Subsets of the Zinc Finger Motifs in dsRBP-ZFa Can Bind Double-Stranded RNA†

Patrick J. Finerty, Jr.‡ and Brenda L. Bass*

Department of Biochemistry and Howard Hughes Medical Institute, University of Utah, 50 North Medical Drive, Salt Lake City, Utah 84132

Received October 22, 1998; Revised Manuscript Received January 21, 1999

ABSTRACT: dsRBP-ZFa is a *Xenopus* zinc finger protein that binds dsRNA and RNA−DNA hybrids with high affinity and in a sequence-independent manner. The protein consists of a basic N-terminal region with seven C_2H_2 zinc finger motifs and an acidic C-terminal region that is not required for binding. The last four zinc finger motifs, and the linkers that join them, are nearly identical repeats, while the first three motifs and their linkers are each unique. To identify which regions of the protein are involved in nucleic acid binding, we examined the ability of five protein fragments to bind dsRNA and RNA−DNA hybrids. Our studies reveal that a fragment encompassing the three N-terminal, unique zinc finger motifs and another encompassing the last three of the nearly identical motifs have binding properties similar to the full-length protein. Since these two fragments do not share zinc finger motifs of the same sequence, dsRBP-ZFa must contain more than one type of zinc finger motif capable of binding dsRNA. As with the full-length protein, ssRNA and DNA do not significantly compete for dsRNA binding by the fragments.

Zinc finger proteins (ZFPs)[1] of the C_2H_2 type represent one of the most common nucleic acid binding motifs found in nature (*1*). Proteins containing these motifs are generally DNA-binding transcription factors that recognize specific sequences in the context of the B-form helix. However, some ZFPs are also able to bind single-stranded RNA (ssRNA; *2*, *3*), double-stranded RNA (dsRNA; *4*), and RNA−DNA hybrids (*4*, *5*). TFIIIA, the founding member of the C_2H_2 zinc finger family, binds both 5S rRNA and the internal control region (ICR) of the 5S rRNA gene. However, it does not bind an RNA−DNA hybrid that is the same sequence as the 5S rRNA gene ICR, indicating that RNA−DNA hybrid-binding is not a property of all C_2H_2 ZFPs (*6*).

We previously reported the identification and characterization of the dsRNA-binding ZFP cited above (*4*). This protein, double-stranded RNA-binding protein ZFa (dsRBP-ZFa), is the first ZFP observed to bind dsRNA and RNA−DNA hybrids with high affinity. dsRBP-ZFa is a 55.6 kDa protein, originally isolated from a *Xenopus* cDNA expression library by screening with radiolabeled dsRNA. The biological role of dsRBP-ZFa is currently unknown, although the protein is known to localize to the nucleus of *Xenopus* oocytes (*4*).

dsRBP-ZFa can be divided into two main regions: an N-terminal region (residues 1−432) that includes seven zinc finger motifs, and a C-terminal region (residues 433−524) that is rich in serine, proline, and glycine. The C-terminal region of the protein is not necessary for dsRNA binding (*4*), lending support to the idea that the zinc fingers, or their linkers, mediate binding to dsRNA. The zinc fingers in dsRBP-ZFa generally follow the motif $Cx_2Cx_3\Phi x_5\Phi x_2Hx_5H$, where Φ is a hydrophobic residue and x is any amino acid. The first three zinc fingers clearly differ in sequence from each other, as well as from the last four fingers. In addition to being separated from the N-terminal fingers by a short region containing 5 proline residues, the last 4 fingers are nearly identical repeats of 58 amino acids with 91−100% sequence identity to each other. In contrast to most zinc fingers, which are joined by a conserved 6−8 amino acid linker (*1*, *7*), the zinc fingers in dsRBP-ZFa are joined by much longer linkers, which range from 34 to 44 amino acids.

Two ZFPs besides dsRBP-ZFa, the transcription factor Sp1, and ZFQQR, a protein of synthetic design, have been shown to bind RNA−DNA hybrids; these proteins, however, are sequence-specific (*5*). Furthermore, Sp1 and ZFQQR primarily recognize the DNA strand of the RNA−DNA duplex and show negligible binding to a dsRNA substrate that is the same sequence (*5*). While Sp1 and ZFQQR recognize DNA and RNA−DNA hybrids by making sequence-specific contacts in the major groove, dsRBP-ZFa is likely limited to binding in the minor groove, since the major groove of an A-form helix is very deep and quite narrow (*8*).

Given the differences between dsRBP-ZFa and previously characterized ZFPs, it is difficult to predict how dsRBP-ZFa binds A-form helices. One possibility is that dsRBP-ZFa is able to bind both dsRNA and RNA−DNA hybrids

† This study was supported by funds to B.L.B. from the David and Lucile Packard Foundation. B.L.B. is an HHMI Associate Investigator. Oligonucleotides were synthesized by the Howard Hughes Medical Institute oligonucleotide synthesis facility at the University of Utah supported by the Department of Energy (Grant DE-FG03-94ER61817) and HHMI; dsRBP-ZFa fragments were sequenced by the University of Utah Health Sciences DNA Sequencing Facility supported in part by the National Cancer Institute (Grant 5P30CA42014).

* Corresponding author: Department of Biochemistry and Howard Hughes Medical Institute, University of Utah, 50 N. Medical Dr., Salt Lake City, UT 84132. E-mail: bbass@howard.genetics.utah.edu. Phone: (801) 581-4884. Fax: (801) 581-5379.

‡ Current address: Structural Biology and Biochemistry, Hospital for Sick Children, 555 University Ave., Toronto, Ontario, Canada, M5G 1X8.

[1] Abbreviations: dsRNA, double-stranded RNA; dsRBP-ZFa, dsRNA-binding protein ZFa; ss, single-stranded; ZFP, zinc finger protein; rRNA, ribosomal RNA; TFIIIA, transcription factor IIIA; PCR, polymerase chain reaction; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; PAG608, p53-activated gene 608.

10.1021/bi9825293 CCC: $18.00
Published on Web 03/06/1999

with high affinity by using all seven zinc finger motifs equivalently. Alternatively, different fingers could have different substrate preferences and contribute different amounts of binding energy. On a superficial level, this seems to be the case with TFIIIA: the first three fingers are sufficient for specific, high-affinity binding to the 5S rRNA gene, and the middle three sufficient for tight binding to 5S rRNA (*9*). However, analyses of the TFIIIA·ICR complex that forms with the full-length protein show that all fingers make specific contacts with the DNA and thus contribute some binding energy (*10*−*12*). Still, subsets of the TFIIIA fingers seem to provide the bulk of the binding energy for interaction with DNA, while others are of primary importance for specificity of the 5S rRNA interaction.

To gain insight into how dsRBP-ZFa recognizes its substrates, we have analyzed the binding properties of dsRBP-ZFa protein fragments encompassing three to four zinc finger motifs each. Because the first three motifs are different from the last four, we were interested if either subset would bind dsRNA or RNA−DNA hybrids alone, and, if so, what differences there might be in their binding affinity and substrate preferences. We show that a fragment encompassing the N-terminal three unique motifs and also a fragment encompassing the C-terminal repeated zinc finger motifs bind dsRNA and RNA−DNA hybrids with high affinity. However, the fragments show slightly different affinities and substrate preferences. Our studies indicate that dsRBP-ZFa contains multiple examples of zinc fingers that bind dsRNA, suggesting studies of this protein may provide paradigms for how zinc fingers can bind the A-form helix.

MATERIALS AND METHODS

Oligonucleotides and Plasmids. Oligonucleotides used as primers in the PCR were not gel-purified. The oligonucleotides used were: (zf-2-zifi) 5′-GAACCATTTAAAGCATATGACACCGTTGTCTGAG-3′; (zf-3-zifi) 5′-GAATCAATTAATGCATATGACACCATTGTCTGAG-3′; (zf-4-zifi) 5′-GAATTTAAACTCTCATATGCCAGGCTCAGGATCA-3′; (pET-c-term) 5′-GGGTTATGCTAGTTATTGCTC-3′; (zf-F9-Bam) 5′-GGGGATCCCTATATAGTCATGATCTGTGTTTTGG-3′; (T7) 5′-TAATACGACTCACTATAGGG-3′.

The T7-based expression vector pET-16b (Novagen) was used for expression of all proteins. Plasmids were purified using CsCl/ethidium bromide equilibrium centrifugation as described (*13*). Expression constructs used in this study were sequenced by the University of Utah Health Sciences Sequencing Facility.

Five different expression constructs were built; one encompassed the first three finger motifs, and four encompassed the last four, three, two, or final zinc fingers from dsRBP-ZFa (Genbank accession number AF005083). pET-16b DNA was digested with *Nde*I (Boehringer Mannheim) and *Bam*HI (NEB) and the 5701 bp fragment gel-purified for use in ligation reactions (below). PCR products specific for each protein were generated using dsRBP-ZFaΔpro (*4*) as a template, digested with *Nde*I and *Bam*HI, ligated to the 5701 bp pET-16b *Nde*I−*Bam*HI fragment, and used to transform DH5α cells. All expressed proteins had an N-terminal 21 amino acid tag (His-tag) derived from pET-16b (MGHHHHHHHHHHSSGHIEGRH).

pET-zf-123, residues 1−198 of dsRBP-ZFa in pET-16b, was generated using primers F9-Bam and T7. zf-123 is 219 amino acids in length. pET-zf-4567, residues 219−435 of dsRBP-ZFa in pET-16b, was generated using primers zf-4-zifi and pET-c-term. zf-4567 is 238 amino acids in length. pET-zf-567, residues 260−435 of dsRBP-ZFa in pET-16b, was generated using primers zf-3-zifi and pET-c-term. zf-567 is 197 amino acids in length. pET-zf-67, residues 317−435 of dsRBP-ZFa in pET-16b, was generated using primers zf-2-zifi and pET-c-term. zf-67 is 140 amino acids in length. pET-zf-7, residues 374−435 of dsRBP-ZFa in pET-16b, was also generated using primers zf-2-zifi and pET-c-term. zf-7 is 83 amino acids in length.

Expression and Purification of dsRBP-ZFa Protein Fragments. To facilitate protein expression, all constructs were transformed into BL21(DE3) cells (Novagen). Protein expression and purification were similar for all proteins except zf-123, which required additional purification. Protein was overexpressed as described (pET manual) using 2 L of media. SDS−PAGE analysis of lysed cells indicated that overexpressed proteins were primarily located in the insoluble fraction. Cells were resuspended in HMGK buffer [25 mM HEPES (pH 8.3), 10% (v/v) glycerol, 12.5 mM $MgCl_2$, 100 mM KCl, 0.1% (v/v) NP-40] containing 0.2 mg mL^{-1} lysozyme and stirred at 4 °C. After 20 min, deoxycholic acid was added to 1.2 mg mL^{-1} and the solution stirred 5 min. Following lysis, the mixture was sonicated to shear chromosomal DNA, the insoluble material isolated by centrifugation at 40000*g*, and the pellet washed twice with HMGK buffer.

The pellet was dissolved in a buffer containing 40 mM HEPES (pH 8.3) and 8 M guanidine hydrochloride by stirring at 25 °C. Following centrifugation at 40000*g* to remove material that did not dissolve, NaCl was added to 0.5 M and imidazole to 5 mM, and the solution was diluted to 6 M guanidine hydrochloride with 1× binding buffer [32 mM HEPES (pH 8.3), 0.5 M NaCl, 5 mM imidazole] and brought to a final volume of 200 mL. The solution was further clarified by passage through a 0.2 μm filter before being subjected to Ni^{2+} affinity chromatography. Ni^{2+} chromatography was carried out as described for dsRBP-ZFaΔpro (*4*) except that buffers did not contain urea or guanidine hydrochloride, so that proteins were refolded on the column. Fractions containing overexpressed protein were pooled, and the protein was precipitated by addition of $(NH_4)_2SO_4$ to 3 M and pelleted by centrifugation at 13000*g*. Ni^{2+} was removed by washing the pellet twice with 3 M $(NH_4)_2SO_4$, 10 mM EDTA and twice more with 3 M $(NH_4)_2SO_4$ without EDTA. The pellet was dissolved in acetate buffer [50 mM sodium acetate (pH 5.0), 20 mM KCl, 1 mM $MgCl_2$, 100 μM $ZnCl_2$] and 8 M guanidine hydrochloride, dialyzed against 5 × 4 L of acetate buffer + 5 mM DTT, and centrifuged at 50000*g* to remove insoluble material.

At this point all proteins except zf-123 were sufficiently pure for binding studies. After Ni^{2+} chromatography, the zf-123 preparation contained three major species. Following Ni^{2+} removal (as above), the zf-123 pellet was dissolved in EKMDZ-9.5 buffer [50 mM ethanolamine (pH 9.5), 5 mM KCl, 1 mM $MgCl_2$, 5 mM DTT, 100 μM $ZnCl_2$] + 8 M guanidine hydrochloride, and dialyzed against EKMDZ-9.5 buffer to refold the protein. During dialysis a significant fraction of the protein precipitated. Analysis of the soluble

and insoluble fractions showed that the insoluble fraction consisted primarily of full-length zf-123 while the soluble fraction contained the smaller proteins as well as some full-length protein (data not shown). The soluble fraction was discarded and the insoluble fraction dissolved in acetate buffer + 8 M guanidine hydrochloride and 100 mM DTT and further purified by two passes through a 370 mL HW-50F (TosoHaas) gel filtration column (70 cm × 2.6 cm) at 1.2 mL min^{-1}. Fractions containing zf-123 protein were pooled.

All proteins were concentrated using a stirred ultrafiltration cell (Amicon, Model 8050). zf-123, zf-4567, and zf-567 were concentrated using YM3 Diaflo ultrafiltration membranes (Amicon). zf-67 and zf-7 were concentrated using YM1 Diaflo ultrafiltration membranes (Amicon). After concentration, proteins were 0.2 μm filtered, and the purity was determined by densitometry of a SYPRO Red (Molecular Probes) stained gel (Figure 1B) using a Storm PhosphorImager (Molecular Dynamics) and the relative amounts of each species determined using ImageQuaNT software (Molecular Dynamics). The percentage of full-length protein for each preparation was as follows: zf-123 = 69%, zf-4567 = 67%, zf-567 = 76%, zf-67 = 89%, zf-7 = 87%.

The protein concentration for each preparation was determined by the absorbance at 280 nm using the following predicted extinction coefficients: zf-123, 1 A_{280} = 4.1 mg mL^{-1}; zf-4567, 1 A_{280} = 4.1 mg mL^{-1}; zf-567, 1 A_{280} = 4.6 mg mL^{-1}; zf-67, 1 A_{280} = 4.9 mg mL^{-1}; zf-7, 1 A_{280} = 5.9 mg mL^{-1}. The final concentration used to prepare dilutions for mobility shift assays was obtained by multiplying the concentration obtained by spectrophotometry by the purity determined by densitometry. Proteins were stored as 45% glycerol stocks at −20 °C.

Mobility Shift Assays. All aspects of the mobility shift assays, including preparation of substrates and competitors, experimental procedures for binding assays and competition studies, and analysis of the binding data, were performed as described (*4*, and references cited therein). K_d values are the result of three independent experiments except for the second value listed for zf-123 (end) which is the result of two independent experiments.

Sequence Analysis. Predicted extinction coefficients were obtained with the LaserGene package from DNASTAR. Searches of public databases were conducted using the Linux versions of blastp and tblastn provided by the National Center for Biotechnology Information (*14*). The accession numbers for the sequences with significant similarity to dsRBP-ZFa, besides PAG608, are AA317148, AA699468, AA393154, G15698, AA448922, and R55900.

RESULTS

Cloning and Expression of dsRBP-ZFa Fragments. Five fragments of dsRBP-ZFa coding sequence were generated using PCR, cloned into the *Escherichia coli* T7-based expression vector pET-16b (Novagen), and subsequently used for protein expression. The protein fragments were named according to the zinc finger motifs from dsRBP-ZFa they include. For instance, zf-123 contains the region of dsRBP-ZFa encompassing the 3 most N-terminal zinc finger motifs (the N-terminal 198 residues of dsRBP-ZFa). The other proteins, zf-4567 through zf-7, are successively smaller

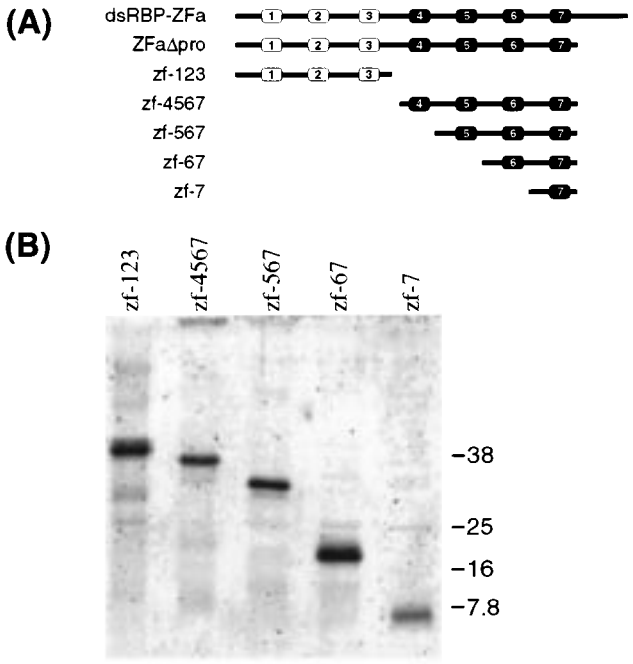

FIGURE 1: dsRBP-ZFa protein fragments. (A) The full-length protein (dsRBP-ZFa), a C-terminal truncation (dsRBP-ZFaΔpro), and various protein fragments are shown schematically. The zinc finger motifs are shown as rectangles and numbered, 1−7, N-terminal to C-terminal. The last four zinc fingers, indicated by the filled rectangles, are nearly identical repeats while each of the first three are unique. (B) Following purification and concentration, each protein preparation was separated by SDS−16% PAGE. A PhosphorImager image of the SYPRO Red stained gel is shown, and positions of molecular mass markers in kDa are indicated to the right of the gel. The purity of each protein was determined by densitometry to be 69% for zf-123, 67% for zf-4567, 76% for zf-567, 89% for zf-67, and 87% for zf-7.

fragments of the four C-terminal zinc finger motifs and surrounding residues (see Materials and Methods). An illustration of the various protein fragments is shown in Figure 1A. In all overexpressed proteins, the first 21 residues are derived from the His-tag in pET-16b.

For all proteins, except zf-7, the majority of overexpressed protein was found in the insoluble fraction (data not shown). We had successfully obtained active protein following refolding from guanidine hydrochloride in the past, so this method was employed with these proteins as well. Additionally, except for zf-123, overexpression resulted in the production of a single species, and Ni^{2+} affinity chromatography alone yielded protein preparations that were at least 67% pure (Figure 1B). Following Ni^{2+} affinity chromatography, the zf-123 preparation was found to contain three major species (data not shown), so it was subjected to additional purification procedures (see Materials and Methods). Purified proteins (Figure 1B) were stored as 45% glycerol stocks at −20 °C.

Mobility Shift Assays. Full-length dsRBP-ZFa is able to bind dsRNA and RNA−DNA hybrids in a sequence-independent manner with high affinity (0.5 nM and 1.7 nM, respectively) (*4*). In preliminary binding studies, we observed that zf-67 did not bind dsRNA until ∼1 μM protein concentrations, and zf-7 showed negligible binding even at ∼10 μM protein (data not shown). Since we were primarily interested in determining the region of dsRBP-ZFa responsible for high-affinity binding, these proteins were not further characterized.

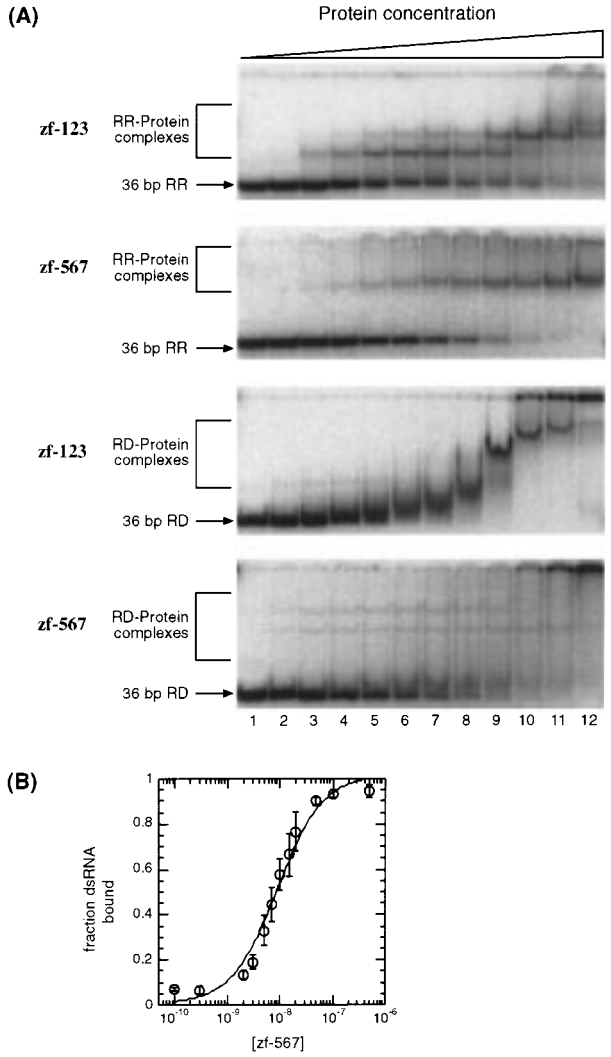

FIGURE 2: Mobility shift analyses of zf-123 and zf-567 binding to 36-mer dsRNA and 36-mer RNA−DNA hybrid. (A) A 36 bp dsRNA (RR, 10 pM) or a 36 bp RNA−DNA hybrid (RD, 10 pM) was incubated with increasing concentrations of zf-123 or zf-567. The protein concentrations for lanes 1−12 of the zf-123 + RR reactions are, respectively, 10 pM, 50 pM, 150 pM, 400 pM, 600 pM, 800 pM, 1.1 nM, 1.5 nM, 3 nM, 5 nM, 30 nM, 100 nM. The protein concentrations for lanes 1−12 of the zf-567 + RR reactions are, respectively, 100 pM, 300 pM, 2 nM, 3 nM, 5 nM, 7 nM, 10 nM, 15 nM, 20 nM, 50 nM, 100 nM, 500 nM. The protein concentrations for lanes 1−12 of the zf-123 + RD reactions are, respectively, 370 pM, 1 nM, 2 nM, 4 nM, 7 nM, 12 nM, 20 nM, 40 nM, 100 nM, 200 nM, 400 nM, 1 μM. The protein concentrations for lanes 1−12 of the zf-567 + RD reactions are, respectively, 3.7 nM, 15 nM, 30 nM, 60 nM, 98 nM, 170 nM, 270 nM, 500 nM, 1 μM, 2 μM, 4 μM, 10 μM. The electrophoretic mobilities of free and bound substrate are indicated to the left of these representative PhosphorImager images. The band appearing in all samples near the top of the autoradiogram derives from material caught in the well of the gel; radioactivity from this band was included in the value calculated for the total radioactivity of the lane (see Materials and Methods and refs *4* and *17*). (B) K_d determination. Primary data for zf-567 binding to 36-mer dsRNA were generated in independent experiments and quantified. Each point in this typical graph is an average of three experiments and error bars represent ± standard deviation from the mean. Data from other binding reactions were similarly analyzed (see Materials and Methods).

We determined the ability of zf-123, zf-4567, and zf-567 to bind dsRNA and RNA−DNA hybrids using a gel mobility shift assay (Figure 2A). Two different substrates were used in the mobility shift assays: a 36-mer dsRNA and a 36-mer RNA−DNA hybrid, both derived from the same sequence. 36-mer is not related to any biologically relevant sequence. Mobility shift assays were performed by mixing increasing amounts of purified protein with a ^{32}P-labeled substrate followed by electrophoresis on a native 6% polyacrylamide gel. zf-4567 and zf-567 had very similar binding properties, so only data for zf-567 are shown, and for convenience, these proteins are sometimes referred to together as zf-(4)567.

Table 1: Summary of Binding Data

protein	RR[a]	RD[b]	RD/RR[c]
dsRBP-ZFa[d]	5×10^{-10}	1.7×10^{-9}	3.4
zf-123 (beg)[e]	4.9×10^{-10}	–	–
zf-123 (end)[f]	1.5×10^{-9}	8.6×10^{-9} [g]	5.7
zf-4567	8.4×10^{-9}	1.6×10^{-7}	19.1
zf-567	9.0×10^{-9}	8.6×10^{-8}	9.6

[a] 36 bp dsRNA. [b] 36 bp RNA−DNA hybrid. [c] The result of dividing the K_d obtained with the RNA−DNA hybrid by the K_d obtained with dsRNA. [d] From (*4*). [e] The K_d value initially obtained, [f]and following a loss in binding activity. [g] The K_d value for zf-123 binding to the RNA−DNA hybrid was determined after the loss in activity.

Using the data from the 36-mer binding assays, apparent equilibrium dissociation constants (K_d) were determined for each protein and substrate combination (Figure 2B, values listed in Table 1). Our results show that both the N-terminal zinc finger motif region and the C-terminal zinc finger motif region of dsRBP-ZFa are able to bind dsRNA and RNA−DNA hybrids with high affinity. For all fragments, dsRNA was bound with a higher affinity than RNA−DNA hybrids. Additionally, zf-(4)567 binds less well to both substrates than zf-123 (Table 1). For all fragments, gel shift analyses with RNA−DNA hybrids showed poorly resolved complexes, probably because the lower affinity binding resulted in complexes that were unstable during electrophoresis. Note that this did not affect measurement of the K_d values since these were determined by quantifying the disappearance of free substrate.

The K_d values were based on the assumption that the protein used in the assays was fully active, but we do not know if this is true. Over the space of time in which the binding studies were conducted, the activity of the zf-123 protein preparation decreased approximately 3-fold. As such, we have reported two K_d values for this protein. The K_d values reported for the binding of zf-123 to RNA−DNA hybrids were measured after the decrease in activity, and, thus, comparisons of substrate binding between the various fragments should use the second K_d value reported for zf-123 (Table 1, end). The activity of zf-(4)567 remained constant through the course of binding studies.

Substrate Preference. We further characterized the binding properties of zf-123 and zf-(4)567, by using tritiated or nonradioactive dsRNA, ssRNA, RNA−DNA, and DNA substrates, all derived from a 100-mer sequence, as competitors of binding to ^{32}P-labeled 36-mer dsRNA. The various 100-mer substrates were derived from the same sequence and have no relationship to 36-mer or to any naturally occurring sequences. The ability of each 100-mer substrate to compete at 1×, 10×, and 100× molar excess over 36-mer dsRNA was measured in gel mobility shift assays (Figure 3A). These data show that, for zf-123, dsRNA and RNA−DNA hybrids compete equivalently, while dsRNA competes better than RNA−DNA hybrids for zf-(4)567

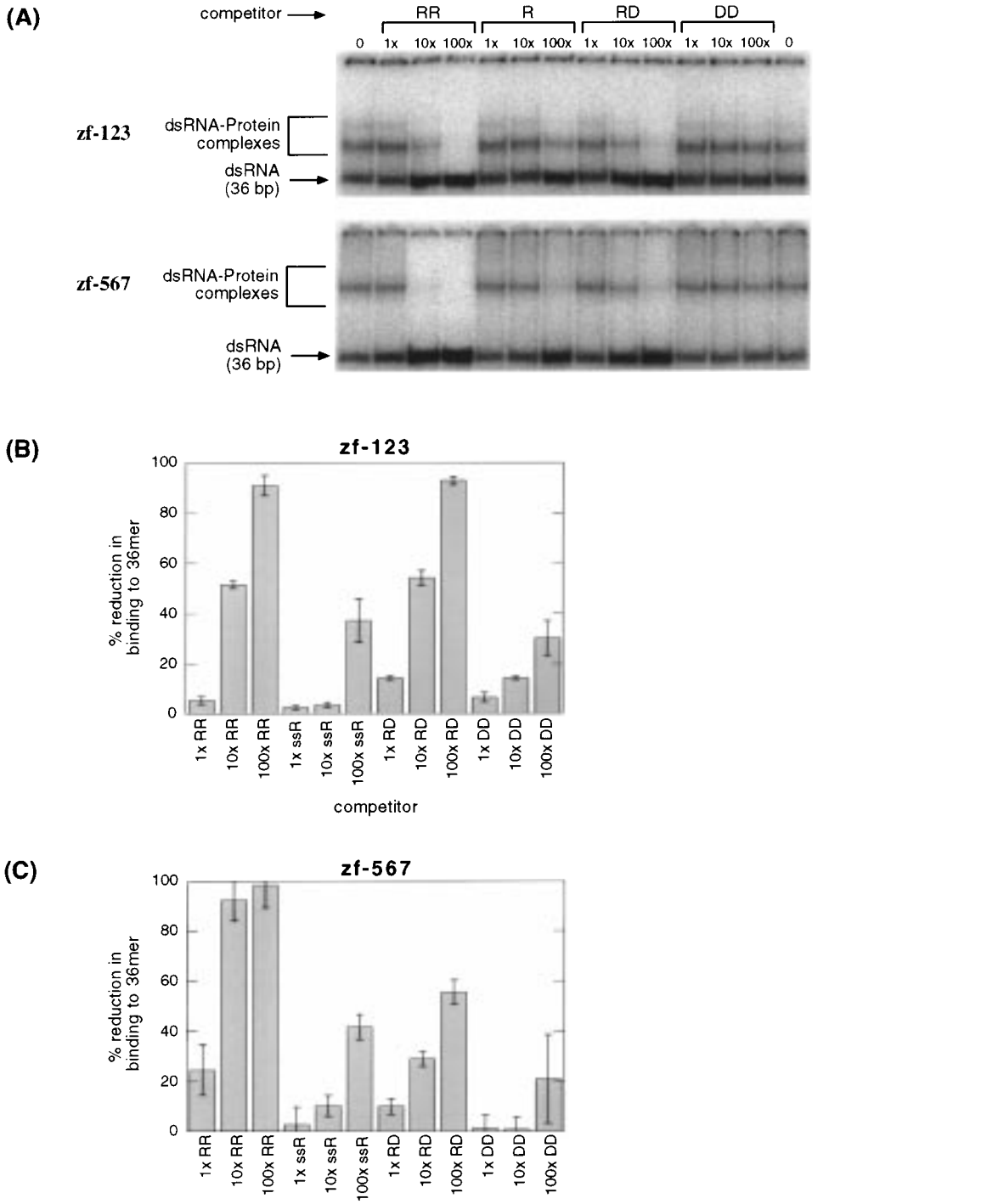

FIGURE 3: Competition studies with zf-123 and zf-567. (A) PhosphorImager images of two representative mobility shift competition assays are shown. The ability of four substrates (RR = 100 bp dsRNA, R = 100 base ssRNA, RD = 117 bp RNA−DNA hybrid, DD = 143 bp dsDNA) to compete for either zf-123 or zf-567 binding to radiolabeled 36-mer dsRNA (10 pM) was assayed. zf-123 was present at 1.5 nM and zf-567 at 9.0 nM in all lanes, and competitors were added at 1×, 10×, and 100× the concentration of 36-mer dsRNA. (B and C) The ability of each substrate to compete was determined as in (A) in three experiments and the average shown graphically. Error bars represent ± standard deviation from the mean.

binding (Figure 3B,C). For all three proteins, ssRNA and DNA do not compete significantly until present at 100× molar excess.

DISCUSSION

dsRBP-ZFa is the only characterized ZFP known to bind dsRNA with high affinity. On the basis of previous results, and the fact that the first three zinc finger motifs in dsRBP-ZFa are different from the last four, we examined the nucleic acid binding properties of five different fragments of dsRBP-ZFa. We showed that fragments of dsRBP-ZFa encompassing the three N-terminal unique zinc finger motifs, and another encompassing the last three of the nearly identical motifs, bind dsRNA and RNA−DNA hybrids with high affinity (Figure 2A). The fragments have K_d values similar to that for the full-length protein and, like the full-length protein, do not significantly bind ssRNA or DNA (Figure 3A). However, zf-123 and zf-(4)567 differ in their relative

affinities for dsRNA and RNA−DNA hybrids (see below). Although K_d values were not determined for the proteins with one and two finger motifs, the C-terminal region appears to require at least three of the four zinc fingers for high-affinity binding (nM). Future studies to determine how many of the three unique N-terminal motifs are sufficient for dsRNA binding will be informative.

The partial cDNA clone initially isolated in our expression library screen encoded only the five C-terminal zinc finger motifs (*4*). As such, when designing dsRBP-ZFa fragments, we concentrated on this region of the protein since it seemed most likely to be involved in binding dsRNA. Additionally, the presence of four nearly identical repeats in this region seemed more than coincidental and suggested that dsRNA binding might require a specific zinc finger sequence. We speculated that dsRBP-ZFa binding might be modular like TFIIIA, and that the C-terminal zinc finger motifs might bind one type of nucleic acid, such as dsRNA, while the N-terminal finger motifs might fill another role, such as binding DNA or another protein (*9*). However, unexpectedly, our studies with fragments of dsRBP-ZFa showed that both the N-terminal and C-terminal zinc finger motifs bind dsRNA and RNA−DNA hybrids with high, but slightly different, affinities.

dsRBP-ZFa Fragments Differentially Bind dsRNA and RNA−DNA Hybrids. Our previous studies with the full-length protein, in combination with the K_d measurements (Table 1) and competition studies (Figure 3) reported here, indicate that dsRBP-ZFa and its fragments bind dsRNA better than RNA−DNA hybrids. However, the specificity for dsRNA over RNA−DNA hybrids is most apparent with the fragment zf-(4)567 (see RD/RR ratio, Table 1 and Figure 3). Possibly, the C-terminal zinc fingers represented in zf-(4)567 are less suited for binding to RNA−DNA hybrids than the N-terminal fingers. For example, the interactions that occur with zf-123 may be less sensitive than zf-(4)567 to the lack of a 2′-OH on one strand. RNA−DNA hybrids have a conformation in which the DNA strand forms a structure intermediate between A- and B-form while the RNA strand maintains an A-form conformation (*15, 16*). Thus, alternatively, the N-terminal finger regions of zf-123 may be more suited to induce the DNA strand of an RNA−DNA hybrid to assume an A-form helix.

One important consequence of the binding studies reported here is the expansion of the number of zinc finger sequences (or linker sequences) that are known to bind dsRNA. Since zf-123 and zf-(4)567 do not share zinc finger regions of the same sequence, and they both bind dsRNA, dsRBP-ZFa must contain at least two zinc finger regions capable of high-affinity binding to the A-form helix. Since it is likely that more than one finger in each fragment interacts with dsRNA, dsRBP-ZFa probably provides examples of several different zinc finger sequences with the capability to bind dsRNA. Of course, it is important to keep in mind that, because the linkers joining the zinc finger motifs in dsRBP-ZFa are much longer than in other ZFPs, it is not possible to attribute the novel binding properties solely to the zinc finger motifs; possibly the linkers do the actual binding. It might be informative to substitute the conserved linker found in many ZFPs for those in dsRBP-ZFa to determine if they are required for binding.

dsRBP-ZFa was not the first protein found to bind dsRNA; rather, several investigators have reported dsRNA binding by proteins containing another sequence motif named the dsRNA-binding motif (dsRBM) (*17*−*19*). These proteins have binding characteristics similar to dsRBP-ZFa: they bind dsRNA with high affinity, and while some also bind RNA−DNA hybrids (*17*), they do not bind ssRNA or DNA.

A Model for dsRBP-ZFa Binding. In our previous report, we determined an apparent binding site size for full-length dsRBP-ZFa of 18−25 base pairs (bp) by considering the number of shifts observed with 36-mer (2 shifts) and 100-mer (4 shifts) dsRNA substrates. At that time we did not know if the first three fingers were able to bind dsRNA, so we could not predict the site size of each individual finger. Although we still do not know if the zinc finger motifs themselves (rather than the linkers) are responsible for the nucleic acid binding properties of dsRBP-ZFa, the binding site size calculated per finger is consistent with that determined for fingers within other ZFPs. For example, in DNA-binding ZFPs, each zinc finger motif typically occupies 3 bp, and in the case of ZFPs such as Sp1 or Zif268, which contain three zinc finger motifs, a contiguous 9 bp sequence is recognized (*20, 21*). If the zinc finger motifs in dsRBP-ZFa bind similarly, the seven zinc finger motifs would be expected to occupy 21 bp, a number consistent with the binding site size for the full-length protein. However, for reasons explained below, dsRBP-ZFa likely binds dsRNA differently than Zif268 and similar proteins bind DNA.

The cocrystal structure of Zif268 with its DNA site revealed that the protein has a C-shape which wraps around the DNA, following the path of the major groove. This mode of binding would be topologically problematic for multifinger proteins such as dsRBP-ZFa or TFIIIA since it would result in the protein twisting around the DNA for up to three helical turns. When the cocrystal structure of the first six fingers from TFIIIA in complex with 31 bp of the 5S ICR was determined, it revealed a unique mode of DNA binding (*12*); while fingers 1−3 of TFIIIA wrap around the major groove of DNA and bind in a manner similar to Zif268 (*12, 22*), fingers 4−6 do not (*12*). Instead, fingers 4−6 bind in an extended conformation in which only finger 5 makes contacts with bases in the major groove, while fingers 4 and 6 straddle the flanking minor grooves.

In our mobility shift assays with dsRBP-ZFa fragments, up to three shifts are visible for zf-123 binding to 36 bp dsRNA, while only one or two shifts are observed for zf-(4)567 binding to 36 bp dsRNA. This suggests zf-123 has a maximum binding site size of 12 bp, while zf-(4)567 has a binding site size between 18 and 36 bp. If zf-123 were an Sp1-like ZFP, its three fingers would be expected to occupy a minimum of 9 bp, and to occupy at least 27 bp when three zf-123 proteins simultaneously bind one dsRNA molecule. Thus, data for zf-123 are consistent with a model in which it binds dsRNA like the first three fingers of TFIIIA or Zif268 binds DNA. On the other hand, the dsRNA binding of zf-(4)567 is more consistent with this protein binding in an extended conformation similar to that found with fingers 4−6 from TFIIIA.

Finally, it is important to note that since zf-123 and zf-(4)567 bind dsRNA with an affinity very close to that of the full-length protein, the binding energy of the full-length protein cannot merely be a sum of the interactions that occur

with the fragments. Given our data, it is possible that the full-length protein gains all of its binding energy using only a subset of its zinc fingers. However, an alternative explanation is that there are unfavorable interactions between the various zinc fingers when they exist together in the intact protein, or between the basic zinc finger region and the C-terminal acidic domain. Recent data suggest such "thermodynamic interference" occurs between the zinc fingers of TFIIIA (*10*). Interestingly, similar to what we observed for dsRBP-ZFa, subsets of the zinc fingers from different regions of TFIIIA show affinities very close to that of the intact, full-length protein (*10*).

Biological Relevance of dsRBP-ZFa. In our initial description of dsRBP-ZFa, we noted that there were four sequences in various databases with similarities to dsRBP-ZFa (*4*). More recent database searches have revealed several additional mammalian sequences with significant similarities to short regions of dsRBP-ZFa (see Materials and Methods). Together these sequence similarities include each of the zinc finger motifs in dsRBP-ZFa, suggesting there might be mammalian ZFPs that have binding properties similar to dsRBP-ZFa. Unfortunately, at this time, none of these database sequences have been completely sequenced or characterized.

In our original characterization of dsRBP-ZFa we also noted that its zinc finger motifs are widely spaced relative to those found in most ZFPs. The majority of C_2H_2 ZFPs have zinc finger motifs that are separated by an evolutionarily conserved six to eight amino acid linker (*1*). Previously, several proteins were grouped together on the basis of the long linkers between their zinc fingers, and we added dsRBP-ZFa to this group (*4*). In addition to long linkers, many of the zinc fingers in these proteins have an inter-histidine distance of five residues as opposed to the more common three to four residues.

Recently, the sequence of a novel p53-induced ZFP that can induce apoptosis was reported (*23, 24*; accession numbers Y13148 and AF012923). This protein, PAG608, has three zinc finger motifs with interesting similarities to dsRBP-ZFa. Although the linker sequences are not very similar, the zinc finger motifs in PAG608 are widely spaced and also have five residues separating the histidine residues involved in zinc coordination. Additionally, the first two zinc finger motifs in PAG608 have strong identities to the first two zinc finger motifs in dsRBP-ZFa. Amino acids 72−105 and 149−179 in PAG608 show 47% and 58% identity (58% and 74% similarity) with dsRBP-ZFa, respectively. Furthermore, PAG608, like dsRBP-ZFa, is a nuclear protein. We have conducted a preliminary binding study using bacterially expressed PAG608 protein (data not shown). In this study, PAG608 bound 36-mer dsRNA with an affinity comparable to that obtained using zf-123. Further experiments to measure the K_d for the interaction of PAG608 with 36-mer dsRNA and to determine if PAG608 can also bind RNA−DNA hybrids are planned.

The implications of dsRNA binding by PAG608 are 2-fold. First and foremost, it shows that dsRBP-ZFa is not the only ZFP capable of binding dsRNA and suggests that dsRNA binding may be a property of any protein that contains zinc fingers similar to those in dsRBP-ZFa. Second, since the sequence similarity between dsRBP-ZFa and PAG608 is limited to the zinc finger motifs, this result suggests that the fingers, not the linkers, are responsible for nucleic acid binding. Future experiments that investigate the nucleic acid binding preferences of other ZFPs similar to dsRBP-ZFa will continue to expand our understanding of the versatile zinc finger motif.

ACKNOWLEDGMENT

We are grateful to Scott Beeser for helpful discussions and critiques of the manuscript and to David Israeli for providing the PAG608 cDNA clones. P.J.F. is grateful to Star Division Corp. for providing an excellent office suite for Linux.

REFERENCES

1. Hoovers, J. M., Mannens, M., John, R., Bliek, J., van Heyningen, V., Porteous, D. J., Leschot, N. J., Westerveld, A., and Little, P. F. (1992) *Genomics 12*, 254−263.
2. Andreazzoli, M., De-Lucchini, S., Costa, M., and Barsacchi, G. (1993) *Nucleic Acids Res. 21*, 4218−4225.
3. Koster, M., Kuhn, U., Bouwmeester, T., Nietfeld, W., el, B. T., Knochel, W., and Pieler, T. (1991) *EMBO J. 10*, 3087−3093.
4. Finerty, P. J., Jr., and Bass, B. L. (1997) *J. Mol. Biol. 271*, 195−208.
5. Shi, Y., and Berg, J. M. (1995) *Science 268*, 282−284.
6. Nightingale, K. P., and Wolffe, A. P. (1995) *J. Biol. Chem. 270*, 22665−22668.
7. Berg, J. M. (1990) *Annu. Rev. Biophys. Biophys. Chem. 19*, 405−421.
8. Delarue, M., and Moras, D. (1989) in *Nucleic Acids and Molecular Biology* (Eckstein, F., and Lilley, D. M. J., Eds.) pp 182−196, Springer-Verlag, Heidelberg.
9. Theunissen, O., Rudt, F., Guddat, U., Mentzel, H., and Pieler, T. (1992) *Cell 71*, 679−690.
10. Kehres, D. G., Subramanyan, G. S., Hung, V. S., Rogers, G. W., Jr., and Setzer, D. R. (1997) *J. Biol. Chem. 272*, 20152−20161.
11. Neely, L., Trauger, J. W., Baird, E. E., Dervan, P. B., and Gottesfeld, J. M. (1997) *J. Mol. Biol. 274*, 439−445.
12. Nolte, R. T., Conlin, R. M., Harrison, S. C., and Brown, R. S. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 2938−2943.
13. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1990) *Current Protocols in Molecular Biology*, Greene Publishing Associates and Wiley-Intersciences, New York.
14. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol. 215*, 403−410.
15. Salazar, M., Fedoroff, O. Y., Miller, J. M., Ribeiro, N. S., and Reid, B. R. (1993) *Biochemistry 32*, 4207−4215.
16. Gonzalez, C., Stec, W., Kobylanska, A., Hogrefe, R. I., Reynolds, M., and James, T. L. (1994) *Biochemistry 33*, 11062−11072.
17. Bass, B. L., Hurst, S. R., and Singer, J. D. (1994) *Curr. Biol. 4*, 301−314.
18. Gibson, T. J., and Thompson, J. D. (1994) *Nucleic Acids Res. 22*, 2552−2556.
19. St. Johnston, D., Brown, N. H., Gall, J. G., and Jantsch, M. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 10979−10983.
20. Berg, J. M., and Shi, Y. (1996) *Science 271*, 1081−1085.
21. Pieler, T., and Bellefroid, E. (1994) *Mol. Biol. Rep. 20*, 1−8.
22. Wuttke, D. S., Foster, M. P., Case, D. A., Gottesfeld, J. M., and Wright, P. E. (1997) *J. Mol. Biol. 273*, 183−206.
23. Israeli, D., Tessler, E., Haupt, Y., Elkeles, A., Wilder, S., Amson, R., Telerman, A., and Oren, M. (1997) *EMBO J. 16*, 4384−4392.
24. Varmeh-Ziaie, S., Okan, I., Wang, Y., Magnusson, K. P., Warthoe, P., Strauss, M., and Wiman, K. G. (1997) *Oncogene 15*, 2699−2704.

BI9825293