P]}{2}\right) - \left(\frac{[L]}{2}\right) \] (6)

Comparison of CSPs between AluSx1Jo titration and (CpG)2 DNA and RNA. For comparison of our measured CSPs from the 15N-HSQC titration of Z.1 Jo with the AluSx1Jo, all other data were measured using the (CpG)2 DNA and RNA and AluSx1Jo DNA junction23, CSPs were estimated from the published bar graphs showing CSPs vs residue number.

References
1. Wagner, R. W., Smith, J. E., Cooperman, B. S. & Nishikura, K. A double-stranded RNA unwinding activity introduces structural alterations by means of adenosine to inosine conversions in mammalian cells and Xenopus eggs. Proc. Natl. Acad. Sci. USA https://doi.org/10.1073/pnas.88.8.2647 (1991).
2. Nishikura, K. A-to-I editing of coding and non-coding RNAs by ADARs. Nat. Rev. Mol. Cell Biol. https://doi.org/10.1038/nrm.2015.4 (2016).
3. Bass, B. L. & Weintraub, H. An unwinding activity that covalently modifies its double-stranded RNA substrate. Cell https://doi.org/10.1016/0092-8674(88) 90253-X (1988).
4. Chung, H. et al. Human ADAR1 prevents endogenous RNA from triggering translational shutdown. Cell https://doi.org/10.1016/j.cell.2017.12.038 (2018).
5. Mannion, N. M. et al. The RNA-editing enzyme ADAR1 controls innate immune responses to RNA. Cell Rep. https://doi.org/10.1016/j.celrep.2014.10.041 (2014).
6. Patterson, J. B. & Samuel, S. E. Expression and regulation by interferon of a double-stranded RNA-specific adenosine deaminase from human cells: evidence for two forms of the deaminase. Mol. Cell. Biol. 15, 5376–5388 (1995).
7. O’Connell, M. A. et al. Cloning of cDNAs encoding mammalian double-stranded RNA-specific adenosine deaminase. Mol. Cell. Biol. https://doi.org/10.1128/mcb.15.3.1389 (1995).
8. O’Connell, M. A. & Keller, W. Purification and properties of double-stranded RNA-specific adenosine deaminase from calf thymus. Proc. Natl. Acad. Sci. USA https://doi.org/10.1073/pnas.91.22.10596 (1994).
9. George, C. X. & Samuel, C. E. Human RNA-specific adenosine deaminase ADAR1 transcribes possess alternative exon 1 structures that initiate from different promoters, one constitutively active and the other interferon inducible. Proc. Natl. Acad. Sci. USA https://doi.org/10.1073/pnas.96.8.4621 (1999).
10. Levanon, K., Eisenberg, E., Rechavi, G. & Levanon, E. Y. Adenosine-to-inosine RNA editing in Alu repeats in the human genome. EMBO Rep. 6, 831–835 (2005).
11. Levanon, E. Y. et al. Systematic identification of abundant A-to-I editing sites in the human transcriptome. Nat. Biotechnol. 22, 1001–1005 (2004).
12. Athanasiadis, A., Rich, A. & Maas, S. Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. PLOS Biol. https://doi.org/10.1371/journal.pbio.0020391 (2004).
13. Figeys, D., Petre, R. & Batzer, M. The impact of retrotransposons on human genome evolution. Nat. Rev. Genet. 10, 691–703 (2009).
14. Heinrich, M. J. et al. Endogenous double-stranded Alu RNA elements stimulate IFN-responses in relapsing remitting multiple sclerosis. J. Autoimmun. https://doi.org/10.1016/j.jaut.2019.02.003 (2019).
15. Hung, T. et al. The B067 autoantigen binds endogenous retroelements and regulates inflammatory gene expression. Science https://doi.org/10.1126/science.aac7442 (2015).
16. Vlachogiannis, N. I. et al. Increased adenosine-to-inosine RNA editing in rheumatoid arthritis. J. Autoimmun. https://doi.org/10.1016/j.jaut.2019.02.029 (2020).
17. Zhao, Y. & Karjolich, J. Know thyself: RIG-1-like receptor sensing of DNA virus infection. J. Virol. https://doi.org/10.1128/jvi.01085-19 (2019).
18. Zhao, Y., Ye, X., Dunker, W., Song, Y. & Karjolich, J. RIG-1-like receptor sensing of host RNA facilitates the cell-intrinsiv immune response to KSHV infection. Nat. Commun. https://doi.org/10.1038/s41467-018-07314-7 (2018).
19. Pfäfler, C. K., Donohue, R. C., Nersisyan, S., Brodsky, L. & Cattaneo, R. Extensive editing of cellular and viral double-stranded RNA structures accounts for immune community support to the (and the proviral activity of ADAR1 p150. PLOS Biol. https://doi.org/10.1371/journal.pbio.2006577 (2018).

20. Herbert, A. et al. The Z domain from human ADAR1 binds to the Z-DNA conformer of many different sequences. Nucleic Acids Res. https://doi.org/10.1093/nar/26.15.3486 (1998).

21. Schwartz, T., Rich, A. & Lowenhaupt, K., Herbert, A. & Rich, A. Crystal structure of the Zalpha domain of the human editing enzyme ADAR1 bound to left-handed Z-DNA. Science 11, 1841–1845 (1999).

22. Brown, B. A., Lowenhaupt, K., Wilbert, C. M., Hanlon, E. B. & Rich, A. The Za domain of the editing enzyme dsRNA adenosine deaminase binds left-handed Z-RNA as well as Z-DNA. Proc. Natl. Acad. Sci. USA https://doi.org/10.1073/pnas.1728607997 (2000).

23. Lee, A. R. et al. NMR dynamics study reveals the Za domain of human ADAR1 associates with and dissociates from Z-RNA more slowly than Z-DNA. ACS Chem. Biol. https://doi.org/10.1021/acschembio.8b00914 (2019).

24. Koeris, M., Funke, L., Shrestha, J., Rich, A. & Maas, S. Modulation of ADAR1 editing activity by Z-RNA in vitro. Nucleic Acids Res. https://doi.org/10.1093/nar/gkx849 (2015).

25. Placido, D., Brown, B. A., Lowenhaupt, K., Rich, A. & Athanasiadis, A. A left-handed RNA double helix bound by the Za domain of the RNA-editing enzyme ADAR1. Structure https://doi.org/10.1016/j.str.2007.03.001 (2007).

26. Kruse, H., Mrazikova, K., D’Ascenzo, L., Sponer, J. & Auffering, P. Short but weak: The Z-DNA–Z-DNA interaction. Nucleic Acids Res. https://doi.org/10.1093/nar/gkx388 (2016).

27. Herbert, A. Mendelian disease caused by variants affecting recognition of Z-DNA and Z-RNA by the Za domain of the double-stranded RNA editing enzyme ADAR. Eur. J. Hum. Genet. 28, 114–117 (2020).

28. Rice, G. J. et al. Mutations in ADAR1 cause Aicardi-Goutières syndrome associated with a type I interferon signature. Nat. Genet. https://doi.org/10.1038/nj.2014.212 (2014).

29. Ha, S. C. et al. The structures of non-CG-repeat Z-DNAs co-crystalized with the Z-DNA-binding domain, hZaADAR1. Nucleic Acids Res. https://doi.org/10.1093/nkx697 (2016).

30. D’Ascenzo, L., Leonarsi, F., Vicens, Q. & Auffering, P. ‘Z-DNA like’ fragments in RNA: a recurring structural motif with implications for folding, RNA/protein recognition and immune response. Nucleic Acids Res. https://doi.org/10.1093/nkx697 (2016).

31. Lee, Y. M. et al. NMR investigation on the DNA binding and B-Z transition of the Z-DNA–Z-DNA complex. Nucleic Acids Res. https://doi.org/10.1093/nar/gky784 (2018).

32. Lee, Y. M. et al. NMR study on the DNA binding and B-Z transition pathway of the Z domain of human ADAR1. Biophys. Chem. https://doi.org/10.1016/j.bjrc.2012.12.002 (2013).

33. Lee, Y. M. et al. NMR study on the B-Z junction formation of DNA duplexes induced by Z-DNA binding domain of human ADAR1. J. Am. Chem. Soc. https://doi.org/10.1021/ja11581b (2012).

34. Kim, D. et al. Sequence preference and structural heterogeneity of BZ junctions. Nucleic Acids Res. https://doi.org/10.1093/gkx784 (2018).

35. Koeris, M., Funke, L., Shrestha, J., Rich, A. & Maas, S. Characterization of the crystal structure of a junction between B-DNA and Z-DNA reveals two extruded base pairs. Nature 437, 1183–1186 https://doi.org/10.1038/nature04088 (2005).

36. Kim, D. et al. Base extrusion is found at helical junctions between right- and left-handed forms of DNA and RNA. Nucleic Acids Res. https://doi.org/10.1093/nkx697 (2016).

37. Herbert, A. Z-DNA and Z-RNA in human disease. Commun. Biol. https://doi.org/10.1038/s42003-018-0237-x (2019).

38. Feng, S. et al. Alternate rRNA secondary structures as regulators of translation. Nat. Struct. Mol. Biol. https://doi.org/10.1038/nmb1962 (2011).

39. Jiao, H. et al. Z-nucleic-acid sensing triggers ZBP1-dependent necroptosis and Lamm equation modeling. Biophys. J. https://doi.org/10.1016/j.bpj.2012.12.002 (2013).

40. Herbert, A. Z-DNA and Z-RNA in human disease. Commun. Biol. https://doi.org/10.1038/s42003-018-0237-x (2019).

41. Lehmann, K. A. & Bass, B. L. The importance of internal loops within RNA substrates of ADAR1. J. Mol. Biol. https://doi.org/10.1016/j.jmb.1999.2914 (1999).

42. Daniel, C., Wüthrich, K. & Ohman, M. Editing inducer elements increase A-to-I editing efficiency in the mammalian transcriptome. Genome Biol. https://doi.org/10.1186/s13059-017-1324-x (2017).

43. Mitahara, T., Nakatani, H. & Sugiyama, H. Similarities and differences between RNA and DNA double-helical structures in circadian dioxorhodopsin spectroscopy: a SAC-CI study. J. Phys. Chem. A https://doi.org/10.1021/acs.jpca.6b08023 (2016).

44. Schade, M., Turner, C. J., Lowenhaupt, K., Rich, A. & Herbert, A. Structure-function analysis of the Z-DNA-binding domain Zalpha of dsRNA adenosine deaminase type I reveals similarity to the (alpha + beta) family of helix-turn-helix proteins. EMBO J. 18, 470–479 (1999).

45. Kang, Y. M. et al. NMR spectroscopic elucidation of the B-Z transition of a DNA double helix induced by the Za domain of human ADAR1. J. Am. Chem. Soc. https://doi.org/10.1021/ja902654a (2009).

46. Lee, A. R. et al. Solution structure of the Z-DNA binding domain of PKR-like protein kinase from Carassius auratus and quantitative analyses of the intermediate complex deactivating the B-Z transition. Nucleic Acids Res. https://doi.org/10.1093/gkx025 (2016).

47. Mathews, D. H. et al. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. Proc. Natl Acad. Sci. USA https://doi.org/10.1073/pnas.0401799101 (2004).

48. Jiao, H. et al. Structure and hydrogen exchange during the B-Z transition of a DNA duplex induced by the Za domains of yatapoxivirus E3L. FEBS Lett. https://doi.org/10.1016/S0014-5793(10)00303 (2010).

49. Lee, A. R. et al. Solution structure of the Z-DNA binding domain of PKR-like protein kinase from Carassius auratus and quantitative analyses of the intermediate complex deactivating the B-Z transition. Nucleic Acids Res. https://doi.org/10.1093/gkx025 (2016).

50. Herbert, A. & Rich, A. Structure and hydrogen exchange during the B-Z transition of a DNA duplex induced by the Za domains of yatapoxivirus E3L. FEBS Lett. https://doi.org/10.1016/S0014-5793(10)00303 (2010).

Acknowledgements

The authors thank Peter Dröge and Aleksos Athanasiadis for the gift of the 6x His-tag (N-terminal) Za plasmid [pet-28a(+)]; current and former Kieft lab and Vogel lab members for thoughtful discussions and technical assistance; Jared Lindenberg for help with ARCG measurement; the biochemistry core and NMR facilities at the University of Colorado Denver; and Pascal Auffering, Chrysa Latrick, Ben Akiyama, and David Costanzo for helpful discussions and critical reading of the manuscript. This research is funded by NSF grant 1917254 for Infrastructure Innovation for Biological Research and a start-up package from the University of Colorado to B.V.H; NIH grant R53GM118070 to J.K.; NSF grant 1502002 for shared and high-end instrumentation award. The University of Colorado Anschutz Medical Campus NMR facility is partially supported by NIH grant P30CA4693.
Author contributions
P.N. performed all experiments and data analysis with the help of S.B. (for CD, AUC, ITC) and M.H. (for NMR). P.N., J.S.K., Q.V., and B.V. interpreted the data. P.N., Q.V. and B.V. conceptualized the study, wrote and revised the manuscript. All authors provided comments on the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-21039-0.

Correspondence and requests for materials should be addressed to Q.V. or B.Vög.

Peer review information Nature Communications thanks Hashim Al-Hashimi, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021