Za is a peptide motif that binds to Z-DNA with high affinity. This motif binds to alternating dC-dG sequences stabilized in the Z-conformation by means of bromination or supercoiling, but not to B-DNA. Za is part of the N-terminal region of double-stranded RNA adenosine deaminase (ADAR1), a candidate enzyme for nuclear pre-mRNA editing in mammals. Za is conserved in ADAR1 from many species; in each case, there is a second similar motif, Zb, separated from Za by a more divergent linker. To investigate the structure-function relationship of Za, its domain structure was studied by limited proteolysis. Proteolytic profiles indicated that Za is part of a domain, Zab, of 229 amino acids (residues 133–361 in human ADAR1). This domain contains both Za and Zb as well as a tandem repeat of a 49-amino acid linker module. Prolonged proteolysis revealed a minimal core domain of 77 amino acids (positions 133–209), containing only Za, which is sufficient to bind left-handed Z-DNA; however, the substrate binding is strikingly different from that of Zab. The second motif, Zb, retains its structural integrity only in the context of Zab and does not bind Z-DNA as a separate entity. These results suggest that Za and Zb act as a single bipartite domain. In the presence of substrate DNA, Zab becomes more resistant to proteases, suggesting that it adopts a more rigid structure when bound to its substrate, possibly with conformational changes in parts of the protein.

Many protein domains that recognize DNA in both sequence- and conformation-specific manners have been characterized (for a review, see Ref. 1). These studies have resulted in an understanding of the variety of ways in which protein-DNA interactions can result in function. Identification of a peptide motif, Za, which binds specifically to Z-DNA, opens up a new vista and invites the investigation of the similarities and differences between domains that bind right- and left-handed DNAs. The conformation specificity of Za binding has been characterized in many ways. Peptides including this motif bind to alternating dC-dG that has been stabilized in the Z-conformation using bromination or supercoiling, as shown by band shift assays, competition experiments, and BIAcore measurements (2). When linked to the nuclease domain from FokI, the resulting chimeric nuclease cuts supercoiled plasmid DNA to bracket a d(C-G)13 in the Z-conformation (3). The protein also binds to short oligonucleotides of suitable sequence and canverts them from the B- to the Z-conformation, as detected by CD and Raman spectroscopy (4, 5). The binding of Z-DNA by Za occurs even in the presence of a 10^5-fold excess of B-DNA (6). Za binds poly(dC-dG), stabilized in the Z-conformation by bromination, with an equilibrium dissociation constant (Kd) in the lower nanomolar range, as shown by BIAcore measurements (2).

Although many properties of Za have been studied, its biological function in the context of ADAR1 remains unknown. The Z-DNA binding activity of Za was first identified in proteolytic fragments of double-stranded RNA adenosine deaminase (ADAR1) (6) and then in the full-length enzyme (7). Za has been shown to be a conserved feature of human, rat, bovine, chicken, and Xenopus ADAR1 (2). A second related motif, Zb, has been identified in all the ADAR1 enzymes whose sequences are known. These two motifs are separated by a linker region of conserved size; an exception is the human enzyme, in which the linker is twice as long and consists of two nearly identical copies of a module (8). The presence of a conserved N-terminal region containing these motifs distinguishes ADAR1 from other members of the ADAR family (9, 10), and the N terminus has been shown to be differentially expressed (8). Therefore, we conclude that this region is of importance for the biological function of ADAR1.

The ADAR family of enzymes converts adenosine to inosine within double-stranded regions of RNA (11). In mRNA, inosine is read as guanosine by the translation apparatus, resulting in codon changes within the synthesized protein. A-to-I editing has been shown to occur in vivo in a number of mRNAs from higher animals (12–18). The best characterized of these, the editing of pre-mRNAs for subunits of the glutamate-gated cation channels in the brain, results in channels with dramatically altered functional properties (19). Double-stranded RNA structures required for ADAR activity are formed by base pairing of an exonic sequence around the editing site with a complementary sequence in the downstream intron; therefore, editing must take place in the nucleus before splicing removes the respective intron(s). It has been proposed that Za serves to target ADAR1 to its preferred substrates by binding to Z-DNA formed close to actively transcribing genes (20).

To better understand the role of Za, we have characterized the N-terminal region of ADAR1 functionally and structurally. Using human ADAR1 as a model, the classical approach of limited proteolysis was employed to define the boundaries of this domain. Both motifs, Za and Zb, together are shown to form a single functional domain, Zab; Zab is stable and protected from proteolysis. Za, containing Za, but not Zb, can be regarded as a stable subdomain; this subdomain contributes the main binding activity. There is no equivalent subdomain containing Zb; this region is poorly structured and unstable when isolated. The intervening linker region is unexpectedly well structured, In humans, the second copy of the linker mod-

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Supported by a predoctoral fellowship from the German Academic Exchange Service.

§ To whom correspondence should be addressed.

This paper is available on line at http://www.jbc.org
ule appears to have a structure similar to the first. Removing one copy of the linker modules reduces the DNA binding affinity, indicating the importance of the distance between the Za and Zβ motifs. Za binds Z-DNA in a conformationally specific, but not sequence-specific manner. The binding is modified by the presence of the entire Zab domain to confer preference for d(C-G)n over d(C-A)n-d(T-G)n.

**EXPERIMENTAL PROCEDURES**

**ADAR1: DNA Constructs and Protein Purification**—Different portions of the cloned cDNA coding for human ADAR1 (GenBank accession number U10403) were polymerase chain reaction-amplified and inserted into the pET28a (Novagen) vector, resulting in N-terminal His6-tagged fusion proteins. In detail, Za131 (residues 96–226), Za77 (133–209), and Za236 (133–368) were amplified using complementary primers flanked with restriction sites at their termini. Polymerase chain reaction products were analyzed on an agarose gel; bands of the correct size were extracted and subcloned into the Ndel-HindIII sites (Za131 and Za77) or the Ndel-HindIII sites (Za236) of the multiple cloning site of pET28a, resulting in the vectors pZa131, pZa77, and pZa236, respectively. Another construct, ZabΔ1, lacking one of the two 49-amino acid linker modules separating the Za and Zβ motifs, was created from pZab236 as follows. The 1.1-kilobase Sphi-HindIII restriction fragment was digested with the restriction enzyme Drd1, resulting in two cleavage sites at identical locations at nucleotides 759 and 926 (amino acids 197 and 244 for GenBank accession number U10403). The resulting DNA fragments were depurinized and precipitated (21). After incubation with T4 DNA ligase (25 °C, 4 h), the reaction mixture was analyzed on an agarose gel. The 930-base pair ligation product was isolated and subcloned in the 5-kilobase Sphi-HindIII restriction fragment of pET28a, resulting in the vector pZabΔ1. To ensure that the plasmids were correct, they were analyzed by restriction digestion, and the coding regions were sequenced using Sequenase Version 2.0 (U. S. Biochemical Corp.). According to the manufacturer’s instructions. The proteins were overproduced in *Escherichia coli* strain NovablaDE3 (Novagen). Bacteria were grown at 37 °C in Luria-Bertani medium and induced with 1 mM isopropyl-β-D-thiogalactopyranoside at 0.7–0.9 A600 nm units. Cells were harvested after a further 3 h of growth at 37 °C. All subsequent steps were done at 4 °C. The proteins were purified essentially to homogeneity under nondenaturing conditions as follows. A cell pellet obtained from a 1-liter culture was resuspended in 15 ml of buffer A (50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 10 mM imidazole, 5 mM β-mercaptoethanol, 20 μg/ml RNase A, and 100 μM phenylmethylsulfonyl fluoride), and the cells were lysed using a French press. The lysate was then centrifuged for 25,000 × g, and the supernatant was separated and incubated with 2 ml of Ni2+-nitrilotriacetic acid metal affinity resin (QIAGEN Inc.) for 1 h. The resin was washed three times with 20 ml of buffer B in a batch and then washed with 40 ml of buffer B (50 mM Tris-HCl (pH 8.0), 1 mM NaCl, 10 mM imidazole, and 5 mM β-mercaptoethanol) in a column. Overproduced His6-tagged fusion protein was eluted with an imidazole step gradient in buffer C (50 mM Tris-HCl (pH 8.0), 300 mM NaCl, and 5 mM β-mercaptoethanol). Steps were 30, 50, and 200 mM imidazole, respectively. Fractions were analyzed by denaturing SDS-polyacrylamide gel electrophoresis (PAGE) on 15 or 18% gels. Fractions containing protein were pooled and dialyzed against buffer D (20 mM Tris-HCl (pH 8.0), 150 mM NaCl, and 2 mM dithiothreitol (DTT)). After 1 h of dialysis, 15 units of thrombin (Calbiochem) were added in buffer D to cleave the remaining His6 tag. 12 h later, the cleaved protein was dialyzed against buffer E (20 mM HEPES (pH 7.5), 20 mM NaCl, and 2 mM DTT) and finally purified by cation-exchange chromatography on a Mono S HR5/5 column (Amersham Pharmacia Biotech). Proteins were eluted with a 30-ml linear gradient of NaCl (0.05–0.3 M) in 20 mM HEPES (pH 7.5) and 1 mM DTT at a flow rate of 0.7 ml/min, resulting in sharp peak profiles. Za77 eluted at 220 mM NaCl, ZabΔ1 at 200 mM, and Zab236 at 180 mM. The yield of electrophoretically homogeneous protein was determined using extinction coefficients of 14,000 M−1 cm−1 (Za77 and ZabΔ1), 22,400 M−1 cm−1 (ZabΔ1), and 28,020 M−1 cm−1 (Zab236) at the absorbance maximum at 278 nm (calculated as described in Ref. 22). 8–12 mg of protein were obtained per liter of bacterial culture.

**Limited Proteolysis**—Protease digestion was performed by treating 50 μg of protein in 20 mM Tris-HCl, pH 7.5, with trypsin, chymotrypsin, thermolysin, or *Staphylococcus aureus* endoprotease Glu-C in 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, and 2 mM DTT at a protein/protease mass ratio in the range of 50:1 to 1000:1 for various times at 24 °C. Reactions were stopped by heat denaturation at 100 °C for 5 min. To examine the effect of various DNA conformers on Zab digestion, the reaction was performed in 10 mM Hepes (pH 7.5), 20 mM NaCl, 5 mM DTT, and 10 mM poly[d(A-T)] or poly[d(G-C)] was used as substrate DNA, and poly[d(5-MeC-G)] was used as substrate DNA, and poly[d(A-G)]-poly[d(C-T)] as unspecific DNA. The digests were separated by SDS-PAGE on 18% gels, followed by staining with Coomassie Brilliant Blue G-250. In the case of protein digested for the experiment shown in Fig. 6 (lanes 10–13), the reactions were stopped by adding phenylmethylsulfonyl fluoride (1 mM) instead of heat inactivation to ensure nondenatured protein. Protein fragments were analyzed by mass spectrometry as a Voyager DE Workstation (PerSeptive) using matrix-assisted laser desorption ionization-time of flight technology. As a matrix, sinapinic acid (10 μg/μl) in acetone/ H2O/trifluoroacetic acid (70:29:9:0.1) was used. Alternatively, for fragments smaller than 10-kDa, the matrix was prepared with α-cyanoaminic acid (10 μg/μl) instead of sinapinic acid. Various fragments were further analyzed by amino terminal sequencing on an Applied Biosystems Model 477A/477A protein sequencer.

**DNA Binding Assay**—DNA binding was assayed by native PAGE (23). The assay was carried out using d(5-MeC-G)n as the substrate, which is stable in the left-handed Z-DNA conformation under the applied conditions (24). The substrate was end-labeled with 32P and purified by native PAGE prior to the experiment. A reaction mixture of 10 μg of the ADAR1 construct Za77 (133–209) or Za236 (133–368) was incubated with 10 mM Tris-HCl (pH 7.8), 20 mM NaCl, 5 mM DTT, 5% glycerol, 100 μg/ml bovine serum albumin, and 50 μg/ml poly[d(A-G)]-poly[d(C-T)] (Amersham Pharmacia Biotech) as an unspecific competitor was incubated for 30 min at 24 °C. The mixture was analyzed on a 6% native polyacrylamide gel using 0.5% Tris borate (22.5 mM) as the running buffer. After electrophoresis (10 V/cm, 90 min), the gel was dried and autoradiographed at −70 °C on Kodak X-Omat Blue film with intensifying screens.

**CD Measurements**—CD spectra were recorded at 24 °C on an Aviv Model 62DS spectrometer. Conformational changes in DNA oligomers was monitored between 235 and 305 nm. DNA samples used were annealed prior to the experiment. For this purpose, a concentrated solution of the self-complementary sequence d(C-G)n or an equimolar amount of d(C-A)n, and d(T-G)n, was heated to 85 °C for 10 min and slowly cooled to −20 °C over 1 h. Measurements were carried out in 10 mM sodium phosphate (pH 7.0), 10 mM NaF, 1 mM EDTA, and 2 mM DTT using a DNA concentration of 30.0 μM base pairs and an optical path length of 5 mm. Spectra were recorded in 10-nm steps and averaged over 4 s. Protein was added to the sample from a concentrated stock solution, in aliquots never exceeding 5% of the total volume, and the mixture was equilibrated for 5 min before each measurement. The mixture was corrected for the buffer and protein. Software provided by Aviv. Protein spectra were recorded between 190 and 250 nm. Za77 was measured at a concentration of 10.0 μM, and ZabΔ1 and Za were measured at 5.0 μM and an optical path length of 1 mm. Spectra were measured in 1-nm steps and averaged over 10 s.

For comparisons of the spectra of Zab between 190 and 250 nm in the presence and absence of substrate, poly[d(5-MeC-G)] was used as substrate. A 2:1 base pair/protein molar ratio was used.

**RESULTS**

**Defining the Boundaries of the Minimal Z-DNA-binding Domain of Human ADAR1**—Protein domains are usually well structured regions of 50–200 amino acids (25, 26). Larger proteins are built from multiple, mostly independently folded domains. The regions connecting those domains are often flexible and solvent-exposed. Limited proteolysis is a classical approach to define domain organization (27–30). It takes advantage of the fact that site-specific proteases will cleave proteins preferentially in solvent-exposed unstructured regions, rather than within a folded domain.

Limited proteolysis was used to define a structured core containing Za, the Z-DNA-binding motif present in the N-terminal region of ADAR1. Za has been defined to comprise residues 121–197 of human ADAR1 using functional assays (2). However, a variety of results from nondenaturing electrophoresis, chromatographic elution, and NMR studies have suggested
that the recombinantly produced peptide is not stably folded.\(^2,3\)

An N- and C-terminally extended portion of the ADAR1 N terminus comprising Gly\(^{96}\)-Ser\(^{226}\) of human ADAR1, was overproduced as a His\(_6\) tagged fusion protein in E. coli, and its digestion with four different proteases (endoproteinase Glu-C, chymotrypsin, thermolysin, and trypsin) was analyzed. Each of these enzymes has a different sequence specificity; therefore, using them in concert results in complementary information. The use of this combination of proteases results in an even distribution of potential cleavage sites throughout the studied protein, with gaps no longer than 4 residues between adjacent sites. A time course of cleavage with endoproteinase Glu-C is shown in Fig. 1A. An 11-kDa band appeared rapidly and increased in intensity over the observed time; the full-length protein band gradually disappeared over the same period. The intensity of the 11-kDa band was comparable to that of the full-length band, indicating a stoichiometric conversion to a stable product. The cleavage site was mapped to a preferential endoproteinase Glu-C site, C-terminal to Asp\(^{142}\), using N-terminal sequencing. Similar results were obtained using trypsin and chymotrypsin to cleave this protein (data not shown).

To ensure that a minimum domain had been identified, the protein was cleaved sequentially with two different proteases. Fig. 1B shows the digestion with endoproteinase Glu-C followed by chymotrypsin. Before addition of the second protease, only the 11-kDa band was detectable. Chymotrypsin further truncated the fragment, producing the stable product V8/Ch-8. The N and C termini of these fragments were identified unambiguously using matrix-assisted laser desorption ionization-time of flight mass spectrometry. The V8–11 fragment was shown to contain residues 133–226. Chymotrypsin cut after Trp\(^{204}\) of V8/Ch-8 consists of residues 133–204. A similar digestion, carried out with endoproteinase Glu-C and thermolysin, produced a stable product extending from amino acids 133 to 209 (data not shown). Other combinations of enzymes produced consistent results in all cases. From this, we conclude that there is a core domain containing Za. Trp\(^{204}\) is a potential target for cleavage by both chymotrypsin and thermolysin.

Chymotrypsin cut well, but thermolysin cut only marginally at Trp\(^{204}\). Therefore, we define the core domain as comprising Leu\(^{133}\)–Gly\(^{209}\). This core was in no case significantly degraded, whereas the regions on either end were rapidly degraded to pieces too small to detect.

These results were used to design a stable construct, Za, comprising Leu\(^{133}\)–Gly\(^{209}\). This protein was purified from E. coli underdaged under nonnondenaturing conditions. Za showed superior chromatographic behavior over previous Za constructs, purifying from a Mono S cation-exchange column homogeneously as a sharp peak; this indicated structural uniformity. Samples yielded a single band when analyzed by native PAGE.\(^4\) When challenged with exogenous proteases, only Za showed striking stability; other Za constructs were rapidly degraded (data not shown).

The Two Motifs, Za and Zβ, Form a Single Structural Entity—Both Za and Zβ are present in every species in which the sequence of ADAR1 is known. The motifs are separated by one or two copies of a module, weakly conserved in sequence, but consistently lacking positively charged residues and 43–49 amino acids in length. 12 residues from this module are an essential part of Za, the stable Za core domain. It seemed possible that Za, Zβ, and the linker module(s) together form a single structural and functional unit. To investigate this possibility, we examined the structural organization of a peptide spanning both DNA-binding motifs. This peptide, termed Zab, comprising Leu\(^{133}\) (the previously defined N terminus of Za) to Asn\(^{368}\) (C-terminal to Zβ, from human ADAR1), was soluble when overproduced in E. coli, and full-length protein could be obtained with high yield. These results indicate proper folding with no significant instability. Improper folding often leads to the formation of inclusion bodies inside the overproducing bacterial cell.\(^31\) Highly flexible proteins are frequently degraded if expressed in a foreign host.\(^32\)

The results of the digestion of Zab with four different proteases are shown in Fig. 2. Each enzyme cleaved in a characteristic pattern and produced a small number of very stable bands. Time points were selected to allow the identification of all stable products, using mass spectrometry and N-terminal sequencing where appropriate; minor products were identified wherever possible. In each case, well resolved spectra were recorded. Table I lists the peptides produced by each enzyme, as determined from the molecular mass. For chymotrypsin, trypsin, and endoproteinase Glu-C, the assignments are unambiguous and in good agreement with SDS-PAGE analysis. Minor exceptions are fragments Tr8 and Ch5, which were detected only by mass spectrometry, as discussed below. In the case of thermolysin, it was not possible to unambiguously assign the multiple transitory fragments; however, the major fragments seen after 60 min of digestion could be identified. A schematic diagram of the major transitory products and the stable proteolytic fragments is shown in Fig. 3.

Endoproteinase Glu-C cleaved Zab rapidly at a single site, Glu\(^{361}\) at the extreme C terminus (Fig. 2A). The resulting peptide was very stable to further proteolysis, despite an abundance of potential cleavage sites, including Asp\(^{132}\), which is exquisitely sensitive in the shorter construct used to define Za, as described above. After a long incubation with large amounts of enzyme, additional cleavage occurred at Glu\(^{209}\), Glu\(^{380}\), and Leu\(^{367}\). Glu\(^{380}\) lies within the first 49-amino acid repeat; remarkably, the equivalent site in the second repeat, Glu\(^{209}\), was uncut.

Chymotrypsin cleaved the protein after Trp\(^{204}\) and Trp\(^{253}\). These sites are at equivalent positions in the two copies of the

---

\(^2\) K. Lowenhaupt and T. Schwartz, unpublished results.

\(^3\) M. Schade and I. Berger, personal communication.

\(^4\) T. Schwartz, unpublished results.
Proteolytic Dissection of the Z-DNA-binding Domain, Zab

Fig. 2. Digestion of Zab with different proteases. Zab (residues 133–368 from human ADAR1) was incubated with endoproteinase Glu-C (V8) (A), chymotrypsin (B), thermolysin (C), and trypsin (D) at the indicated protease/protein mass ratios for the indicated times. Fragments were resolved by SDS-PAGE (18% gels) and visualized by staining with Coomassie Brilliant Blue. Arrows indicate stable fragments, which were further analyzed by mass spectrometry. The results of this analysis are shown in Table I. Lane M, molecular mass markers.

Table I

| Proteolytic fragment | Measured mass | Sequence assignment | Calculated mass |
|----------------------|---------------|---------------------|-----------------|
|                      | Da            |                     | Da              |
| V8                   | 25,650        | 133L_E            | 25,666          |
| V26                  | 19,756        | 133L_L            | 19,716          |
| V20                  | 18,841        | 133L_K            | 18,850          |
| V19                  | 13,536        | 240D_E            | 13,539          |
| V14                  | 12,173        | 133L_E            | 12,126          |
| Chymotrypsin         |               |                     |                 |
| Ch18                 | 18,082        | 205N_N            | 18,119          |
| Ch13                 | 12,918        | 205N_N            | 12,907          |
| Ch6                  | 7865          | 133L_W            | 7851            |
| Ch5                  | 5189          | 205N_W            | 5230            |
| Thermolysin          |               |                     |                 |
| Ti12                 | 12,239        | 205R_M            | 12,186          |
| Ti8                  | 8495          | 135G_G            | 8460            |
| Trypsin              |               |                     |                 |
| Tr19                 | 19,103        | 133L_K            | 19,102          |
| Tr15                 | 15,309        | 233N_N            | 15,310          |
| Tr12                 | 11,600        | 133L_R            | 11,587          |
| Ty8                  | 7559          | 233N_K            | 7539            |

* Fragment contains 6 additional N-terminal vector-encoded residues.

tandem repeat. The three generated fragments were stable (Fig. 2B). The 5-kDa fragment was not visible on SDS-polyacrylamide gel, although it generated a signal in mass spectrometry of comparable intensity to Ch12 and Ch8. Coomassie Blue staining depends largely on positive charges present in the peptide (33). The 49-amino acid repeat contains only 1 positively charged residue. Therefore, we speculated that although Ch5 was resolved on the gel, it was not stained. Two other transitory fragments, Ch18 and Ch*, were separated on the gel. Ch18 could be assigned to be the product of a single cutting site. Ch* could not be unambiguously determined by mass spectrometry.

Thermolysin produced similar stable products (Fig. 2C). Again, symmetrical sites in the repeated linker, Gly209 and Gly258, were cut, resulting in two stable products on an SDS-polyacrylamide gel. Because thermolysin has low sequence specificity, many transitory products were seen, especially at early time points. Because of these products, it was not possible to unambiguously identify the Ti15 fragment from among several candidates seen by mass spectrometry.

Trypsin attacked the protein at two preferred sites, Arg232 in the first repeat and Lys302 near the N terminus of Zβ (Fig. 2D). (The site equivalent to Arg232 in the second repeat is Ser280, not a substrate for trypsin.) Two sites near the C terminus, Lys366 and Arg367, resulted in heterogeneity of the full-length protein and in the Tr15 fragment. Most of the expected products were stable; however, the C-terminal region peptide, starting at Ile303, was not detected by SDS-PAGE or mass spectrometry. A similar result was seen after extensive endoproteinase Glu-C digestion; again, the C-terminal fragment was not stable. It appears that Zβ, intrinsically more accessible than Za, is stable only in the context of the larger domain.

One of the Two 49-Amino Acid Repeats Can Be Removed without Destabilizing the Domain—In all the ADAR1 genes sequenced, there is a single copy of a 43–49-amino acid linker module between Za and Zβ. In human ADAR1, this module is repeated. To determine the effect of this repeat on the structure of Zab, a protein lacking one module was constructed. This protein, ZahΔl, was produced in high yields as a soluble protein in E. coli and could be purified to homogeneity. Protease mapping showed results similar to those for Zab (data not shown). Trypsin and endoproteinase Glu-C cleaved at identical residues. Chymotrypsin and thermolysin had only a single site each. Therefore, the overall structure of the domain was not altered by the presence of the repeated module.

Zab Consists of Two Ends with Regular Secondary Structure, Connected by an Unconventionally Folded Linker—Circular dichroic measurements in the region between 190 and 250 nm are a useful tool to assess the secondary structure of a protein. This method was used to analyze Za, Zab, and ZahΔl. The spectra are shown in Fig. 4, along with a difference spectrum between Zab and ZahΔl, which reflects the contribution of a single copy of the linker module. Results of the analysis of the curves using the program K2d (34) are shown in Table II. The Za and Zβ motifs contain significant amounts of α-helix and β-sheet structures. In contrast, to a large extent, the linker adopts an alternate structure. This is consistent with the secondary structure analysis of the primary sequence with computer programs, such as PHD (35). These analyses predict that no significant areas in the linker are structured as α-helices or β-sheets. It must be emphasized that the proteolysis studies clearly indicate that the linker module is structurally well defined, although in the majority, neither α-helical nor β-pleated.

Zab Is Protected from Proteolysis When Bound to Its Substrate—The presence of substrate can affect the protease sen-
sensitivity of a protein either because of a direct interference by the substrate molecule or by altering the conformation of the protein. To test whether this is the case for Zab, protease digestions were carried out in the presence of either B-DNA or Z-DNA. Although there were no dramatic changes in the digestion profiles, B-DNA stabilized Zab slightly against proteolysis, and Z-DNA had a very marked stabilizing effect. Chymotrypsin, thermolysin, and trypsin all cut at their established sites, but to a 50-fold lower extent in the presence of Z-DNA (data not shown).

Binding to Z-DNA offered striking protection against digestion with endoproteinase Glu-C (Fig. 5). Although there was no protection of the C-terminal site (Glu361), the internal cleavage sites were strongly protected. Cleavage sites at residues 301 and 307 were completely protected in the presence of Z-DNA, resulting in the absence of the V20 and V19 bands. Cleavage at residue 239 was reduced, with an ~50-fold increase in the

![Diagram of Z-DNA-binding domain of ADAR1](image)

**Fig. 3. Structure and protease cleavage map of the Z-DNA-binding domain of ADAR1.** At the top is a schematic representation of human ADAR1 (hADAR1). Below are the stable fragments produced by limited proteolysis. Numbers above ADAR1 are residue positions. The illustrations are proportional. dsRNA, double-stranded RNA.

![Protein CD spectra](image)

**Fig. 4. Protein CD spectra.** Spectra were recorded as described under "Experimental Procedures" and are expressed in terms of mean residue ellipticity in units of degrees cm$^2$ dmol$^{-1}$. The curves show the protein spectra of Zab (---), ZabΔ1 (— — —), and Za (----). A difference spectrum, Zab minus ZabΔ1, is also shown (—). The corresponding percentages of secondary structure motifs were calculated using the program K2d (34) and are listed in Table II.

**Table II: Secondary structure analysis of recorded CD spectra (Fig. 4)**
The values were calculated using the algorithm K2d (35).

| Protein | α-Helix | β-Sheet | Other |
|---------|---------|---------|-------|
| Za      | 40      | 16      | 44    |
| ZabΔ1   | 37      | 16      | 47    |
| Zab     | 29      | 16      | 55    |
| 49-aa repeat | 8      | 27      | 65    |

* Amino acid.
stability of the full-length Zab protein relative to the absence of DNA. In contrast, B-DNA protected Zab against cleavage only ~5-fold and did not alter the choice of sites. These results suggest that the entire domain becomes more rigid and less accessible in the presence of substrate. The protection of sites within Zab from endoproteinase Glu-C cleavage may occur because these sites are involved in DNA interaction. On the other hand, conformational changes occurring in the protein as a consequence of binding to DNA could prevent endoproteinase Glu-C from cutting. It is of note that the nearby trypsin site, Lys302, was protected in the presence of Z-DNA, but not to the same extent (data not shown).

When Zab in the presence and absence of Z-DNA was compared, there was no change in the CD spectra between 190 and 250 nm (data not shown). This indicates that there is no major change in the secondary structure of the protein when substrate is bound.

The Intact Zab Domain Forms a Stable Complex with Z-DNA and Binds with Sequence Preference—The binding of Za to Z-DNA has been previously characterized using electrophoretic mobility shift assays (2, 6, 7). This assay was used to compare the binding of Za with that of Zab. d(5-BrC-G)20 was used as a substrate; this oligonucleotide is stabilized in the Z-form by the presence of bromine in the 5-position of cytosine (24). Binding was tested in the presence of a 104-fold excess of B-DNA. The results are shown in Fig. 6. Four different proteins were compared at four concentrations (500, 100, 20, and 4 nM protein). Zab bound well at 500 and 100 nM, producing a stable, high molecular mass protein-DNA complex (Fig. 6A, lanes 2–5). At lower concentrations, the complex appeared to break down during electrophoresis, resulting in a smear. This behavior suggests that the most stable complex is formed when the sites on the probe are saturated. ZabΔ showed a similar behavior, although the stable complex was formed only at the highest concentration (Fig. 6A, lanes 14–17). In contrast, Za produced two complex bands, which appeared smeared at all concentrations (Fig. 6A, lanes 6–9). Compared with Zab, Za had a slightly higher affinity for the substrate. The smearing is the result of the instability of the complex under electrophoretic conditions and the longer migration path of Za-DNA complexes as compared with Zab-DNA complexes.

These results may indicate that binding of the Zβ moiety of Zab is responsible for the difference in binding behavior between Za and Zab.

![Fig. 5](image_url) Zab is protected from proteolysis in the presence of Z-DNA. Zab was digested with endoproteinase Glu-C without DNA (lanes 3–5), in presence of B-DNA (lanes 6–8), and in the presence of Z-DNA (poly[d(5-MeC-G)]) (lanes 9–11). Digestion and analysis were performed as described for Fig. 1, except that the protein/protease ratio was 1:30. Lane M, molecular mass markers.

![Fig. 6](image_url) Binding of Z-DNA by Zab and subdomains. A, electrophoretic mobility shift assays were performed with 32P-labeled d(5-BrC-G)20 as the substrate, which is stable in the Z-conformation under the applied conditions (24). Zab (a, lanes 2–5), Za (b, lanes 6–9), Zab digested with chymotrypsin to separate the Za and Zb motifs (c, lanes 10–13), and ZabΔ (d, lanes 14–17) were each assayed in a 5-fold dilution series (500, 100, 20, and 4 nM). Lanes 1 and 18 show the migration of free substrate. Reaction conditions are described under “Experimental Procedures.” The spot at the top of lane 16 is an artifact, which was not reproducible. B, 200 nmol of the protein preparation used for the band shift assay in A were subjected to SDS-PAGE on 18% gels. The bands were visualized by Coomassie Brilliant Blue staining. Proteins are labeled as described for A. The digestion of Zab with chymotrypsin (lane c) was complete, leaving no full-length protein. The difference in the size of Za and the small digestion fragment in lane c is due to 5 additional C-terminal residues present in the Za expression construct. Lane M, molecular mass markers.
and Zab. To test this hypothesis, Zab was digested with chymotrypsin and then assayed in the band shift (Fig. 6A, lanes 10–13). Complete digestion, yielding the Za and Zβ motifs as separate peptides, was confirmed by SDS-PAGE (Fig. 6B). This mixture showed a band shift pattern very similar to that for Za. No additional bands were observed; therefore, Zβ alone does not bind to the substrate. Since the molecular masses of the Za- and Zβ-containing fragments differ substantially, it is extremely unlikely that any complex formed by Zβ and DNA would comigrate with the observed Za-DNA complexes (23). That the isolated Zβ is not capable of binding Z-DNA under these conditions is remarkable considering the conservative substitution of functionally important residues (2).

As a second method of studying the binding of Za to DNA, circular dichroism was used to monitor the transition of the DNA conformation from the B- to the Z-form (2, 4, 5). The spectrum of Z-DNA is inverted as compared with that of B-DNA in the near-UV region between 240 and 300 nm (36, 37). Fig. 7 shows the spectra of two Z-DNA-forming oligomers of different sequence, d(C-G)₆ and d(C-A)₇, in the presence of either Za or Zab. The DNAs adopted the right-handed B-DNA conformation in the absence of protein. Protein was added in aliquots, resulting in protein/base pair molar ratios of 1:6 (-- - -), 1:4 (-- - - -), 1:2 (-- - - - -), and 1:1.5 (-- - - - - -), respectively. Spectra are expressed in absolute values of ellipticity in millidegrees (mdeg).

**DISCUSSION**

Although the exact biological role of ADAR1 has not been established, analysis of the amino acid sequence has allowed the assignment of functions to parts of the protein (38). These functions contribute to the known *in vitro* activity of this enzyme, the conversion of A to I in regions of double-stranded RNA. The central series of double-stranded RNA-binding motifs and the C-terminal catalytic domain are well characterized protein domains. The N-terminal region, on the other hand, contains a novel motif, Za, with a binding specificity for Z-DNA in a sequence-independent manner. However, when Za was in the context of the entire domain, Zab, a sequence preference for d(C-G)₆ was observed. Band shift data suggests that the mode of binding is different between Za and Zab, probably reflecting a difference in the degree of cooperativity.

**Fig. 7.** CD studies of the conformational change of Z-forming sequences in the presence of Za or Zab. Traces were generated as described under “Experimental Procedures.” The spectra show the titration of d(C-G)₆ (A and C) and d(C-A)₇-d(T-G)₃ (B and D) with Zab (A and B) and Za (C and D), respectively. The curves represent the spectra of the DNA alone (→) and in presence of protein at protein/base pair molar ratios of 1:6 (-- - -), 1:4 (-- - - -), 1:2 (-- - - - -), and 1:1.5 (-- - - - - -), respectively. Spectra are expressed in absolute values of ellipticity in millidegrees (mdeg).
smaller core domain, Za, from Leu133 to Gly209, contains Za and a portion of the linker. C-terminally further shortened Za peptides are functional in binding specifically to Z-DNA (2). However, they lack the structural uniformity seen only for the proteolytically defined domain Za. These earlier constructs also contain additional N-terminal residues (positions 121–132). These residues have been reported to modulate the results of band shift assays (5); however, CD experiments are unaffected by their presence. Although Za is stable to proteolysis, we conclude that Za is the functional domain. Domain boundaries are frequently protease-hypersensitive, with cleavage sites for different proteases clustered in close proximity. This is observed both N-terminal to Leu133 and C-terminal to Glu361. In contrast, Za produces a less stable product, although it is not clear that Za and Zβb need interact within the Zβa domain to form a single binding site, involving both motifs. The results obtained by CD measurements strongly support this conclusion. Za binds with sequence preference for alternating d(C-G)₅, whereas Za does not discriminate between Z-forming sequences, but rather is conformation-specific. This is in agreement with previous studies on Za (121–201) (5). The preference of Za for d(C-G)₅ needs further investigation. Identification of the optimal substrate in vitro may elucidate the role of Za as part of ADAR1 in vivo and lead to the identification of the actual binding sites of ADAR1 on chromatin.

Acknowledgments—We thank Marina Voulova for help in the proteolytic digestion experiments. We especially thank Dr. Stefan Maas for many helpful discussions and for carefully reading the manuscript.

REFERENCES

1. Lilley, D. M. J. (1995) DNA-Protein: Structural Interactions, Oxford University Press, Oxford

2. Herbert, A., Alfenken, J., Kim, Y.-G., Mian, I. S., Nishikura, K., and Rich, A. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 8421–8426

3. Kim, Y.-G., Kim, P. S., Herbert, A., and Rich, A. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 12975–12980

4. Berger, I., Winston, W., Manoharan, R., Schwartz, T., Alfenken, J., Kim, Y.-G., Lowenhaupt, K., Herbert, A., and Rich, A. (1998) Biochemistry 37, 13213–13221

5. Herbert, A., Schade, M., Lowenhaupt, K., Alfenken, J., Schwartz, T., Skyllakhtenko, L. S., Lyubchenko, Y. L., and Rich, A. (1998) Nucleic Acids Res. 26, 3486–3493

6. Herbert, A. G., Spitzer, J. R., Lowenhaupt, K., and Rich, A. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 3339–3342

7. Herbert, A., Lowenhaupt, K., Spitzer, J., and Rich, A. (1996) Proc. Natl. Acad. Sci. U. S. A. 92, 7550–7554

8. Patterson, J. B., and Samuel, C. E. (1995) Mol. Cell. Biol. 15, 5376–5388

9. Melcher, T., Maas, S., Herb, A., Sprengel, R., Seeburg, P. H., and Higuchi, M. (1996) Nature 379, 660–664

10. Melcher, T., Maas, S., Herb, A., Sprengel, R., Higuchi, M., and Seeburg, P. H. (1996) J. Biol. Chem. 271, 31795–31798

11. Maas, S., Melcher, T., and Seeburg, P. H. (1997) Curr. Opin. Cell Biol. 9, 343–349

12. Verdoorn, T. A., Burnashev, N., Monyer, H., Seeburg, P. H., and Sakmann, B. (1991) Science 252, 1715–1718

13. Hume, R. I., Dingledine, R., and Heinemann, S. F. (1991) Science 253, 1028–1031

14. Sommer, B., Kohler, M., Sprengel, R., and Seeburg, P. H. (1991) Cell 67, 11–19

15. Kohler, M., Burnashev, N., Sakmann, B., and Seeburg, P. H. (1993) Neuron 10, 191–200

16. Loenli, H., Mosbacher, J., Melcher, T., Heger, T., Geiger, J. R., Kuner, T., Monyer, H., Higuchi, M., Bach, A., and Seeburg, P. H. (1994) Science 266, 1709–1713

17. Petschek, J. P., Mermer, M. J., Scheckelhoff, M. R., Simone, A. A., and Vaugham, J. C. (1996) J. Mol. Biol. 259, 885–890

18. Burns, C. M., Chu, H., Ruster, S. M., Hutchinson, L. K., Canton, H., Sanders-Bush, E., and Emeson, R. B. (1997) Nature 387, 305–308

19. Seeburg, P. H. (1996) J. Neurochem. 66, 1–5

20. Herbert, A., and Rich, A. (1996) J. Biol. Chem. 271, 11595–11598

21. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

22. Gill, S. C., and von Hippel, P. H. (1989) Anal. Biochem. 182, 319–326

23. Lane, D., Pretkki, P., and Chandler, M. (1992) Microbiol. Rev. 56, 509–528

24. Malfoy, B., Rousseau, N., and Leng, M. (1982) Biochimie 64, 543–547

25. Janin, C., and Chothia, C. (1985) Methods Enzymol. 115, 420–430

26. Schulz, G., and Schirmer, R. (1979) Principles of Protein Structure, Springer-Verlag New York Inc., New York

27. Porter, R. (1973) Science 180, 713–716

28. Jovin, T., Geisler, N., and Weber, K. (1977) Nature 269, 668–672

29. Roy, R., Kumar, A., Lee, J. C., and Mitra, S. (1996) J. Biol. Chem. 271, 23690–23697

30. Nakagawa, N., Masui, R., Kato, R., and Kuramitsu, S. (1997) J. Mol. Biol. 272, 22703–22713

31. Wilkinson, D. L., and Harrison, R. G. (1989) J. Biol. Chem. 264, 55–72

32. Tal, M., Silverstein, A., and Russe, E. (1985) J. Biol. Chem. 260, 9976–9980

33. Andrade, M. A., Chao, P., Merelo, J. J., and Moran, F. (1993) Protein Eng. 6, 383–390

34. Rost, B., and Sander, C. (1994) Proteins 19, 55–72

35. Pohl, F. M., and Jovin, T. M. (1977) J. Mol. Biol. 106, 375–396

36. Riazance-Lawrence, J. H., and Johnson, W. C., Jr. (1992) Cell 70, 271–276

37. Kim, U., Wang, Y., Sanford, T., Zeng, Y., and Nishikura, K. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 11457–11461

38. Schade, M., Turner, C. J., Lowenhaupt, K., Rich, A., and Herbert, A. (1999) EMBO J., in press