BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2010,** *49,* 679–686 679
DOI: 10.1021/bi9018626

Characterization of the DNA Binding Activity of the ZFY Zinc Finger Domain†

Jennifer Grants, Erin Flanagan, Andrea Yee, and Paul J. Romaniuk*

Department of Biochemistry and Microbiology, University of Victoria, P.O. Box 3055, Victoria, British Columbia V8W 3P6, Canada

Received October 29, 2009; Revised Manuscript Received December 11, 2009

ABSTRACT: The ZFY protein is a member of one of the most interesting classes of polydactyl zinc finger proteins. It has a domain of 13 tandem zinc fingers that is organized with an internal repeat of odd−even finger pairs. It has been proposed that each finger pair interacts with six base pairs within a turn of the double helix, the downstream linker crossing the minor groove to place the next finger pair on the following turn of the DNA. Yet putative binding sites for the full-length ZFY protein appear to consist of a six-base AGGCCY consensus sequence that is present in one or two copies. In this study the equilibrium binding of two ZFY-derived zinc finger peptides to 4R DNA with tandem copies of the consensus sequence was investigated. The ZFY5 peptide contains fingers 5−13, including four odd−even finger pairs, and the ZFY11 peptide contains fingers 11−13 and has one odd−even finger pair. Both peptides bound to 4R DNA with equal affinities, forming a bimolecular complex that is mediated by the downstream AGGCCY motif. The additional odd−even finger pairs in ZFY5 made no measurable difference in the mechanism of DNA binding compared to ZFY11. The effects on the DNA−protein interaction of mutations in the 4R DNA and in the key α-helical residues of fingers 11−13 indicate that the binding of ZFY to DNA is mediated by the interaction of the GGCC core base pairs with fingers 12 and 13. These results demonstrate that the even−odd repeats in the ZFY zinc finger domain do not make significant contributions to DNA binding.

One of the most common structural motifs in eukaryotic proteins is the C_2H_2 zinc finger (*1*−*3*). Each zinc finger motif is roughly 30 amino acids in length and consists of a $\beta\beta\alpha$-fold that is stabilized by the tetrahedral coordination of the conserved cysteines and histidines to a zinc ion and by a hydrophobic core consisting of conserved amino acids located at specific positions within the motif. It is generally assumed that a protein with a domain consisting of two or more linked zinc fingers binds to DNA and acts as a transcription factor. Three-dimensional structures of zinc finger domains consisting of three to five zinc fingers bound to DNA have provided a detailed picture of how this protein motif forms sequence-specific interactions with the base pairs (*4*−*7*).

Many mammalian zinc finger proteins have domains consisting of much longer tandem arrays of zinc fingers. Since zinc finger proteins can be multifunctional, binding to DNA and/or RNA and/or proteins (*3*, *8*), it is not immediately obvious for any particular polydactyl zinc finger protein how the individual fingers are involved in binding to more than one type of cellular ligand. The one clear example is TFIIIA, which consists of nine tandem zinc fingers. TFIIIA binds to the internal promoter of the 5S rRNA gene, which consists of box C, intermediate, and box A elements contained within a 50 base pair region (*9*, *10*). Fingers 1−3, 5, and 7−9 bind to these three elements respectively; fingers 4 and 6 cross over the minor groove of the DNA and provide the necessary displacement along the double helix (*11*−*16*). TFIIIA also binds to the 5S rRNA transcript, an interaction that is mediated primarily by fingers 4 and 6 (*17*−*22*). While providing valuable insight into how different functional roles can be accommodated within a polydactyl zinc finger domain, it is clear that the domain of TFIIIA is highly specialized and is unlikely to be a general model for the interaction of other polydactyl zinc finger proteins with DNA or RNA or protein partners.

ZFY is a member of an intriguing class of highly conserved polydactyl zinc finger proteins (*23*). The ZFY domain consists of 13 tandem zinc fingers, organized as a repeat of a two-finger pair (*24*). The sequences of the odd-numbered fingers conform to the general zinc finger consensus, whereas the sequences of the even-numbered fingers have some unique features which result in a slightly different folded structure. It was proposed that ZFY would interact with DNA in such a way that each odd−even repeat would contact approximately six base pairs of DNA and the linker after the even finger would cross the minor groove to position the following odd−even finger pair on the next turn of the DNA (*25*, *26*). This model predicts that the ZFY zinc finger domain would interact with up to six turns of the DNA double helix. However, high-affinity binding sites identified by *in vitro* selection methods (*27*), and DNase I footprinting of the HLA-A11 promoter (*28*), have identified ZFY binding sites that consist of a tandem repeat of an AGGCCY box. In order to understand the mechanism by which ZFY interacts with DNA, we have characterized the equilibrium binding of zinc finger peptides containing four odd−even repeats (fingers 5−13) and one odd−even repeat (fingers 11−13) to a consensus DNA site. Site-directed mutagenesis of both the DNA consensus site and the protein zinc fingers was used to measure the contribution of individual base pairs and amino acids to the overall free energy of forming the ZFY−DNA complex. Our results indicate that DNA binding by ZFY is mediated exclusively by the interaction of zinc fingers 12 and 13 with one AGGCCY box.

†This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.
*Corresponding author. Telephone: (250) 721-7088. Fax: (250) 721-6227. E-mail: pjr@uvic.ca.

Published on Web 12/22/2009

MATERIALS AND METHODS

Bacterial Strains and Plasmid Vectors. The plasmid vector pET30a was used to express recombinant wild-type and mutant zinc finger proteins in *Escherichia coli* strain BL21 DE3 using methods that have been described elsewhere (*29*).

Construction of Zinc Finger Protein Expression Vectors. The zinc finger expressing cDNAs used in this study were constructed by the polymerase chain reaction (PCR) using upstream primers containing a recognition sequence for the restriction enzyme *Nco*I and downstream primers containing a recognition sequence for the restriction enzyme *Eco*RI. PCR products were digested with *Nco*I and *Eco*RI and ligated into plasmid pET30a that had been digested with the same restriction enzymes. Ligation reactions were used to transform *E. coli* strain DH5α. Putative constructs were identified by colony PCR using primers specific to pET30a sequences flanking the multiple cloning site. Plasmids were isolated from appropriate colonies and characterized by DNA sequencing before protein expression and purification. Truncations of the mouse ZFY protein containing the cDNA regions encoding zinc fingers 1−13 (designated ZFY1, encoded by nucleotides 1517−2671 of GenBank M24401), zinc fingers 5−13 (designated ZFY5, encoded by nucleotides 1886−2671 of GenBank M24401), and zinc fingers 11−13 (designated ZFY11, encoded by nucleotides 2399−2671 of GenBank M24401) were constructed by PCR using a plasmid containing the ZFY cDNA (ATCC 63041) as a template.

Construction of Mutant ZFY11 Expression Vectors. Alanine substitution mutants of the ZFY11 zinc finger peptide were constructed by overlap extension PCR using primers encoding the desired codon change (*30*). Final PCR products were digested with *Nco*I and *Eco*RI and then ligated into pET30a as described above.

Expression and Purification of Recombinant Zinc Finger Proteins. Preparation of N-terminal His-tagged zinc finger proteins was carried out as described previously (*29*). Protein purity was confirmed by polyacrylamide gel electrophoresis, and the concentration of each protein preparation was determined by the method of Bradford (*31*).

Radiolabeling of DNA Ligands. All DNA ligands were synthesized as complementary oligonucleotides designed to have 5′ overhanging *Bam*HI and *Eco*RI sites upon annealing. The resulting double-stranded DNAs were 3′ end labeled as described previously (*29*, *32*).

Equilibrium Binding of DNAs to Zinc Finger Proteins. The apparent association constants for the binding of radiolabeled DNAs to zinc finger proteins were determined using a double filter binding assay (*33*). The binding buffer consisted of 20 mM Tris-HCl, pH 7.5 (20 °C), 5 mM $MgCl_2$, 100 mM KCl, 10 μM $ZnCl_2$, 0.5 mM tris(2-carboxyethyl)phosphine hydrochloride, 100 μg/mL BSA, and 1 μg/mL poly(dI·dC). When the effects of altering binding conditions such as pH, divalent salt concentration, or monovalent salt concentration were being investigated, the binding buffer was altered accordingly.

The affinity of each zinc finger protein for DNA was determined using three or more independent assays. Apparent dissociation constants (K_d) for the binding of the mutant and wild-type proteins to DNA were calculated by fitting the data to a simple bimolecular equilibrium model using the general curve fitting function of Kaleidagraph software (Synergy Software, Reading, PA) and the equation:

$$\frac{[\text{DNA} - \text{protein}]}{[\text{DNA}]_{\text{total}}} = \frac{[\text{protein}]_{\text{total}}}{[\text{protein}]_{\text{total}} + K_d} \quad (1)$$

where $[\text{DNA}]_{\text{total}} \ll K_d$ and $[\text{DNA}-\text{protein}]/[\text{DNA}]_{\text{total}}$ is reported as the fraction of DNA bound. Association constant (K_a) values were derived as the reciprocal of the measured K_d values and are reported as the mean of three independent determinations with the associated standard deviations. Relative affinities were arrived at by dividing the apparent K_a for the mutant protein by the apparent K_a for the wild-type protein determined in parallel or by dividing the apparent K_a for the mutant DNA by the apparent K_a for the wild-type DNA determined in parallel. The errors for relative affinities are given by the expression $\sigma = \{(\sigma 1/\text{M1})^2 + (\sigma 2/\text{M2})^2\}^{1/2}(\text{M2}/\text{M1})$, where M1 and M2 are the respective association constants for wild-type and mutant protein or DNA and the σ values are the corresponding standard deviations for these determinations.

For some of the mutant proteins, the apparent affinity for DNA was significantly decreased. The standard filter binding protocol relies upon the assumption that 100% of the protein is in an active form, and therefore for these mutant proteins it was necessary to confirm that this assumption was valid. In these cases the filter binding assay was modified so that the wild-type or mutant protein was held at a constant concentration and incubated with a series of DNA concentrations ranging from 2.2×10^{-10} M to 8.2×10^{-8} M. The determination of apparent association constants using this assay does not rely upon the concentration of active protein. All other conditions of this assay are identical to the standard filter binding assay. After a 90 min incubation at 20 °C, each sample was filtered through the double nitrocellulose/nylon filter stack to provide a determination of protein-bound DNA (nitrocellulose filter) and free DNA (nylon filter). Apparent association constants were calculated by fitting the data to a simple bimolecular equilibrium model using the appropriate equation and the general curve fitting function of Kaleidagraph software (Synergy Software, Reading, PA). For all of the mutant proteins tested the results from the two different assays were identical within experimental error.

RESULTS

The ZFY protein has a DNA binding domain consisting of 13 tandem C_2H_2 zinc fingers, with several unique properties (Figure 1). The linkers that joint the first four fingers are 10−11 amino acids long, rather than the typical linker of 7 amino acids. For zinc fingers 3−12, the odd-numbered zinc fingers conform to the consensus zinc finger sequence motif, having three amino acids between the zinc coordinating histidines. The even-numbered fingers diverge from this consensus, notably by having a glutamine residue immediately preceding the first cysteine, by having a noncanonical positioning of the central aromatic amino acid, and also by having four residues between the two zinc coordinating histidines (Figure 1). The unique architecture of the ZFY zinc finger domain is believed to have implications for its interactions with DNA, the proposed model predicting that the overall binding site would span up to six turns of the double helix, with each odd−even finger pair interacting with six base pairs per turn (*25*, *26*). However, binding site selection experiments carried out with the entire ZFY zinc finger domain resulted in the identification of a relatively short binding site of AGGCCY, typically presented in a tandem repeat (*27*).

The Zinc Finger Domain of Mouse ZFY2 Protein

```
                              -1  2 3    6
                               |  | |    |
ZF01:  V Y P C M F C G K K F K T K R F L K R H I K N - H P E Y L A N K
ZF02:  K Y H C T E C D Y S T N K K I S L H N H M E S - H K L T I K T E K
ZF03:  T T E C D D C R K N L S H A G T L C T H K T M - H T E K G V N K
ZF04:  T C K C K F C D Y E T A E Q T L L N H H L L V V H R K K F
ZF05:  P H I C G E C G K G F R H P S A L K K H I R V - H T G E K
ZF06:  P Y E C Q Y C E Y K S A D S S N L K T H I K S K H S K E I
ZF07:  P L K C D I C L L T F S D T K E A Q Q H A V L - H Q E S R
ZF08:  T H Q C S H C N H K S S N S S D L K R H I I S V H T K A Y
ZF09:  P H K C D M C S K G F H R P S E L K K H V A T - H K S K K
ZF10:  M H Q C R H C D F N S P D P F L L S H H I L S A H T K N V
ZF11:  P F K C K R C K K E F Q Q Q C E L Q T H M K T - H S S R K
ZF12:  V Y Q C E Y C E Y S T K D A S G F K R H V I S I H T K D Y
ZF13:  P H R C D F C K K G F R R P S E K N Q H I M R H H K E V G L A
```

FIGURE 1: Amino acid sequence of the zinc finger domain of the mouse ZFY2 protein. Amino acid features that differ between the canonical odd and noncanonical even fingers are highlighted in bold letters. The key α-helical positions involved in DNA binding are indicated by the numbering scheme above the sequence.

To investigate the role of individual zinc fingers of ZFY in DNA binding, we expressed three peptides: ZFY1 containing the complete zinc finger domain, ZFY5 consisting of fingers 5 through 13 with four odd−even pairs that include the even fingers that have all three unique characteristics, and ZFY11 which consists of only fingers 11 through 13 and has one odd−even pair of fingers. The ZFY1 construct produced low yields of protein, which was not stable during the purification procedure and quickly degraded upon storage and therefore was not suitable for quantitative DNA binding studies. The ZFY5 and ZFY11 constructs produced suitable quantities of stable proteins, and the affinities of the ZFY5 and ZFY11 peptides for two putative ZFY DNA binding sites were measured using an equilibrium binding assay. The previously identified 10F binding site has an imperfect repeat of the consensus AGGCCY sequence, while the 4R binding site has a repeat of two consensus sequences (Figure 2A). The 4R site binds to both ZFY5 and ZFY11 with high affinity, the apparent association constants being $(1.51 \pm 0.09) \times 10^9$ M^{-1} for ZFY5 and $(1.08 \pm 0.08) \times 10^9$ M^{-1} for ZFY11 (Figure 2B). At the assay temperature of 20 °C, the free energy of binding to the 4R DNA site is -12.30 ± 0.72 kcal mol^{-1} for ZFY5 and -12.11 ± 0.87 kcal mol^{-1} for ZFY11. In comparison to 4R, the 10F site does not bind ZFY11 and binds with 50-fold lower affinity to the ZFY5 protein, the apparent association constant for the interaction being $(3.07 \pm 0.14) \times 10^7$ M^{-1} (Figure 2B).

The 4R DNA has a tandem repeat of the consensus AGGCCY boxes, each of which could potentially bind a ZFY protein. To determine how the ZFY peptides interact with 4R DNA, equilibrium binding assays were carried out by titrating 2.5 nM protein with concentrations of 4R DNA ranging from 0.2 to 20 nM. The potential for cooperativity in the binding ZFY5 and ZFY11 to 4R DNA was assessed using a Hill plot (*34*). The Hill constant for ZFY5 binding to 4R DNA is 1.06 ± 0.07, and for the binding of ZFY11 the Hill constant is 1.02 ± 0.02, values that indicate the binding of each protein to the DNA is not cooperative (Figure 3A). The stoichiometries of the complexes formed between 4R DNA and the ZFY peptides were determined by Scatchard analysis (*35*) and found to be 0.99 ± 0.03 for ZFY5 and 0.96 ± 0.02 for ZFY11 (Figure 3B). The data from these analyses indicate that ZFY5 and ZFY11 form simple bimolecular complexes with the 4R DNA.

The effects on the association constant of changes in monovalent and divalent salt concentrations and pH were compared for ZFY5 and ZFY11 to determine if the presence of additional odd−even pairs of fingers influences the stability of the protein−DNA complex. The contributions of ionic

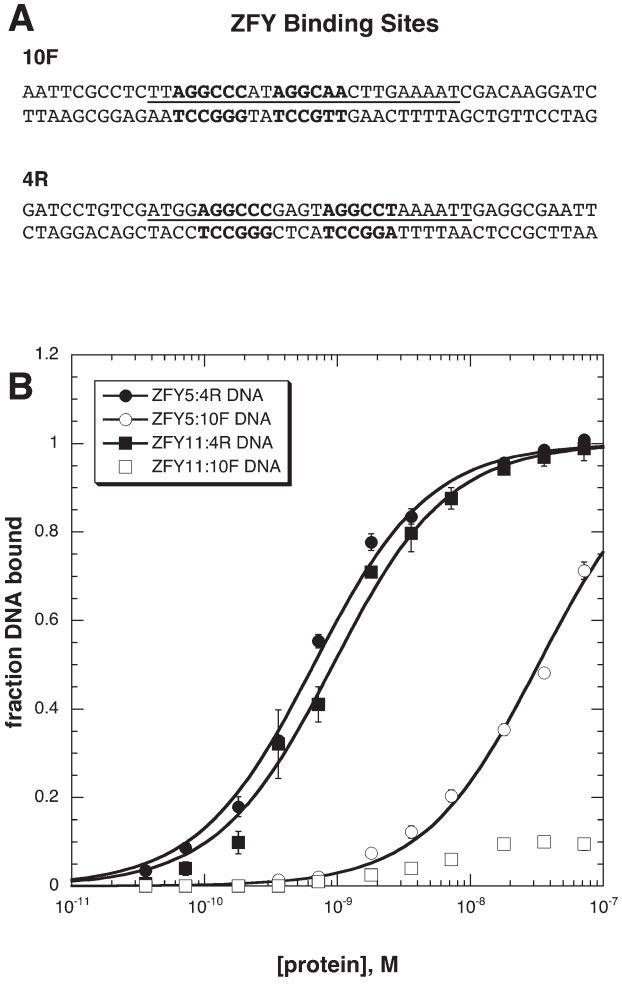

FIGURE 2: Equilibrium binding of the ZFY5 and ZFY11 peptides to DNA. (A) The sequence of the two probes used for equilibrium binding assays, originally isolated by *in vitro* selection from a library of random sequences (*27*). The consensus AGGCCY boxes are shown in bold within the complete region selected as indicated by the line. (B) Equilibrium binding curves for the interactions of ZFY5 and ZFY11 with the 10F and 4R DNA probes. Each data point is the mean value from three independent experiments, with the standard deviation for each point represented by the error bars. Best-fit curves to a simple bimolecular equilibrium are shown.

protein−DNA bonds to the binding energies of the ZFY5−DNA and ZFY11−DNA interactions were determined using an analysis of the salt dependence of K_a based upon an ion displacement model (*36*, *37*). Analysis of the results in Figure 4A indicate that 3.9 ± 1.0 ionic contacts are formed between ZFY5 and 4R, while 4.1 ± 0.7 ionic contacts are formed between ZFY11 and 4R. The binding of both proteins to 4R shows very similar sensitivity to the concentration of $MgCl_2$ in the assay buffer, with ZFY5 retaining slightly more affinity in the 3−5 mM range compared to ZFY11 (Figure 4B). Both interactions have similar pH profiles, with the ZFY11−DNA interaction being slightly less sensitive to pH values below 7.0 than the ZFY5−DNA interaction but being somewhat more sensitive to pH values above 7.5 (Figure 5). This comparison suggests that DNA binding by ZFY5 may require a group with a pK_a value below 7.0 that is not involved in DNA binding by ZFY11. A residue with a pK_a value above 7.5 is involved in the interaction of both proteins with DNA, making a larger contribution to the overall free energy of binding in the ZFY11−DNA interaction (Figure 5).

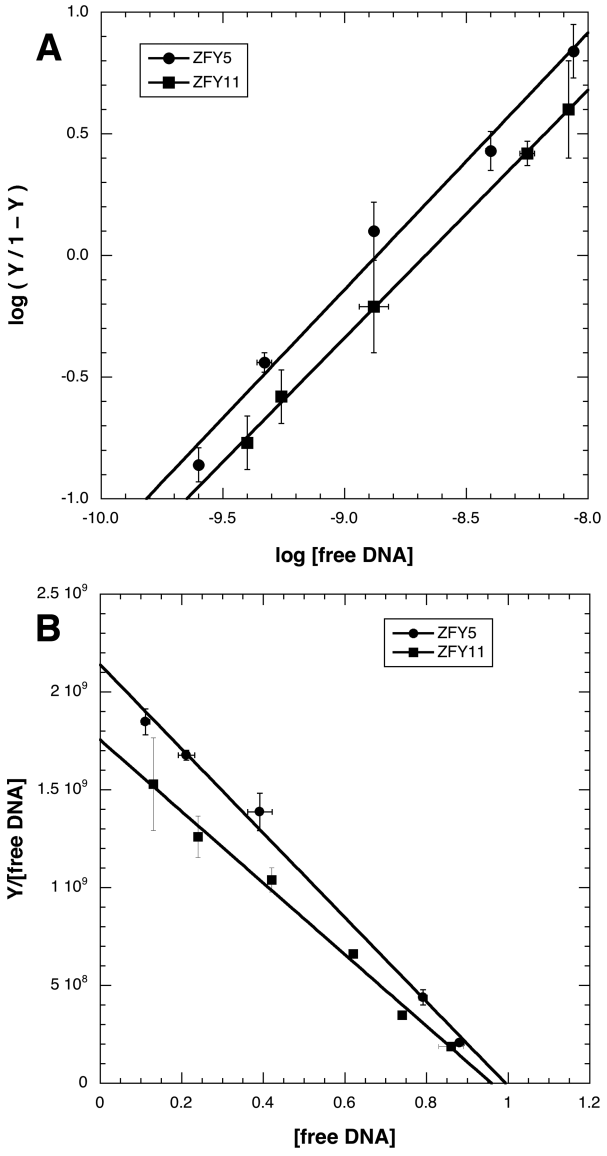

FIGURE 3: Analysis of the cooperativity and stoichiometry of the binding of ZFY5 and ZFY11 to 4R DNA. (A) Hill plot of data obtained with a nitrocellulose filter binding assay where 2.5 nM protein was titrated with 0.2−20 nM DNA. Points corresponding to fraction bound (Y) of 0.1−0.9 are plotted (*34*). Each data point is the mean value from two independent experiments, with the standard deviation for each point represented by the error bars. The lines represent the best fit to the Hill equation determined by linear regression. (B) Scatchard plot of the same data. The lines represent the best fit to the Scatchard equation determined by linear regression.

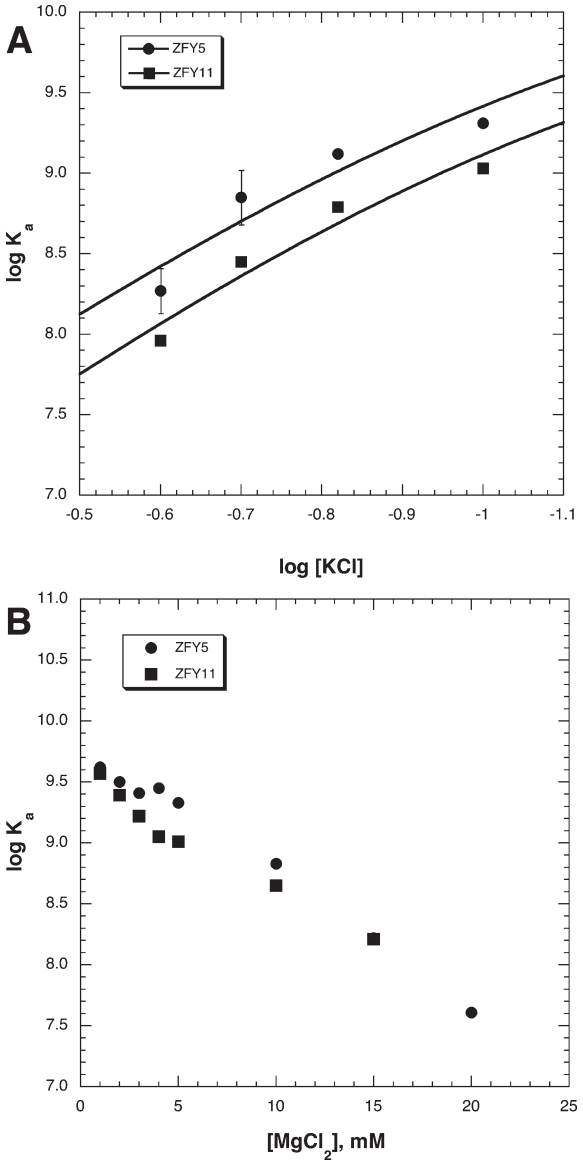

FIGURE 4: Effects of monovalent and divalent salt concentrations on the affinities of ZFY5 and ZFY11 for 4R DNA. (A) Effect of monovalent salt concentration. The experimental data were fit to eq 5 of ref *41* by a regression method that varied the number of ion pairs and the apparent K_a at 1 M salt. The continuous lines represent the best least-squares fit. (B) Effect of magnesium ion concentration. log K_a values are reported for a single experiment, the error bars indicating the level of imprecision in fitting the data to a simple bimolecular equilibrium model.

The 4R DNA that binds ZFY5 and ZFY11 with high affinity consists of a tandem repeat of two consensus AGGCCY boxes separated by four base pairs, while the 10F DNA that binds ZFY5 poorly and does not bind to ZFY11 consists of a single AGGCCC followed two base pairs later by the partial consensus sequence AGGCAA (Figure 2A). To probe the importance of the two AGGCCY sequences in 4R for high-affinity binding of ZFY, scanning mutants were created and assayed for their affinities for ZFY5 and ZFY11 (Figure 6). Replacement of the upstream AGGCCC sequence in mutant M1 resulted in a 5.8−10-fold decrease in affinity for ZFY binding compared to the wild-type 4R sequence. Replacement of the downstream AGGCCT sequence in mutant M3 decreased the affinities for ZFY5 and ZFY11 by over 2 orders of magnitude (Figure 6). In comparison, replacement of either the four base pairs separating the two AGGCCY boxes in mutant M2 or the six base pairs downstream of the second box in mutant M4 decreased the affinity for ZFY5 and ZFY11 by 2.4-fold or less. These results indicate that the downstream AGGCCT box is the binding site used for formation of a bimolecular complex with ZFY5 and ZFY11, and the upstream AGGCCC box is required to reach the full affinity of binding.

Since both ZFY5 and ZFY11 form bimolecular complexes with the 4R DNA, the results obtained with mutants M1−M4 suggest that the binding site extends over both AGGCCY boxes. This possibility was tested by measuring the effects of changing the spacing between the two boxes by reducing it to zero in mutant M5 or doubling it to eight base pairs in mutant M6 (Figure 6). Removing the spacing between the two boxes resulted in a small reduction in binding affinity for ZFY5 and ZFY11 of

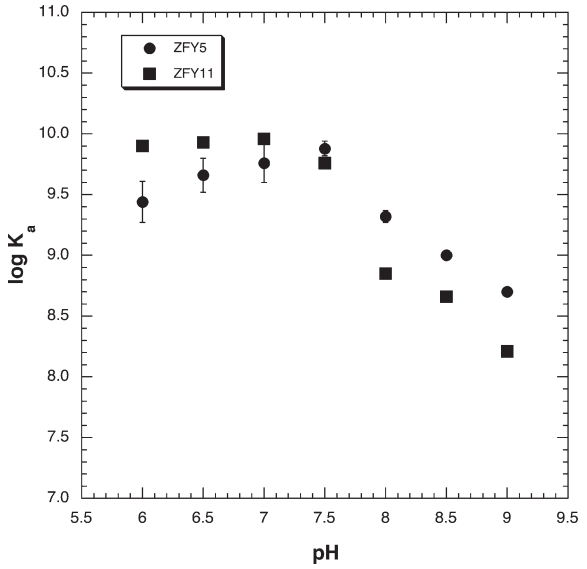

FIGURE 5: Effect of pH on the affinities of ZFY5 and ZFY11 for 4R DNA. log K_a values are reported for a single experiment, the error bars indicating the level of imprecision in fitting the data to a simple bimolecular equilibrium model.

		Relative K_a	
		ZFY5	ZFY11
4R	GATCCTGTCGATGGAGGCCCGAGTAGGCCTAAAATTGAGGCGAATT CTAGGACAGCTACCTCCGGGCTCATCCGGATTTTAACTCCGCTTAA		
M1	GATCCTGTCGATGGCTTAAAGAGTAGGCCTAAAATTGAGGCGAATT CTAGGACAGCTACCGAATTTCTCATCCGGATTTTAACTCCGCTTAA	0.171±0.047	0.101±0.021
M2	GATCCTGTCGATGGAGGCCCTCTGAGGCCTAAAATTGAGGCGAATT CTAGGACAGCTACCTCCGGGAGACTCCGGATTTTAACTCCGCTTAA	0.574±0.029	0.422±0.053
M3	GATCCTGTCGATGGAGGCCCGAGTCTTAAGAAAATTGAGGCGAATT CTAGGACAGCTACCTCCGGGCTCAGAATTCTTTTAACTCCGCTTAA	0.003±0.006	0.007±0.002
M4	GATCCTGTCGATGGAGGCCCGAGTAGGCCTCCCCGGGAGGCGAATT CTAGGACAGCTACCTCCGGGCTCATCCGGAGGGGCCCTCCGCTTAA	0.474±0.056	0.790±0.154
M5	GATCCTGTCGATGGAGGCCCAGGCCTAAAATTGAGGCGAATT CTAGGACAGCTACCTCCGGGTCCGGATTTTAACTCCGCTTAA	0.365±0.127	0.230±0.046
M6	GATCCTGTCGATGGAGGCCCGAGTGAGTAGGCCTAAAATTGAGGCGAATT CTAGGACAGCTACCTCCGGGCTCACTCATCCGGATTTTAACTCCGCTTAA	1.009±0.192	0.366±0.081
M7	GATCCTGTCGATGGAGGCCTGAGTAGGCCTAAAATTGAGGCGAATT CTAGGACAGCTACCTCCGGACTCATCCGGATTTTAACTCCGCTTAA	1.563±0.355	1.369±0.295
M8	GATCCTGTCGATGGAGGCCTGAGTAGGCCCAAAATTGAGGCGAATT CTAGGACAGCTACCTCCGGACTCATCCGGGTTTTAACTCCGCTTAA	1.534±0.338	1.178±0.458
M9	GATCCTGTCGATGGAGGCCCGAGTAGGCCCAAAATTGAGGCGAATT CTAGGACAGCTACCTCCGGGCTCATCCGGGTTTTAACTCCGCTTAA	1.456±0.435	1.066±0.386

FIGURE 6: Effects of scanning, spacing, and point mutations of 4R on the binding affinities of ZFY5 and ZFY11. The wild-type 4R sequence is shown at the top, with the tandem AGGCCY boxes indicated by blue. For the mutants M1−M9, the sequence changes relative to wild-type 4R are shown in red. Relative binding affinities for ZFY5 and ZFY11 were determined from three independent experiments, as outlined in the Materials and Methods section.

2.7−4.3-fold, less than the effect of mutation M1 that completely replaced the upstream AGGCCC box. Increasing the spacing between the two boxes to eight base pairs had no effect on ZFY5 binding and a small 2.7-fold decrease in affinity for ZFY11. The two AGGCCY sequences in 4R differ at the sixth base pair, this being a C in the first box and a T in the second box. Mutants M7−M9 were designed to test whether this base pair difference explained the large difference in contribution of the two boxes to

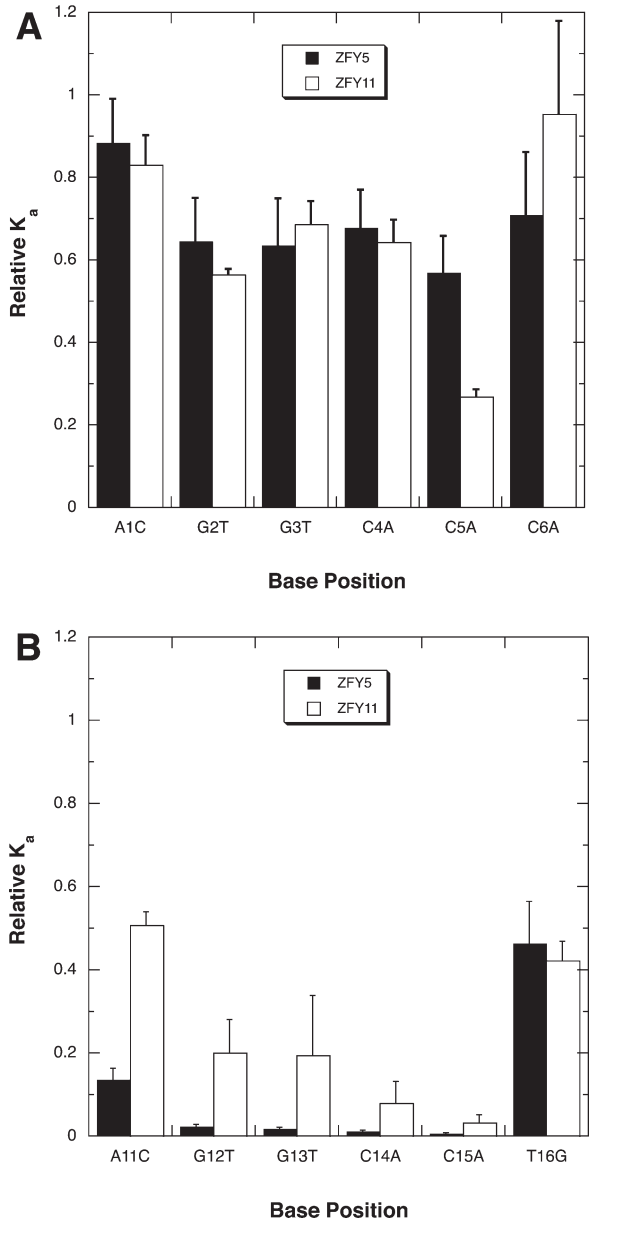

FIGURE 7: Contributions of individual base pairs in the AGGCCY boxes of 4R DNA to the binding affinities of ZFY5 and ZFY11. Relative binding affinities were determined from three independent experiments, as outlined in the Materials and Methods section.

ZFY5 and ZFY11 binding affinities. In mutant M7, the upstream box was mutated to AGGCCT to match the downstream box, in mutant M8 the two boxes were reversed in their positions in 4R, and in mutant M9 the downstream box was mutated to AGGCCC to match the upstream box (Figure 6). All of these mutants showed wild type or slightly elevated affinities for ZFY5 and ZFY11, suggesting that the identity of the sixth base pair does not account for the large differences in the relative contributions of the two AGGCCY boxes to ZFY binding affinity.

The role of individual base pairs in both AGGCCY boxes of 4R in ZFY binding was probed by the creation of point transition mutations (Figure 7). Complete replacement of the upstream box in mutant M1 resulted in a 5.8-fold decrease in affinity for ZFY5 and a 10-fold decrease in affinity for ZFY11 (Figure 6). Transition mutations at each of the six positions within the upstream box individually resulted in less than a 2-fold reduction in binding of ZFY5 and ZFY11, with the exception of a 3.7-fold reduction

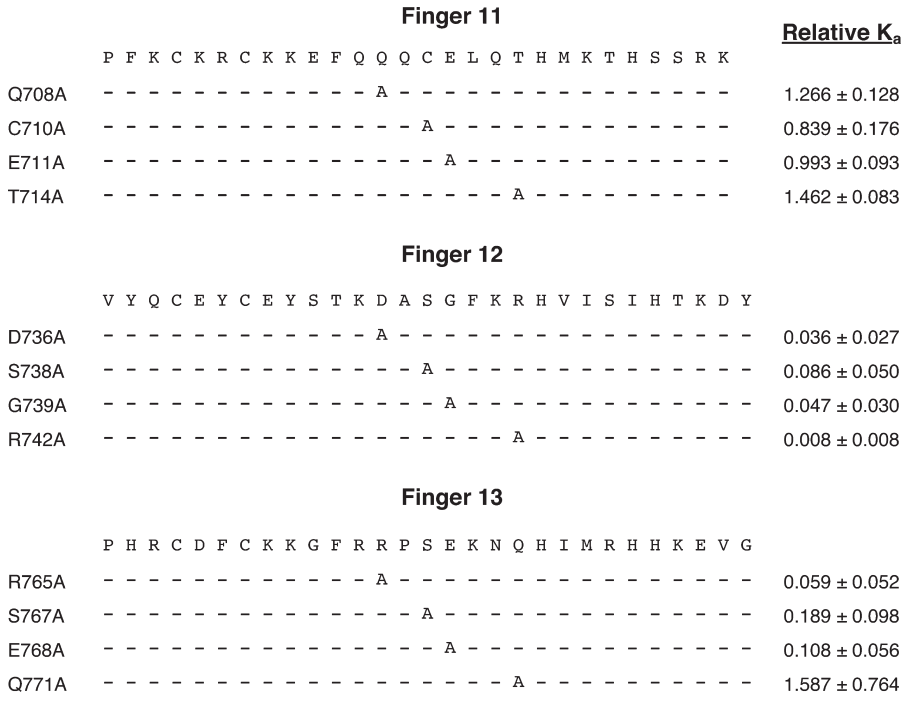

FIGURE 8: Effects of alanine substitution mutations on the affinity of ZFY11 for 4R DNA. The wild-type sequence of each finger is shown, along with the site of the alanine substitutions. Relative binding affinities were determined from three independent experiments, as outlined in the Materials and Methods section.

in the case of the C5A mutation in binding of ZFY11 (Figure 7A). Collectively, the six point mutations would result in a 10-fold reduction in binding affinity for ZFY5 and a 20-fold reduction in affinity for ZFY11, which is reasonably consistent with the effects of the M1 scanning mutation. The difference in the effect of the C5A mutation on ZFY11 vs ZFY5 binding affinity explains the approximately 2-fold difference in the effect of the M1 scanning mutation on the binding affinities of the two proteins. Individual point mutations in the downstream AGGCCT box had a dramatic effect on the binding affinities of both ZFY5 and ZFY11 (Figure 7B). For both proteins, the largest effects are observed for the GGCC core of the hexamer sequence. The sixth base pair would appear to make only a modest contribution to the binding affinity of ZFY5 and ZFY11, while the importance of the first base pair differs for the two proteins. Collectively, the point mutations of the downstream AGGCCT box have a larger effect on the binding of both proteins than the M3 scanning mutation has. This result suggests the possibility that the protein−DNA contacts in this region are interdependent, in which case a single point mutation may weaken the energy of contacts to a neighboring base pair in addition to resulting in the loss of the free energy of the direct contact normally made to the base pair that was mutated. The point mutations at base pairs 11−15 generally have a more deleterious effect on the binding affinity of ZFY5 compared to ZFY11 (Figure 7B), which may indicate that the contacts formed between the downstream AGGCCT box and ZFY5 are more interdependent than those formed between the AGGCCT box and ZFY11.

Since ZFY5 and ZFY11 have virtually identical binding affinities for 4R DNA and the same four to five base pairs of DNA appear to be critical for binding of both proteins, we undertook an alanine mutagenesis study of the key −1, +2, +3, and +6 α-helical positions in the three fingers of the ZFY11 protein. The sequences of the mutated fingers and the effects of the mutations on the affinity of ZFY11 for 4R DNA are shown in Figure 8. Only point mutations in fingers 12 and 13 had a significant effect on the binding of ZFY11 to 4R DNA, consistent with the observation that ZFY11 forms its strongest contacts to the central GGCC core base pairs of the downstream AGGCCT box. All four key α-helical residues of finger 12 are critical for DNA binding by ZFY11, while only residues at positions −1, +2, and +3 of finger 13 are required for high-affinity binding of ZFY11 to DNA (Figure 8). The data are also consistent with a model whereby finger 11 is positioned over the base pairs that flank the AGGCCT box.

DISCUSSION

Our knowledge of the mechanism of DNA binding by zinc finger proteins comes primarily from biochemical and structural studies on complexes formed between a cognate DNA sequence and a domain of three to five zinc fingers (*4*−*7*). Of the 807 human zinc finger genes, 76% contain four or more zinc finger motifs and 45% have at least nine zinc finger motifs. To date, we only have detailed understanding of the interaction with DNA of one polydactyl zinc finger protein, that being TFIIIA with nine zinc fingers (*11*−*16*). ZFY has a domain of 13 tandem zinc finger motifs with the unique feature of a two-finger repeat (*24*). The odd-numbered fingers conform to the consensus zinc finger sequence motif. The even-numbered fingers have a glutamine residue immediately preceding the first cysteine, typically have a tyrosine two residues after the second cysteine rather than the canonical phenylalanine four residues after the second cysteine, and have an additional amino acid between the two zinc coordinating histidines and a hydrophobic residue in the center of the linker. NMR analysis of the structure of zinc finger 6 of ZFY demonstrated that even fingers fold in such a way that they likely interact with DNA via their α-helix exactly as a canonical zinc finger does (*25*, *26*, *38*, *39*). However, while the displaced aromatic tyrosine residue participates in formation of

the hydrophobic core of the zinc finger as does the phenylalanine residue two positions away in the canonical zinc finger, the side chain hydroxyl group of the tyrosine is positioned in such a way as to extend the DNA interaction surface (*38*). In the canonical zinc finger, the two zinc coordinating histidines and the three residues between them extend the recognition α-helix, whereas inclusion of an additional residue between the two histidines in the even finger motif causes a loop to form in order to maintain the optimal geometry for zinc coordination of the two histidines (*26*). Formation of the loop changes the finger-linker exit from the major groove. These observations lead to the proposal that the odd−even pattern of zinc fingers in ZFY results in the binding of an odd−even pair in the major groove, with the linker following the even finger crossing the minor groove to position the next odd−even pair on the next turn of the DNA. It was proposed that the characteristic hydrophobic group in the middle of the linker following the even finger would pack against the deoxyribose sugars of the minor groove to stabilize this orientation (*25*, *26*). This arrangement of the odd−even zinc fingers of ZFY would result in the recognition of a four to five base pair sequence within each successive turn of the DNA, the entire zinc finger domain extending through approximately four to five turns of DNA. Arranging six zinc fingers in pairs of two rather than groups of three enhances the affinity and sequence specificity of artificial proteins (*40*), which suggests that the odd−even finger repeat in ZFY could have similar benefits.

An AGGCCY motif was previously identified as a consensus binding site for ZFY through an *in vitro* selection experiment (*27*). The 4R site is abundant in the resolved library and consists of a tandem repeat of the motif separated by four base pairs, which would place the two motifs on the same face of the DNA double helix. The highly related protein ZFX was shown to regulate expression of the MHC class I HLA-A11 promoter by binding to two AGGCCC sequences three turns apart on the DNA (*28*). We compared the DNA binding properties of two peptides derived from the zinc finger domain of ZFY, one that includes four odd−even finger pairs (ZFY5) and one that includes one odd−even pair (ZFY11). Both proteins bound with high affinity to the 4R DNA but with only poor affinity to 10F DNA, a selected site that only has one consensus AGGCCY box. The influences of monovalent salt concentration, divalent salt concentration, and pH on the binding of ZFY5 and ZFY11 to 4R DNA were very similar, indicating that including additional odd−even pairs of zinc fingers made no significant difference to the mechanism of binding to 4R DNA. Although scanning mutagenesis of the 4R site demonstrated that both AGGCCY boxes are required for high-affinity binding of ZFY5 and ZFY11, it also revealed that the downstream AGGCCT box contributes the majority of the free energy of the interaction regardless of the spacing between the two boxes or the identity of the last base pair of the two boxes. The effects of point mutagenesis of the upstream and downstream AGGCCY boxes on ZFY5 and ZFY11 binding were consistent with the relative contributions of each box to the free energy of the interaction. In the case of the upstream AGGCCC box, in all but one case each point mutation reduced the binding affinity of the DNA for the two proteins by less than a factor of 2. The only exception was the mutation at C5, which reduced the binding of ZFY11 by about a factor of 2 more than it reduced the binding of ZFY5, thus explaining why this upstream box is more critical for ZFY11 binding. In comparison, the point mutants in the downstream AGGCCT box had significant effects on the binding of ZFY5 and ZFY11, consistent with the major contribution this box makes to the overall free energy of binding. In the case of ZFY11, the most significant effects were observed for the core GGCC sequence, while in the case of ZFY5 there would appear to be an additional role for the first base pair in protein binding.

The results obtained with the mutants of the 4R DNA binding site did not seem to be consistent with the docking of tandem odd−even finger repeats to the two AGGCCY boxes. Since ZFY11 has only one such repeat, one would have expected it to bind to 10F and 4R with roughly equal affinity and to bind to 4R with considerably less affinity than ZFY5, which would have the capability of docking two odd−even finger pairs to the two AGGCCY boxes. This model for DNA binding of the ZFY zinc finger domain was tested further by the introduction of alanine substitution mutations at the canonical α-helical positions involved in base recognition in each finger of ZFY11. None of the alanine substitutions in zinc finger 11 had a significant effect on DNA binding, indicating that the odd member of the only odd−even repeat in ZFY11 does not participate in DNA binding. Amino acids at all four canonical positions in the α-helix of finger 12 appear to be involved in DNA binding, as are the first three such residues in zinc finger 13. Therefore, DNA binding by ZFY to AGGCCY boxes is mediated by the last two zinc fingers in the domain and does not involve the odd−even repeat structure of the domain.

While our results have clarified the role of the zinc fingers of ZFY in DNA binding, they also raise some intriguing questions. The role of the upstream AGGCCC box in the binding of the ZFY zinc fingers is unclear. Although it is required for the highest binding affinity, the interaction of ZFY5 and ZFY11 does not rely upon the identity of any of the individual base pairs in this box, unlike the sequence requirements of the downstream AGGCCT box. Changing the spacing between the two boxes by deletion of the intervening sequence, or doubling its length, would put them on opposite faces of the DNA, yet this has little to no effect on the binding of ZFY5 and ZFY11. The isotherms for the binding of both proteins to the 4R DNA show no evidence of cooperativity, and steric hindrance would likely prevent occupation of both boxes simultaneously, particularly in the case of the ZFY5 protein. Furthermore, reversing the orientation of the two boxes in mutant M8 did not result in a change in affinity for the ZFY proteins, nor did making the two boxes identical in sequence in mutants M7 and M9.

Our work has shown that zinc fingers 12 and 13 mediate the interaction of ZFY with a consensus DNA binding site, raising the question of whether the odd−even repeat motifs within fingers 1−10 of ZFY play a role in DNA binding. These fingers all have amino acid residues compatible with DNA binding at the canonical −1, +2, +3, and +6 α-helical positions, yet binding site selection with the full ZFY protein did not identify a consensus binding site longer than 4R. The most straightforward model that explains these results is one where only fingers 12 and 13 make sequence-specific contacts with the DNA. If this model is correct, it will be interesting to identify what features of the odd−even repeat structure prevent interaction of fingers 1−11 with DNA and to determine what roles these fingers play in ZFY function. An alternative model for the ZFY−DNA interaction can be proposed based on our knowledge of the TFIIIA−DNA interaction. In the case of TFIIIA, the majority of the DNA sequence specificity and free energy of binding is provided by the interaction of zinc fingers 1−3 with the box C promoter element (*14*). Zinc finger 4 crosses the minor groove, placing zinc finger 5 in the major groove where it interacts with two base pairs of the

DNA (*16*). Zinc finger 6 is presumed to cross the major groove in order to place zinc fingers 7−9 so that they can interact with the box A element of the 5S RNA gene promoter. Thus the normal mode of interaction of TFIIIA with DNA is more extended and involves fingers 5 and 7−9, even though they do not contribute a great deal to the free energy of binding. It is possible that the normal mode of ZFY binding to DNA might involve some combination of the first ten zinc fingers acting in an analogous fashion to the last six fingers of TFIIIA. Distinguishing between these two models for the interaction of ZFY with DNA will require identification of an authentic genomic DNA binding site.

ACKNOWLEDGMENT

The authors thank Tristen Weiss for many helpful discussions.

REFERENCES

1. Iuchi, S. (2001) Three classes of C_2H_2 zinc finger proteins. *Cell. Mol. Life Sci. 58*, 625–635.
2. Pabo, C., Peisach, E., and Grant, R. (2001) Design and selection of novel Cys2His2 zinc finger proteins. *Annu. Rev. Biochem. 70*, 313–340.
3. Ladomery, M., and Dellaire, G. (2002) Multifunctional zinc finger proteins in development and disease. *Ann. Hum. Genet. 66*, 331–342.
4. Pavletich, N. P., and Pabo, C. O. (1993) Crystal structure of a five-finger GLI-DNA complex: new perspectives on zinc fingers. *Science 261*, 1701–1707.
5. Elrod-Erickson, M., Rould, M. A., Nekludova, L., and Pabo, C. O. (1996) Zif268 protein-DNA complex refined at 1.6 Å: a model system for understanding zinc finger-DNA interactions. *Structure 4*, 1171–1180.
6. Elrod-Erickson, M., Benson, T. E., and Pabo, C. O. (1998) High-resolution structures of variant Zif268-DNA complexes: implications for understanding zinc finger-DNA recognition. *Structure 6*, 451–464.
7. Stoll, R., Lee, B. M., Debler, E. W., Laity, J. H., Wilson, I. A., Dyson, H. J., and Wright, P. E. (2007) Structure of the Wilms tumor suppressor protein zinc finger domain bound to DNA. *J. Mol. Biol. 372*, 1227–1245.
8. Cassiday, L. A., and Maher, L. J., III (2002) Having it both ways: transcription factors that bind DNA and RNA. *Nucleic Acids Res. 30*, 4118–4126.
9. Bogenhagen, D. F. (1985) The intragenic control region of the *Xenopus* 5S RNA gene contains two factor a binding domains that must be aligned properly for efficient transcription initiation. *J. Biol. Chem. 260*, 6466–6471.
10. Pieler, T., Oei, S.-L., Hamm, J., Engelke, U., and Erdmann, V. A. (1985) Functional domains of the *Xenopus laevis* 5S gene promoter. *EMBO J. 4*, 3751–3756.
11. Clemens, K. R., Liao, X. B., Wolf, V., Wright, P. E., and Gottesfeld, J. M. (1992) Definition of the binding sites of individual zinc fingers in the transcription factor IIIA-5S RNA gene complex. *Proc. Natl. Acad. Sci. U.S.A. 89*, 10822–10826.
12. Hayes, J. J., and Clemens, K. R. (1992) Locations of contacts between individual zinc fingers of *Xenopus laevis* transcription factor IIIA and the internal control region of a 5S RNA gene. *Biochemistry 31*, 11600–11605.
13. Del Rio, S., Menezes, S. R., and Setzer, D. R. (1993) The function of individual zinc fingers in sequence specific DNA recognition by transcription factor IIIA. *J. Mol. Biol. 233*, 567–579.
14. Clemens, K. R., Zhang, P. H., Liao, X. B., McBryant, S. J., Wright, P. E., and Gottesfeld, J. M. (1994) Relative contributions of the zinc fingers of transcription factor IIIA to the energetics of DNA binding. *J. Mol. Biol. 244*, 23–35.
15. Neely, L., Trauger, J. W., Baird, E. E., Dervan, P. B., and Gottesfeld, J. M. (1997) Importance of minor groove binding zinc fingers within the transcription factor IIIA DNA complex. *J. Mol. Biol. 274*, 439–445.
16. Nolte, R. T., Conlin, R. M., Harrison, S. C., and Brown, R. S. (1998) Differing roles for zinc fingers in DNA recognition: structure of a six finger transcription factor IIIA complex. *Proc. Natl. Acad. Sci. U.S.A. 95*, 2938–2943.
17. Clemens, K. R., Wolf, V., Mcbryant, S. J., Zhang, P. H., Liao, X. B., Wright, P. E., and Gottesfeld, J. M. (1993) Molecular basis for specific recognition of both RNA and DNA by a zinc finger protein. *Science 260*, 530–533.
18. McBryant, S. J., Veldhoen, N., Gedulin, B., Leresche, A., Foster, M. P., Wright, P. E., Romaniuk, P. J., and Gottesfeld, J. M. (1995) Interaction of the RNA binding fingers of *Xenopus* transcription factor IIIA with specific regions of 5 S ribosomal RNA. *J. Mol. Biol. 248*, 44–57.
19. Zang, W. Q., Veldhoen, N., and Romaniuk, P. J. (1995) Effects of zinc finger mutations on the nucleic acid binding activities of *Xenopus* transcription factor IIIA. *Biochemistry 34*, 15545–15552.
20. Searles, M. A., Lu, D., and Klug, A. (2000) The role of the central zinc fingers of transcription factor IIIA in binding to 5 S RNA. *J. Mol. Biol. 301*, 47–60.
21. Hamilton, T. B., Turner, J., Barilla, K., and Romaniuk, P. J. (2001) Contribution of individual amino acids to the nucleic acid binding activities of the *Xenopus* zinc finger proteins TFIIIA and p43. *Biochemistry 40*, 6093–6101.
22. Lu, D., Searles, M. A., and Klug, A. (2003) Crystal structure of a zinc finger-RNA complex reveals two modes of molecular recognition. *Nature 426*, 96–100.
23. Johnston, C. M., Shimeld, S. M., and Sharpe, P. T. (1998) Molecular evolution of the ZFY and ZNF6 gene families. *Mol. Biol. Evol. 15*, 129–137.
24. Weiss, M. A., Mason, K. A., Dahl, C. E., and Keutmann, H. T. (1990) Alternating zinc finger motifs in the human male-associated protein ZFY. *Biochemistry 29*, 5660–5664.
25. Kochoyan, M., Havel, T. F., Nguyen, D. T., Dahl, C. E., Keutmann, H. T., and Weiss, M. A. (1991) Alternating zinc fingers in the human male associated protein ZFY: 2D NMR structure of an even finger and implications for jumping-linker DNA recognition. *Biochemistry 30*, 3371–3386.
26. Kochoyan, M., Keutmann, H. T., and Weiss, M. A. (1991) Alternating zinc fingers in the human male-associated protein ZFY: HX_3H and HX_4H motifs encode a local structural switch. *Biochemistry 30*, 9396–9402.
27. Taylor Harris, P., Swift, S., and Ashworth, A. (1995) Zfy1 encodes a nuclear sequence-specific DNA binding protein. *FEBS Lett. 360*, 315–319.
28. L'Haridon, M., Paul, P., Xerri, J. G., Dastot, H., Dolliger, C., Schmid, M., de Angelis, N., Grollet, L., Sigaux, F., Degos, L., and Gazin, C. (1996) Transcriptional regulation of the MHC class I HLA-A11 promoter by the zinc finger protein ZFX. *Nucleic Acids Res. 24*, 1928–1935.
29. Hamilton, T., Barilla, K., and Romaniuk, P. (1995) High affinity binding sites for the Wilms' tumour suppressor protein WT1. *Nucleic Acids Res. 23*, 277–284.
30. Ho, S., Hunt, H., Morton, R., Pullen, J., and Pease, L. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene 77*, 51–59.
31. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem. 72*, 248–254.
32. Hamilton, T. B., Borel, F., and Romaniuk, P. J. (1998) Comparison of the DNA binding characteristics of the related zinc finger proteins WT1 and EGR1. *Biochemistry 37*, 2051–2058.
33. Hall, K. B., and Kranz, J. K. (1999) Nitrocellulose filter binding for determination of dissociation constants. *Methods Mol. Biol. 118*, 105–114.
34. Weiss, J. N. (1997) The Hill equation revisited: uses and misuses. *FASEB J. 11*, 835–841.
35. Scatchard, G. (1949) The attractions of proteins for small molecules and ions. *Ann. N.Y. Acad. Sci. 51*, 660–672.
36. Record, M. T., Jr., Lohman, T. M., and de Haseth, P. (1976) Ion effects on ligand-nucleic acid interactions. *J. Mol. Biol. 107*, 145–158.
37. Record, M. T., Jr., Anderson, C. F., and Lohman, T. M. (1978) Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening, and ion effects on water activity. *Q. Rev. Biophys. 11*, 103–178.
38. Kochoyan, M., Keutmann, H. T., and Weiss, M. A. (1991) Alternating zinc fingers in the human male associated protein ZFY: refinement of the NMR structure of an even finger by selective deuterium labeling and implications for DNA recognition. *Biochemistry 30*, 7063–7072.
39. Qian, X., and Weiss, M. A. (1992) Two-dimensional NMR studies of the zinc finger motif: solution structure and dynamics of mutant ZFY domains containing aromatic substitutions in the hydrophobic core. *Biochemistry 31*, 7463–7476.
40. Moore, M., Klug, A., and Choo, Y. (2001) Improved DNA binding specificity from polyzinc finger peptides by using strings of two-finger units. *Proc. Natl. Acad. Sci. U.S.A. 98*, 1437–1441.
41. Lohman, T. M., deHaseth, P. L., and Record, M. T., Jr. (1980) Pentalysine-deoxyribonucleic acid interactions: A model for the general effects of ion concentrations on the interactions of proteins with nucleic acids. *Biochemistry 19*, 3522–3530.