Design of polyzinc finger peptides with structured linkers

Michael Moore*†, Yen Choo*, and Aaron Klug

Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, United Kingdom

Zinc finger domains are perhaps the most versatile of all known DNA binding domains. By fusing up to six zinc finger modules, which normally recognize up to 18 bp of DNA, designer transcription factors can be produced to target unique sequences within large genomes. However, not all continuous DNA sequences make good zinc finger binding sites. To avoid having to target unfavorable DNA sequences, we designed multizinc finger peptides with linkers capable of spanning long stretches of nonbound DNA. Two three-finger domains were fused by using either transcription factor IIIA for the *Xenopus* 5S RNA gene (TFIIIA) finger 4 or a non-sequence-specific zinc finger as a ''structured'' linker. Our gel-shift results demonstrate that these peptides are able to bind with picomolar affinities to target sequences containing 0–10 bp of nonbound DNA. Furthermore, these peptides display greater sequence selectivity and bind with higher affinity than similar six-finger peptides containing long, flexible linkers. These peptides are likely to be of use in understanding the behavior of polydactyl proteins in nature and in the targeting of human, animal, or plant genomes for numerous applications. We also suggest that in certain polydactyl peptides an individual finger can ''flip'' out of the major groove to allow its neighbors to bind shorter, nontarget DNA sequences.

gene regulation | DNA binding site | zinc finger | polydactyl peptides

With the prospect of gene therapy treatments becoming ever more real, several artificial approaches for controlling endogenous gene expression have emerged (1, 2). In nature, however, both activation and repression of genes is generally accomplished at the level of transcription, through the use of transcription factors and associated proteins. Several well conserved DNA-binding motifs have been characterized, and of these, the most common is the transcription factor IIIA for the *Xenopus* 5S RNA gene (TFIIIA)-type Cys_2-His_2 family of zinc finger peptides (3). It is perhaps their simple, stable, modular structure that makes them so universally applicable in nature, and it is these reasons that have made them, to date, the most promising "designer" transcription factors (4–7). The zinc finger module comprises $\approx$30 aa, which form a $\beta\beta\alpha$-fold stabilized by hydrophobic interactions and the chelation of a zinc ion (3, 8). Base-specific DNA recognition by an array of zinc fingers is through side chains along the α-helices (usually at the −1, 2, 3, and 6 positions) to overlapping 4-bp subsites (9). However, zinc fingers do not only bind to duplex DNA. For example, TFIIIA, a nine-finger protein binds both DNA and RNA (10, 11), and the Ikaros protein contains a two-finger domain that is used in homodimerization (12). Hence, it appears that the functions of zinc fingers are far more diverse than first thought.

Several groups now have attempted to create high-affinity binding peptides to regulate gene transcription by joining six or more zinc fingers in three-finger units (5, 6, 13, 14). However, with the exception of one study, the peptides have showed only modest improvements in affinity, relative to their three-finger components. Kamiuchi *et al*. (14) synthesized six- and nine-finger peptides by multimerizing the three-finger Sp1 protein, which displayed affinity enhancements of only 20- and 30-fold, relative to the wild-type three-finger Sp1 peptide. Similarly Liu *et al*. (13)

produced a six-finger peptide that bound its target site $\approx$70-fold tighter than its three-finger components. Both these studies used canonical linkers between all fingers within the array, and it seems likely that this linker is not sufficiently long to allow the optimal binding of all six fingers. In contrast, Kim and Pabo (6) synthesized a six-finger peptide with an extended linker peptide (of either 9 or 12 aa) between fingers 3 and 4 of the construct. The authors reported that these peptides bound their 18-bp target site at least 6,000-fold tighter than a three-finger peptide, although it appears that the peptides affinities were measured by different procedures. These extended linkers were able to span 1- or 2-bp insertions.

We wondered whether it would be possible to design six-finger peptides that are able to span more than 2 bp (e.g., 3–10 bp) of DNA between their recognition subsites, while still binding with high affinity. One advantage of this would be that the protein engineer would have more scope to select favorable binding sites amongst generally unfavorable regions of genomic DNA sequence. A long, flexible linker would allow the attached zinc finger domains to search out their respective target subsites. However, we were concerned by the possibility of losing binding affinity through the entropic factors associated with by using a flexible linker. To overcome this concern, we sought to employ a relatively rigid or structured linker that might act as a bridge between the two DNA-binding domains. For this purpose we used TFIIIA finger 4, which in the wild-type protein does not bind DNA, but crosses the DNA minor groove, allowing fingers 3 and 5 to bind in the major groove some 6 bp apart (15). We also created a non-sequence-specific zinc finger, called serF (by substituting the DNA binding residues of Zif268 finger 2 for serine), and used this as an alternative structured linker. In this work we synthesized a number of polydactyl peptides by using either long, flexible linkers (composed of runs of glycine and serine) or structured linkers to separate the DNA-binding domains. These peptides were then tested for binding affinity to contiguous and noncontiguous target sequences with up to 10 bp spans of nonbound DNA. Our results show that flexible linkers do not discriminate between different length stretches of nonbound DNA, whereas structured linkers demonstrate subtle preferences for particular DNA spans. In addition, peptides containing structured linkers were found to bind their complete target sequences with greater affinity than similar peptides containing flexible linkers. The results also suggest that individual zinc fingers are able to "flip" out of the major groove of DNA when they do not have a spatially correlated binding site,

---

Table 1. The nomenclature and descriptions of the peptides used in the study

Peptide domain name	Description
TF(1–3)	The first three fingers of TFIIIA
TF(1–4)	The first four fingers of TFIIIA
ZIF	The three fingers of Zif268
MUT	A three finger mutant of Zif268
flex	The flexible linker peptide -TG(GSG)$_5$ERP-
F4	TFIIIA finger 4 with natural flanking linkers
serF	The second finger of Zif268 with the amino acids at positions −1, 2, 3 and 6 mutated to serine, flanked by -TGERP- linkers
mutF4	TFIIIA finger 4 with -TGERP- flanking linkers
mutserF	The serine finger (serF) with TFIIIA finger 4 flanking linkers

to accommodate neighboring zinc fingers, which then bind to their optimal target sequences. We have investigated this phenomenon more thoroughly as reported (16).

Materials and Methods

Design, Construction, and Cloning of Zinc Finger Genes. TF(1–4)-ZIF was created by two stages of PCR amplification. In the first, the N-terminal four fingers of wild-type TFIIIA and the three fingers of wild-type Zif268 DNA were amplified. Overlap PCR was then used to generate the complete seven-finger construct. The ZIF-F4-MUT and ZIF-mutF4-MUT constructs were made by three separate PCR amplifications of the three fingers of Zif268 (ZIF), the three fingers of a Zif268 mutant peptide (MUT), and the fourth finger of TFIIIA. Two sequential overlap PCR reactions were then used to fuse the separate units together, creating seven-finger constructs. The ZIF-serF-MUT construct was made by PCR amplification of the three-fingers of wt Zif268 and the Zif268 mutant, creating *Eag*I sites at their C and N termini, respectively. The structured linker, serF, was created by annealing two compatible oligonucleotides (5′-GG CCG TTC CAG TGT CGA ATC TGC ATG CGT AAC TTC AGT TCT AGT AGC TCT CTT ACC AGC CAC ATC CGC ACC CAC ACA GGT GAG C-3′ and 5′-GG CCG CTC ACC TGT GTG GGT GCG GAT GTG GCT GGT AAG AGA GCT ACT AGA ACT GAA GTT ACG CAT GCA GAT TCG ACA CTG GAA C-3′), which created *Eag*I sites at each end. This finger was inserted between the *Eag*I cut ZIF and MUT three-finger units to create the complete seven-finger construct. The ZIF-mutserF-MUT clone was made by PCR amplification of the ZIF, MUT, and serF-structured linker fragments to create mutant ends. These three fragments were joined by two sequential rounds of overlap PCR as above. The TF(1–3)-flex-ZIF and ZIF-flex-MUT constructs were created by PCR amplification of the first three fingers of TFIIIA, the three fingers of Zif268, or the three fingers of the MUT-clone—using appropriate oligonucleotides—which were designed to generate the flexible 20-aa linker peptide -TG(GSG)$_5$ERP- and *Eag*I sites at the position to be joined. The required six-finger constructs were synthesized by digesting the PCR products with *Eag*I and ligating at that site. The nomenclature and description of each peptide domain described is displayed in Table 1. All zinc-finger constructs were digested with *Xba*I and *Eco*RI restriction enzymes and inserted into the similarly digested, eukaryotic expression vector pcDNA 3.1(−) (Invitrogen). The sequences of all constructs were confirmed by dideoxy sequencing.

Template Preparation and Protein Expression. Plasmids containing zinc finger constructs were purified from *Escherichia coli* XL1-

Table 2. The binding site sequences used in gel-shift experiments with the TFIIIA-ZIF fusion peptides and the binding affinities obtained

Binding site name	Binding site sequence*	Apparent K_d, pM	
		TF(1–4)-ZIF	TF-flex-ZIF
ZIF	GCGTGGGCG	2,000	1,800
TF5Z	GCGTGGGCG**X**$_5$GGATGGGAGAC	21	63
TF6Z	GCGTGGGCG**X**$_6$GGATGGGAGAC	17	68
TF7Z	GCGTGGGCG**X**$_7$GGATGGGAGAC	3	57
TF8Z	GCGTGGGCG**X**$_8$GGATGGGAGAC	3	61
TF9Z	GCGTGGGCG**X**$_9$GGATGGGAGAC	15	58

*Nonbound DNA bases in the target sequence are shown by a boldface **X**. The exact base composition of these gaps was found to have no significant effect on peptide affinity.

Blue cells using the Qiagen spin miniprep system and quantitated by *A* at 260 nm. Protein expression was performed *in vitro* by coupled transcription and translation in the TNT Quick Coupled Transcription/Translation System (Promega), according to the manufacturer's instructions, except that the medium was supplemented with 500 μM ZnCl$_2$.

Gel-Shift Assays. All peptides were assayed by using ^{32}P end-labeled synthetic oligonucleotide duplexes containing the required binding site sequences. The sequences of the binding sites are shown in Tables 2 and 3.

DNA binding reactions contained the appropriate zinc finger peptide, radiolabeled binding site, and 1-μg competitor DNA [poly(dI-dC)] in a total volume of 10 μl, which contained 20 mM Bis-Tris propane (pH 7.0)/100 mM NaCl/5 mM MgCl$_2$/50 μM ZnCl$_2$/5 mM DTT/0.1 mg/ml BSA/0.1% Nonidet P-40. Incubations were performed at room temperature for 1 h.

(i) Active peptide concentration. To determine the concentration of zinc finger peptide produced in the *in vitro* expression system, crude protein samples were used in gel-shift assays against a dilution series of the appropriate binding site. Binding site concentration was always well above the dissociation constant (K_d) of the peptide, but ranged from a higher concentration than the peptide (80 mM), at which all available peptide binds DNA, to a lower concentration (3–5 mM), at which all DNA is bound. Controls were carried out to ensure that binding sites

Table 3. The binding site sequences used in gel-shift experiments with the ZIF-MUT fusion peptides and the binding affinities obtained

Binding site name	Binding site sequence*	Apparent K_d, pM	
		ZIF-F4-MUT	ZIF-serF-MUT
ZIF	GCGTGGGCG	2,200	2,000
ZM	GCGGACGCGGCGTGGGCG	11	7
Z1M	GCGGACGCG**X**GCGTGGGCG	6	4
Z2M	GCGGACGCG**X**$_2$GCGTGGGCG	7	6
Z3M	GCGGACGCG**X**$_3$GCGTGGGCG	5	4
Z4M	GCGGACGCG**X**$_4$GCGTGGGCG	13	3
Z5M	GCGGACGCG**X**$_5$GCGTGGGCG	16	8
Z6M	GCGGACGCG**X**$_6$GCGTGGGCG	17	7
Z7M	GCGGACGCG**X**$_7$GCGTGGGCG	5	3
Z8M	GCGGACGCG**X**$_8$GCGTGGGCG	5	6
Z9M	GCGGACGCG**X**$_9$GCGTGGGCG	5	4
Z10M	GCGGACGCG**X**$_{10}$GCGTGGGCG	4	3

*Nonbound DNA bases in the target sequence are shown by a boldface **X**. The exact base composition of these gaps was found to have no significant effect on peptide affinity.

were not shifted by the *in vitro* extract in the absence of zinc finger peptide. The reaction mixtures then were separated on a 7% native polyacrylamide gel. Radioactive signals were quantitated by PhosphorImager analysis (Molecular Dynamics) to determine the amount of shifted binding site and, hence, the concentration of active zinc finger peptide.

(ii) Binding affinity and specificity. Dissociation constants were determined in parallel to the calculation of active peptide concentration. Serial 3-, 4-, or 5-fold dilutions of crude peptide were made and incubated with radiolabeled binding site (0.1–500 pM, depending on the peptide), as above. Samples were run on 7% native polyacrylamide gels and the radioactive signals were quantitated by PhosphorImager analysis. The data then were analyzed according to linear transformation of the binding equation and plotted in CA-CRICKET GRAPH III (Computer Associates, Islandia, NY) to generate the apparent dissociation constants. The K_d values reported are the average of at least two separate studies.

Results

Design of the Fusion Constructs. The goal was to link two three-finger DNA-binding domains so as to bind two DNA subsites separated by long (3–10 bp) stretches of nonbound DNA, with the high affinity expected from a six-finger construct. We linked the two three-finger domains by using either a long, flexible linker or a zinc-finger domain that does not bind DNA in a sequence-specific manner.

The crystal structure of the first six fingers of TFIIIA bound to DNA reveals that finger 4 does not make any specific base contacts with DNA. Instead, the finger acts as a bridge over the DNA minor groove, allowing fingers 3 and 5 to bind in the major groove 6 bp apart (15). To test the theory that TFIIIA finger 4 could be used to link two DNA-binding domains, the first four fingers of TFIIIA were fused N-terminally to the three fingers of Zif268, creating the construct TF(1–4)-ZIF. TFIIIA finger 4 also was inserted between Zif268 and a mutant Zif268 clone (MUT), to test its effect when no longer in the context of TFIIIA; this product was called ZIF-F4-MUT. MUT is a phage-selected variant of Zif268, which binds the DNA sequence 5′-GCG GAC GCG-3′ (17). The linkers found naturally in TFIIIA between finger 3 and finger 4 (-NIKICV-) and between finger 4 and finger 5 (-TQQLP-) were retained in both the above peptides. To test the role of these natural linker sequences, we engineered the construct ZIF-mutF4-MUT, in which the canonical-like sequences -TGERP- were used to flank TFIIIA finger 4.

The concept of a zinc finger acting as a structured linker was further tested by using a modified zinc finger from Zif268, rather than TFIIIA finger 4. Most natural zinc fingers modules would be predicted to span 3 or 4 bp, so a structured linker based on a Zif268 finger may have a similar span. To test this theory, we took the sequence of Zif268 finger 2 and mutated the DNA-binding residues at positions −1, 2, 3, and 6 to serine residues. Serine probably would not interact specifically with a particular DNA base and, therefore, this modified finger should span any DNA sequence. This new finger was flanked by -TGERP- linkers and inserted between the Zif268 and the Zif268 mutant in the construct ZIF-serF-MUT. A similar construct also was created, ZIF-mutserF-MUT, which contained the TFIIIA finger 4 flanking sequences -NIKICV- and -TQQLP-.

We also made two peptides containing flexible linkers to compare with the structured-linker design for spanning long stretches of DNA. These constructs, TF(1–3)-flex-ZIF and ZIF-flex-MUT, contained the 20-aa sequence -TG(GSG)₅ERP- between their respective three-finger domains. This sequence was chosen simply on the basis that it is flexible and long enough to span 1–10 bp of DNA.

All peptides were targeted against their contiguous binding sites and against sites with 1–10 bp of nonbound DNA between
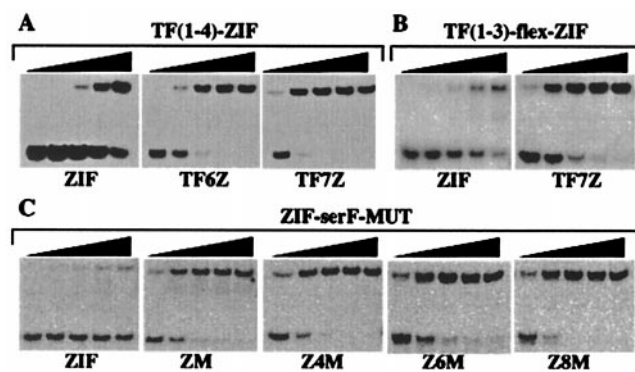

Fig. 1. A selection of DNA binding studies by gel-shift assay. (*A*) Shown are 5-fold dilutions of TF(1–4)-ZIF (from 5.5 nM–9 pM) against 20 pM ZIF binding site, 2 pM TF6Z, and 2 pM TF7Z. (*B*) Shown are 5-fold dilutions of TF(1–3)-flex-ZIF (from 5 nM–8 pM) against 20 pM ZIF and 2 pM TF7Z. (*C*) Shown are 5-fold dilutions of ZIF-serF-MUT (from 1 nM–1.6 pM) against 10 pM ZIF, 0.4 pM ZM, 0.4 pM Z4M, 0.4 pM Z6M, and 0.4 pM Z8M.

their target subsites to determine the optimal span of the structured linkers. The names and sequences of all binding sites are shown in Tables 2 and 3.

Binding Affinity of the TFIIIA/ZIF Fusion Peptides. Peptide samples from the same *in vitro* synthesis reaction were used to calculate active peptide concentration and also to determine binding affinity. A preliminary experiment was conducted with the three-finger Zif268 peptide against its 9-bp binding site as a form of "protocol calibration." This gave a value for the K_d of Zif268 of 0.45 nM, which is within the range expected for this peptide. The TF(1–4)-ZIF and TF(1–3)-flex-ZIF peptides were tested against the noncontiguous TF-5,6,7,8,9-Z sites. In these first experiments, the DNA composition of the nonbound region was based on the endogenous TFIIIA target site. The results clearly show that the TF(1–4)-ZIF peptide has a preference for non-contiguous sites separated by 7- or 8-bp gaps, which were bound with a K_d of ≈3 pM (Table 2). The target sites with 5-, 6-, or 9-bp gaps were bound at least 5-fold weaker (Fig. 1*A*). In contrast, the TF(1–3)-flex-ZIF peptide showed no preference for a particular DNA span, binding all noncontiguous sites with affinities of ≈60 pM (Fig. 1*B*). Further studies were conducted on binding sites with various sequences in the nonbound region of the DNA target site. These studies demonstrated that the peptides had no preference for particular sequence compositions within this nonbound region (data not shown). Both constructs bound the Zif268 half-site with similar affinity, as expected.

Binding Affinity of the ZIF/MUT Fusion Peptides. The first binding study was conducted on ZIF-F4-MUT to determine the optimal span of TFIIIA finger 4 in this construct. This peptide was titrated against the continuous 18-bp ZM binding site and noncontinuous binding sites with 1–10 bp of nonbound DNA. Our results demonstrate that this peptide has little preference for a particular span of DNA, although the highest affinity binding was observed for sites containing 3-bp or ≥7-bp insertions (Table 3). The fact that this peptide was able to bind with such high affinity to sites with <3-bp gaps was highly unexpected. The slight reduction in binding affinity observed in these examples is presumably because the 1- to 2-bp gaps are too small to accommodate a zinc finger in the DNA major groove. In these circumstances, it seems likely that the nonbinding finger actually flips out of the DNA leaving the remaining fingers to bind the target site. The slight reduction in affinity for sites with 5- or 6-bp gaps is probably because TFIIIA finger 4 has to stretch half a

helical turn around the DNA. For longer gaps the finger is likely to span the minor groove as is seen in wild-type TFIIIA.

A further set of binding studies then was carried out on the construct containing the nonspecific zinc finger linker ZIF-serF-MUT. Although this construct was expected to target (primarily) noncontiguous sequences containing three or four base pairs of nonbound DNA, it was tested against all of the binding sites from ZM to Z10M. Our gel-shift data again demonstrate that this peptide is able to bind its optimal targets with very high affinity (3–4 pM) and show a similar trend in binding affinity to the ZIF-F4-MUT peptide (Fig. 1C). However, this peptide was able to bind its least favorable sites with slightly greater affinity than observed for the previous peptide (Table 3).

To determine whether the structured linkers themselves or the linker sequences flanking them were responsible for the slightly lower gap-selectivity of the ZIF-serF-MUT peptide, the -NI-KICV- and -TQQLP- linkers were exchanged with the -TGERP-linkers to give the constructs ZIF-mutF4-MUT and ZIF-mutserF-MUT. These new peptides were targeted against all 11 binding sequences, as above. Our data for these constructs (not shown) demonstrated no significant differences in binding affinity for the different target sequences when compared with the original peptide constructs. Hence, these flanking linkers seem to have no real impact on the gap-selectivity of the structured linkers.

Finally, the ZIF-flex-MUT peptide was examined in the same way as the structured-linker peptides above. This peptide, as with the TF(1–3)-flex-ZIF peptide, displayed no preference for a particular length of DNA span and bound all sites with affinities of ≈50 pM. This 3- to 10-fold reduction in affinity (compared with peptides connected by structured linkers) is probably attributable to the increased conformational freedom of this peptide, which makes DNA binding less entropically favorable.

Discussion

To date, several groups have created six- (or nine-) finger fusion peptides to bind long stretches of DNA with high affinity (6, 13, 14). However, the affinities of these constructs vary greatly and have generally been far weaker than expected. In addition, all of these peptides have targeted either contiguous DNA sequences or those containing just one or two nucleotides of unbound DNA. We sought to increase the utility of polyzinc finger peptides by creating fusion peptides that are able to bind with high affinity to target sequences in which their binding subsites are separated by long (up to 10 bp) stretches of DNA. We also hoped to create linkers that would show a preference for a particular length of DNA span, so that they maintain a high degree of specificity. We decided that a structured or rigid linker might fulfill these requirements. Having examined the crystal structure of the first six fingers of TFIIIA bound to DNA (15), we decided that TFIIIA finger 4 may be a suitable candidate for a structured linker to span long (>5 bp) stretches of DNA.

We first created a fusion peptide comprising the first four fingers of TFIIIA and the three fingers of Zif268, called TF(1–4)-ZIF. This peptide bound DNA with high affinity and showed a preference for sites containing 7 or 8 bp of nonbound DNA. In contrast, a similar construct that contained a 20-residue flexible linker, TF(1–3)-flex-ZIF, bound its full-length target sites somewhat weaker. These data suggest that TFIIIA finger 4 is a suitable structured linker for spanning long stretches of DNA and, furthermore, that TF(1–4)-ZIF would make a good scaffold for "designer" transcription factors that bind DNA with 7 or 8 bp of nonbound DNA.

Noting the utility of the TFIIIA finger 4 linker, we decided to test the ability of a zinc-finger module from Zif268 to act as a structured linker. A zinc finger from Zif268 was mutated to make it non-sequence-specific and then used to link the three wild-type fingers of Zif268 to a three-finger mutant of Zif268 (MUT).

Surprisingly, this new peptide was able to bind with similar affinity to the continuous 18-bp sequence comprising the Zif268 and MUT sites and to all of the noncontiguous sites with 1- to 10-bp gaps. The fact that this peptide could bind tightly to the contiguous binding site and the sites with just 1- or 2-bp gaps suggests that the "serine-finger" is able to flip out of the major groove to make space for the binding of its neighboring fingers. These data indicate that within a zinc finger array, redundant fingers can make way for stronger DNA-binding domains. When the binding subsites are separated by 7–10 bp of DNA it seems likely that the redundant finger lies across the surface of the DNA in a manner analogous to TFIIIA finger 4 (15).

We also created a fusion construct, ZIF-F4-MUT, which used TFIIIA finger 4 as a linker between two Zif-type domains. This peptide displayed little discrimination for the length of DNA span separating the binding subsites, although a trend in the binding affinities of the peptide was apparent. All peptides connected by zinc finger modules show a preference for sequences containing 3 bp or >6-bp gaps. These findings probably correspond to modes in which the zinc-finger-linker can sit "normally" in the major groove or is able to bridge the minor groove.

It has been proposed that the relatively hydrophobic linkers flanking TFIIIA finger 4 may constrain finger 4 into its orientation across the minor groove, as observed in the crystal structure of Nolte et al. (15). Hence, we further investigated the conformational freedom of zinc fingers by swapping the linker sequences flanking wild-type TFIIIA finger 4 and the serine-finger. However, it was found that the linker sequences flanking TFIIIA finger 4 do not appear to make a significant contribution to its preferred span in these constructs.

A predicted benefit of using structured linkers was that of increased binding affinity over peptides containing long, flexible linkers. This finding was confirmed by the binding results from the two peptides containing 20-residue flexible linkers, which bound their full-length targets between 3- and 10-fold weaker than peptides with structured linkers.

Polyzinc finger peptides are likely to become increasingly important in gene therapy and the creation of transgenic organisms. Given the difficulty of engineering zinc finger peptides to bind to all possible DNA sequences (17, 18), it would be advantageous to synthesize peptides capable of spanning long regions of DNA, while still binding with high affinity. This will allow the selection of favorable DNA target sites that may be several nucleotides apart. In this paper, we have presented data that shows that structured linkers can be incorporated into zinc finger fusion peptides. These linkers allow the separate DNA-binding domains to bind with high affinity to sites separated by 1 to 10 bp of nonbound DNA. The ability of these structured-linker fusion peptides to span such long stretches of DNA may be particularly advantageous for the targeting of natural promoter sequences. For example, the zinc finger protein Sp1 binds GC box DNA, which can appear in multiple copies in the promoter sequences upstream of a variety of cellular and viral genes (19, 20). Similarly, the promoter for the herpes simplex virus 40 early genes contains three 21-bp repeats that include GC boxes. Linking zinc finger peptides that recognize such regions could create powerful designer transcription factors. TFIIIA finger 4 may be a particularly useful structured linker as it shows a marked preference for 7- or 8-bp DNA spans.

Our studies also suggest that zinc fingers that do not have a binding site are able to flip in or out of the DNA major groove to accommodate neighboring fingers within the DNA-binding domain. This means that certain zinc finger arrays will bind reasonably tightly to truncated or mutated binding sites. The protein engineer can again take advantage of this feature of zinc-finger arrays, for instance to engineer zinc fingers that bind to a series of related but different binding sites. Nature almost

certainly takes advantage of this phenomenon to evolve zinc finger transcription factors that regulate multiple genes from nonidentical promoters. It is also worth noting that many natural polydactyl proteins that have been isolated contain zinc fingers whose roles are not yet understood. For example, GL1 contains five tandem zinc fingers, but in the crystal structure of this protein only two of these bind to DNA in the classical, base-specific manner (21). The results presented here also suggest that there may be a broad repertoire of roles for zinc finger domains within the cell. Furthermore, as polydactyl peptides are pro-duced more frequently for *in vivo* use, the "flexible" nature of zinc-finger modules raises important questions regarding the specificity of high-affinity binding peptides. Our other research addresses this issue and presents polydactyl peptides that bind with far greater specificity than previously designed six-finger peptides (16).

1. Matteucci, M. D. & Wagner, R. W. (1996) *Nature (London)* **384,** 20–22.
2. Gottesfeld, J. M., Neely, L., Trauger, J. W., Baird, E. E. & Dervan, P. B. (1997) *Nature (London)* **387,** 202–205.
3. Miller, J., McLachlan, A. D. & Klug. A. (1985) *EMBO J.* **4,** 1609–1614.
4. Choo, Y., Sanchez-Garcia, I. & Klug, A. (1994) *Nature (London)* **372,** 642–645.
5. Beerli, R. R., Dreier, B. & Barbas, C. F., III (2000) *Proc. Natl. Acad. Sci. USA* **97,** 1495–1500. (First Published January 31, 2000; 10.1073/pnas.040552697)
6. Kim, J.-S. & Pabo, C. O. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 2812–2817.
7. Kang, J. S. & Kim, J.-S. (2000) *J. Biol. Chem.* **275,** 8742–8748.
8. Lee, M. S., Gippert, G. P., Soman, K. V., Case, D. A. & Wright, P. E. (1989) *Science* **245,** 635–637.
9. Isalan, M., Choo, Y. & Klug, A. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 5617–5621.
10. Hanas, J. S., Gaskins, C. J., Smith, J. F. & Ogilvie, M. K. (1992) *Prog. Nucleic Acid Res. Mol. Biol.* **43,** 205–239.
11. Picard, B. & Wegnez, M. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 241–245.
12. Sun, L., Liu, A. & Georgopoulos, K. (1996) *EMBO J.* **15,** 5358–5369.
13. Liu, Q., Segal, D. J., Ghiara, J. B. & Barbas, C. F., III (1997) *Proc. Natl. Acad. Sci. USA* **94,** 5525–5530.
14. Kamiuchi, T., Abe, E., Imanishi, M., Kaji, T., Nagaoka, M. & Sugiura, Y. (1998) *Biochemistry* **37,** 13827–13834.
15. Nolte, R. T., Conlin, R. M., Harrison, S. C. & Brown, R. S. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 2938–2943.
16. Moore, M., Klug, A. & Choo, Y. (2001) *Proc. Natl. Acad, Sci. USA* **98,** 1437–1441.
17. Choo, Y. & Klug, A. (1994) *Proc. Natl. Acad, Sci. USA* **91,** 11168–11172.
18. Segal, D. T., Dreier, B., Beerli, R. R. & Barbas, C. F., III (1999) *Proc. Natl. Acad. Sci. USA* **96,** 2758–2763.
19. Kadonaga, J. T., Jones, K. A. & Tjian, R. (1986) *Trends Biochem. Sci.* **11,** 20–23.
20. Bucher, P. (1990) *J. Mol. Biol.* **212,** 563–578.
21. Pavletich, N. P. & Pabo, C. O. (1991) *Science* **261,** 1701–1707.