DNA-Bending Finger: Artificial Design of 6-Zinc Finger Peptides with Polyglycine Linker and Induction of DNA Bending†

Miki Imanishi, Yuichiro Hori, Makoto Nagaoka, and Yukio Sugiura*

Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan

Received December 30, 1999

ABSTRACT: DNA structural changes such as bending play an important role in various biological reactions. Not only protein binding to its specific DNA sequence but also DNA bending induced by the protein is indispensable for unique gene expression. Therefore, an artificial protein that induces a DNA conformational change is interesting as a transcriptional regulator of a specific gene. We created 6-zinc finger proteins, Sp1ZF6(Gly)*n* ($n = 4, 7, 10$), by connecting two DNA binding domains of transcription factor Sp1 with flexible polyglycine peptide linkers, and their effects on DNA structure were compared with that of native 3-zinc finger Sp1(530−623). Gel electrophoretic methods revealed that Sp1ZF6(Gly)7 and Sp1ZF6(Gly)-10 bind to two distal GC boxes and result in DNA bending. Evidently, the hydroxyl radical footprinting analysis demonstrated that hypersensitive cleavage was observed at the 5′-TA step in the intervening region bound by Sp1ZF6(Gly)7 or Sp1ZF6(Gly)10. The phasing assays strongly suggested that the induced DNA bending was directed toward the major groove and that Sp1ZF6(Gly)7 caused the most drastic directional change in DNA bending. Of special interest are the facts that the newly designed 6-finger peptides Sp1ZF6(Gly)7 and Sp1ZF6(Gly)10 can induce DNA bending at the intervening region of the two distal binding sites and that the linker length between two 3-zinc finger motifs has a crucial effect on the entire DNA-bending direction. Such DNA-bending fingers may be feasible for use as a gene expression regulator based on the structural change in DNA in the future.

DNA bending is very important for various biological reactions, such as transcription, recombination, and replication. In particular, many transcription factors have been known to induce DNA bending and to support formation of the transcriptional initiation complex (*1*−*4*). A ligand that induces DNA bending in a site-specific manner would be expected to be an artificial regulator of a specific gene transcription. Previously, a man-designed DNA-bending agent was created using tethered TFO[1] (*5*−*7*). DNA modification by charge neutralization at the phosphodiester backbone also brought about DNA bending (*8*−*10*). In the triple helix formation, target sequences are virtually limited to homopurine-homopyrimidine on the basis of the Hoogsteen hydrogen bonding. On the other hand, binding of a protein to DNA is stabilized by various interactions, and hence many kinds of sequences are recognized as target sequences. Such DNA binding proteins bind to their target DNA sequences in a more flexible fashion than TFO. Herein, an artificial zinc finger protein containing two DNA binding domains connected with a variable linker was newly designed, and the effect on the DNA conformation was investigated.

A zinc finger motif of the Cys_2-His_2 type, one of the most common DNA binding motifs found in eukaryotes, offers an attractive framework for the design of a novel sequence-specific DNA binding protein. The X-ray crystal structures of the Zif268− and the GLI−DNA complexes clearly revealed the characteristic DNA binding modes of the Cys_2-His_2 zinc finger proteins as follows: (1) one finger binds to a DNA major groove and typically recognizes the overlapping 4 bp of the DNA sequence; (2) DNA recognition is mediated through base contacts with the side chains of specific amino acids located on the recognition helix; (3) multiple zinc finger domains are tandemly repeated by simple covalent linkage, namely, connected by the consensus linker region; and (4) the monomeric binding mode of zinc finger proteins offers an asymmetric binding sequence unlike other nucleic acid recognition motifs such as the basic region leucine-zipper and the helix−turn−helix (*11*−*15*). Artificial zinc finger proteins with novel sequence specificity have been generated based on a selection strategy via phage display (*16*−*20*), and those with an extended recognition sequence of more than 18 bp have been created by a structure-based hybridization method (*21*−*26*). Fusion of a zinc finger domain to another functional domain also produced new site-specific DNA cleavage agents (*27*, *28*). Therefore, a zinc finger motif is attractive for creating a novel sequence-specific DNA bending protein, aimed at controlling specific gene expression. In particular, the structure-based method could be applicable to the creation of zinc finger proteins causing various site-specific DNA bendings. In this study, novel 6-zinc finger proteins connected by two 3-zinc finger motifs of transcription factor Sp1 were designed, and the effects of their binding on DNA structure were examined.

† This study was supported in part by a Grant-in-Aid for Priority Project "Biometals" (08249103) and Scientific Research (B) (10470493) from the Ministry of Education, Science, Sports, and Culture, Japan. We are also thankful for the grant from the Houansha Foundation.

* To whom correspondence should be addressed. Phone: 81-774-38-3210. Fax: 81-774-32-3038. E-mail: sugiura@scl.kyoto-u.ac.jp.

[1] Abbreviations: bp, base pair(s); TFO, triple helix forming oligonucleotide.

10.1021/bi992989b CCC: $19.00
Published on Web 03/25/2000

MATERIALS AND METHODS

Chemicals. T4 polynucleotide kinase and restriction enzymes were purchased from New England Biolabs, except for *Age*I obtained from Nippon Gene (Tokyo, Japan). Labeled compound [γ-^{32}P]ATP was supplied by DuPont. The plasmid pBS-Sp1-fl was kindly provided by Dr. R. Tjian. All other chemicals were of commercial reagent grade.

Construction of Genes and Peptide Expression. The pUC-Sp1(530−623), which codes the 3-zinc finger region, was constructed as described previously (*29*). Sp1ZF6(Gly)4, Sp1ZF6(Gly)7, and Sp1ZF6(Gly)10 were constructed by exchanging the TGEKP linker region of Sp1ZF6 (*26*) with the Q(G)*n* (n = 4, 7, 10) linkers. Synthesized *Afl*II−*Mfe*I oligonucleotides containing the linker sequences were purchased from Amersham Pharmacia Biotech. These oligonucleotides were annealed and inserted into pUC-Sp1ZF6 (*26*). These plasmids were renamed pUC-Sp1ZF6(Gly)*n* (n = 4, 7, 10). The DNA fragments coding these proteins were cut out and inserted into the similar digested plasmid pEV-3b (*29*). These zinc finger peptides were overexpressed in *Escherichia coli* strain BL21(DE3)pLysS and purified as described previously (*26, 30*).

Gel Mobility Shift Assays. Gel mobility shift assays were carried out under the following conditions. Each reaction mixture contained 10 mM Tris buffer (pH 8.0), 50 mM NaCl, 0.5 mM β-mercaptoethanol, 0.1 mM $ZnCl_2$, 0.05% Nonidet P-40, 5% glycerol, the substrate 112 bp DNA fragment containing the protein binding sequence (150 nM), and 150−450 nM of the zinc finger peptide. After incubation at 4 °C for 15−39 h, the sample was run on an 8% polyacrylamide gel (a 74:1 acrylamide:bisacrylamide ratio) with 1×TB buffer at 4 °C. The gel was stained with 0.5 μg/mL ethidium bromide solution, followed by destaining with H_2O.

Hydroxyl Radical Footprinting Assay. Hydroxyl radical cleavage reactions were carried out as described previously (*31*). Each reaction mixture contained 10 mM Tris buffer (pH 8.0), 50 mM NaCl, 0.5 mM β-mercaptoethanol, 0.1 mM $ZnCl_2$, 0.05% Nonidet P-40, 5′-end-labeled substrate DNA (20 000 cpm), and 0, 25, or 100 nM of the zinc finger peptides. After incubation at 4 °C for 15 h, the sample was treated with [EDTA·Fe(II)] solution (final concentration 100 μM), sodium ascorbate (final concentration 1.1 mM), and hydrogen peroxide (final concentration 0.003%) for 4 min on ice. The cleavage reaction was stopped by adding thiourea (final concentration 0.05 M) and EDTA (final concentration 2 mM), and ethanol precipitation was performed. Dried DNA was dissolved in urea dye, and each sample was run on a 15% denaturing gel.

Phasing Analysis. For the phasing analyses, the sequence which contains the protein binding site, a spacer region (8−18 bp) of different lengths, 6-poly AT tracts, and an adopter region (18−8 bp) of corresponding length to the spacer region, in that order, was inserted into *Sph*I/*Hin*dIII site of pUC19. The plasmid was renamed the pPhase-protein binding site-[*S*] (where *S* is the number of bp of spacer region).

The phasing analysis was carried out under conditions similar to those of the gel mobility shift assay with 34 nM substrate DNA. The concentration of each protein was 40 nM Sp1(530−623) or 90 nM Sp1ZF6(Gly)*n* (n = 4, 7, 10) for a single GC box, and 25 or 100 nM Sp1(530−623), 70 nM Sp1ZF6(Gly)4, 50 nM Sp1ZF6(Gly)7, or 35 nM Sp1ZF6(Gly)10 for two distal GC boxes. DNA fragments generated from the *Pvu*II digestion of the pPhase-protein binding site-[8−18] plasmid and a PCR product created from the amplification of a pPhase-protein binding site-[12] plasmid as the template with a *Pvu*II primer set (5′-CTGGCACGACAGGTTTCCCG-3′ and 5′-CTGGCGAAAGGGGGATGTGC-3′) were used as the substrate DNA. The relative electrophoretic mobility of a set of phase-sensitive DNA fragments has been expressed as a cosine function of spacer length (*32*). The electrophoretic mobility of each band was normalized to the apparent bp (*7*), and the relative mobilities (R_L) were fitted to the cosine function:

$$R_L = 1 + (\mathrm{Aph}/2)\cdot\cos\{2\pi(S - 8)/10.5 + b\} \quad (1)$$

where Aph is the phasing amplitude, S is the spacer length in bp (8−18 bp), and b reflects the total bending direction. The spacer length between the two bends predicted in-phase (S_i) was determined by eq 2:

$$S_i = (b/2\pi) \times 10.5 + 8 \quad (2)$$

The number of helical turns between the bend centers of the 6-poly AT tract sequence and the examined sequence was determined based on eq 3:

$$H = 33.5/10.34 + (S_i + n)/10.5 \quad (3)$$

where n is the number of bp from the 5′-end of the GC box sequence to the bend center.

RESULTS

Design of 6-Zinc Finger Proteins with Flexible Polyglycine Linker. An attempt to induce DNA bending at the intervening region by protein binding was performed. By linking two 3-zinc fingers of transcription factor Sp1, novel three 6-zinc finger proteins including flexible polyglycine linkers, Sp1ZF6-(Gly)4, Sp1ZF6(Gly)7, and Sp1ZF6(Gly)10, were created (Figure 1). As the target of the artificial proteins, the DNA sequence contains two protein binding sites with a one helical turn intervening. The length of simple linearized 10 amino acids is slightly shorter than that of 1 helical turn of DNA.

Difference in DNA Binding Mode among 6-Zinc Finger Proteins. Sp1(530−623) strongly binds to the GC (5′-GGG-GCG-GGG-3′) sequence (*33*). Sp1ZF6, which has a conserved Krüppel-type linker (TGEKP) between two 3-zinc finger domains of Sp1, binds to a contiguous two-GC (5′-GGGGCGGGG-GGGGCGGGG-3′) sequence (*26*). The target DNA for a newly designed 6-zinc finger protein contains two GC boxes separated by one turn of the DNA helix (10 bp). On the other hand, the DNA involving one GC box is also used as a control because these proteins can bind to a single GC box using only one DNA binding domain in a similar manner.

To test the DNA binding mode of these proteins, gel mobility shift assays were performed. In general, the electrophoretic mobility of a DNA fragment in a polyacrylamide gel depends on the mean square end-to-end distance for the fragment (*34*) and/or the DNA flexibility. A DNA fragment with a stiff or bent structure in the neighborhood of its center migrates more slowly than a flexible and straight DNA fragment with the same bp. The DNA fragment containing

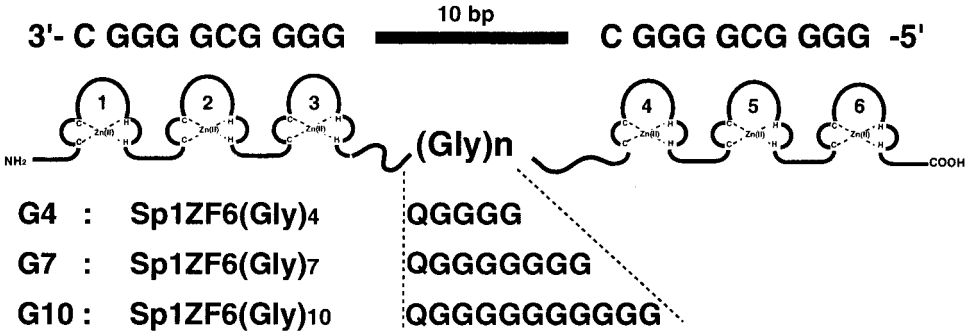

FIGURE 1: Artificial proteins Sp1ZF6(Gly)4, Sp1ZF6(Gly)7, and Sp1ZF6(Gly)10.

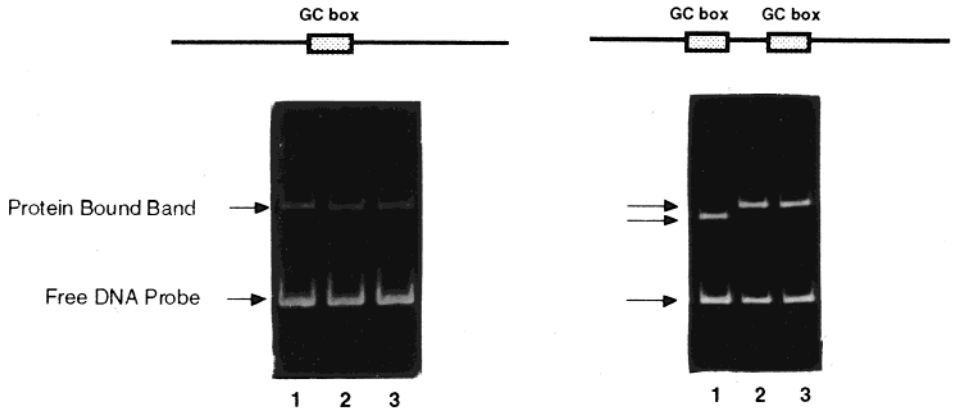

FIGURE 2: Gel mobility shift assays of Sp1ZF6(Gly)*n* binding to a single GC box (left) and to two separated GC boxes (right). Lanes 1, 2, and 3: Sp1ZF6(Gly)4, Sp1ZF6(Gly)7, and Sp1ZF6(Gly)10, respectively.

a single GC box sequence, which formed complexes with Sp1ZF6(Gly)4, Sp1ZF6(Gly)7, or Sp1ZF6(Gly)10, migrated equally in the nondenaturing gel (Figure 2, left). This result suggests that the flexibility or the structure of the protein-bound DNA fragment including a single GC box is almost the same among these 6-zinc finger proteins. On the other hand, the mobilities of the protein complexes with two distal GC boxes separated by one helical turn were apparently different; that is, the complex of Sp1ZF6(Gly)4 migrated faster than that of Sp1ZF6(Gly)7 or Sp1ZF6(Gly)10 (Figure 2, right). Probably, the slower mobility in the Sp1ZF6-(Gly)7− or Sp1ZF6(Gly)10−DNA complex is attributable to an increase in DNA bending or a decrease in DNA flexibility due to the binding of Sp1ZF6(Gly)7 or Sp1ZF6-(Gly)10 in comparison with that of Sp1ZF6(Gly)4.

Positional Analysis of DNA Bending by Hydroxyl Radical Footprinting. To explore whether any distortions of the DNA are induced by binding of these proteins, a hydroxyl radical footprinting experiment was performed. The target DNA fragment contained two distal GC boxes with one helical turn intervening. Figure 3 shows the results of the cleavage patterns by hydroxyl radicals. Enhancement of cleavage activity in the intervening region between the two GC boxes was evidently detected in the complex with Sp1ZF6(Gly)7 or Sp1ZF6(Gly)10. The intensities of the bands were analyzed by NIHimage software, and line scans corresponding to each track on the gel were obtained (Figure 4). The Sp1(530−623)- and Sp1ZF6(Gly)4-bound DNAs showed weak footprints at the GC box sites, but apparent differences in the intervening region between the two GC boxes were not observed compared with the cleavage pattern of the free DNA. Because each protein binds fully to the target DNA at 100 nM (data not shown), both Sp1(530−623) and Sp1ZF6(Gly)4 have little effect structurally on the intervening region of the target DNA. In the case of Sp1ZF6(Gly)7- or Sp1ZF6(Gly)10-bound DNAs, on the other hand, the enhancement of cleavage was clearly detected at two sites, namely, the 5′-TA-3′ sequences in the intervening region of the two GC boxes. This result indicates that Sp1ZF6(Gly)7 and Sp1ZF6(Gly)10 induced DNA distortion at the 5′-TA-3′ sequences. The bands of the two sites slightly shifted in the 3′-direction in comparison with the corresponding bands of the free DNA. In addition, the intensities of Sp1ZF6-(Gly)7-bound bands were stronger than those of Sp1ZF6-(Gly)10-bound bands.

DNA Bending Direction and Magnitude Estimated from Phasing Analysis. Phasing analysis is based on the phase-dependent interaction between a protein-induced bend and a reference DNA bend localized on the same DNA fragment (*35*). We employed DNA fragments, which contained a protein binding sequence, a reference sequence of defined curvature, and a spacer region between them. An array of six phased A_6 tracts with 108° of bending toward the DNA major groove in the center (*36*) was used as the defined curvature, and the length of the spacer region was varied from 8 to 18 bp in increments of 2 bp. If DNA bending occurs at the protein binding site, the overall shape of the DNA fragment would be altered by the spacer length, in the range of conformational isomers between the in-phase and out-of-phase orientations of the reference and induced bend (*6*). Four kinds of fragments termed MIXatGC, MIXmixGC, GCatGC, or GCmixGC were used for the protein binding sequence. The DNA fragments contained a single GC box or two distal GC boxes separated by one turn of DNA helix (10 bp) with AT-rich (at) or randomized (mix) base composition (Figure 5A). Figure 5B summarizes the results

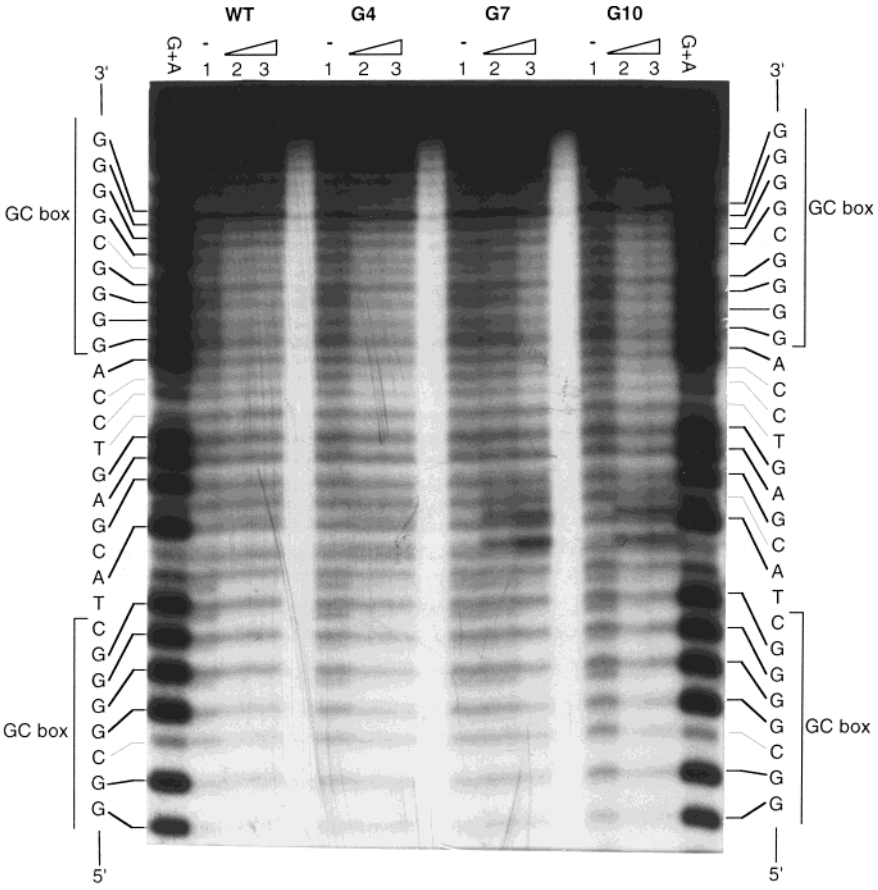

FIGURE 3: Hydroxyl radical footprinting assays of Sp1(530−623) and Sp1ZF6(Gly)*n* for two separated GC boxes. The panel shows the results for the G-strand. The samples (20 μL) contained the following concentrations of peptide: lanes 1, 2, and 3, 0, 25, and 100 nM, respectively. The first and last lanes indicate G+A of the Maxam−Gilbert sequencing reactions.

of phasing analysis. The relative electrophoretic mobility of a set of phase-sensitive DNA fragments was expressed as a cosine function of spacer length (*32*). The data were analyzed by converting each mobility to the corresponding relative mobility and fitting a cosine function (Figure 5C, see also eq 1). The amplitude and the phasing obtained from the fitted cosine curves reflect the overall bending magnitude and the bending direction of the protein binding site, respectively. At the employed concentration of each protein, both protein-bound and free DNAs, or both one and two Sp1(530−623)-bound DNAs, were detected at the same time. In all the cases, differences in the electrophoretic mobility dependent on the spacer length were observed, suggesting the existence of intrinsic and protein-induced DNA bendings. In MIXmixGC, sequence-dependent intrinsic bending was detected under protein-free conditions. A slight structural change in DNA arose due to the binding of Sp1(530−623), and all artificial 6-zinc finger proteins showed almost the same cosine curve as Sp1(530−623) (Figure 5C, left). The analogous tendency was also obtained when each protein bound to MIXatGC sequences (data not shown). These results clearly suggest that the bindings of these zinc finger proteins have remarkably similar effects on DNA structure and that the residual three fingers within the 6-zinc finger motifs of Sp1ZF6(Gly)-*n*, which do not contribute to DNA binding under this condition, showed little influence on DNA. On the other hand, the 6-zinc finger proteins presented different cosine curves in GCmixGC (Figure 5C, right). The phasing amplitude was approximately the same among Sp1ZF6(Gly)7, Sp1ZF6(Gly)10, and the two Sp1(530−623)s, although their phases differed considerably. This result indicates that the DNA bending induced by the binding of these proteins to the two distal GC boxes is different in direction and affects the entire DNA shape.

From the footprinting experiment, the bending center is predicted to locate at the 5′-TA-3′ step in the intervening region between the two GC boxes when Sp1ZF6(Gly)7 or Sp1ZF6(Gly)10 binds to DNA. The distance between the two bend centers, where the control bending of the poly AT tracts and the examined bending of the protein binding site are in-phase, can be determined from the obtained phase value (eqs 2 and 3). When Sp1ZF6(Gly)7 or Sp1ZF6(Gly)-10 bound to two separated GC boxes and formed an "in-phase" isomer, the number of helical turns between the two bending centers was 6.0. The length was almost the same, regardless of the sequence composition of the intervening region in GCmixGC or GCatGC. The number (6.0) of helical turns, namely an integral value, means that the examined bending is directed toward the DNA major groove similar to the bending direction of the control bending of the 6-poly AT tracts. This finding is appreciably in agreement with the result of the hydroxyl radical footprinting.

DISCUSSION

The results of the gel mobility shift assay reveal that the binding mode to the two distal GC boxes is evidently different between Sp1ZF6(Gly)4, and Sp1ZF6(Gly)7 [or

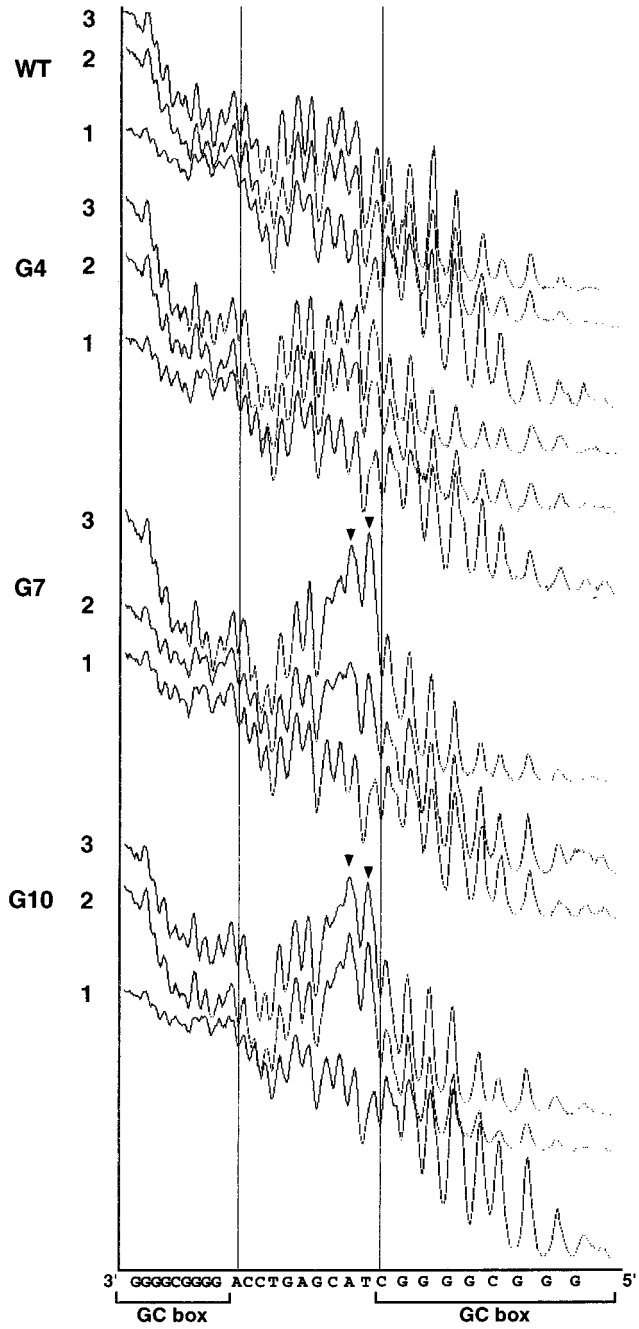

FIGURE 4: Line scans corresponding to tracks on the gel in Figure 3. The DNA sequence is shown at the bottom. Lanes 1, 2, and 3: 0, 25, and 100 nM of each peptide, respectively. Arrows indicate the hypersensitive cleavage nucleotides.

Sp1ZF6(Gly)10], from the standpoint of the induced DNA structural change. These 6-zinc finger proteins contain two GC box-binding units. Therefore, they can bind to two distal GC boxes in three binding modes: (1) one protein binding to two sites; (2) one protein binding to only one site; and (3) two proteins binding to two sites. Because the linker length of Sp1ZF6(Gly)4 is too short to bind stably to both separated GC boxes with all its zinc finger motifs, only one GC box is covered with the protein. The small DNA region covered with Sp1ZF6(Gly)4 may result in increased flexibility of the DNA fragment, in other words, faster electrophoretic mobility. On the other hand, Sp1ZF6(Gly)7 and Sp1ZF6(Gly)10 can bind to the two distal GC boxes with DNA bending which causes slower electrophoretic mobility.

Evidence for DNA bending induced by Sp1ZF6(Gly)7 and Sp1ZF6(Gly)10 was revealed from hydroxyl radical footprinting. The method provides a probe at the minor groove edge of the backbone at every bp regardless of sequence, and generally a hydroxyl radical cleaves the DNA backbone directly by attack at C(4′) in the deoxyribose ring (*37*, *38*). Therefore, the method has been applied to detect not only the protein-binding region but also the DNA conformational change (*39*, *40*). Indeed, Sp1ZF6(Gly)7 and Sp1ZF6(Gly)-10 showed a specific hypersensitive locus, suggesting that the enlarged minor groove at the 5′-TA-3′ site in the intervening region becomes more accessible to hydroxyl radical attacks due to the protein binding. Accordingly, DNA-bending narrowing the major groove is induced by binding of Sp1ZF6(Gly)7 or Sp1ZF6(Gly)10 to the two distal GC boxes. The observed shifts of the hypersensitive bands to a lower mobility indicate that another kind of cleavage product is formed at these positions. The position of a hydrogen atom abstracted by hydroxyl radicals at a target nucleotide has an effect on the modification of the 3′-end of the cleavage product, resulting in differences in the electrophoretic mobility (*41*). Binding of Sp1ZF6(Gly)7 or Sp1ZF6(Gly)10 to the two distal GC boxes surely causes DNA bending with an unusual sugar conformation of nucleotides at the sites of the intervening region. The distinction in the intensities of the hypersensitive bands between Sp1ZF6(Gly)7 and Sp1ZF6-(Gly)10 suggests that Sp1ZF6(Gly)7 induces larger DNA bending than Sp1ZF6(Gly)10.

The phasing analysis detected DNA bending with high sensitivity. As for the native 3-zinc finger Sp1(530−623), a slight DNA structural change was observed when Sp1(530−623) bound to a single GC box. Furthermore, the observed DNA structural alteration became clearer when two Sp1-(530−623) molecules bound to the two distal GC boxes. Several reports demonstrate that the particular region induces the DNA bending (*29*, *42*), unwinding (*43*), and local distortion (*44*). This fact supports the DNA structural change induced by the Sp1 zinc finger region. The DNA conformational alteration produced by each 6-zinc finger protein was compared with that by the binding of two 3-zinc finger Sp1(530−623)s. The DNA structural change caused by the binding of Sp1(530−623) to both of the two distal GC boxes was regarded as standard. A cosine curve was converted into a circular graph (Figure 6), where the coordinate (r, π) means (Aph, b') from the cosine curve [eq 1; b' is the magnitude of phase deviation compared with two Sp1(530−623) bindings]. Therefore, the distance from the center to each plot and the deviation from the x-axis reflect the overall bending magnitude and the relative entire bending direction, respectively. The position of a plot indicates the virtual position of the end of a DNA fragment through structural changes in the target DNA sequences. As for the phasing amplitude of the cosine curve, Sp1ZF6(Gly)10, Sp1ZF6-(Gly)7, and Sp1(530−623) exhibited almost the same values. However, the deviation values were evidently different among these 6-zinc finger proteins and Sp1(530−623), indicating that the linkage of the two 3-zinc finger motifs affects the entire DNA-bending direction. The magnitude of the deviation in Sp1ZF6(Gly)7 was more than that in Sp1ZF6(Gly)10. This tendency was independent of the intervening sequence composition between the two GC boxes, whereas Sp1ZF6(Gly)4 differed in both amplitude and

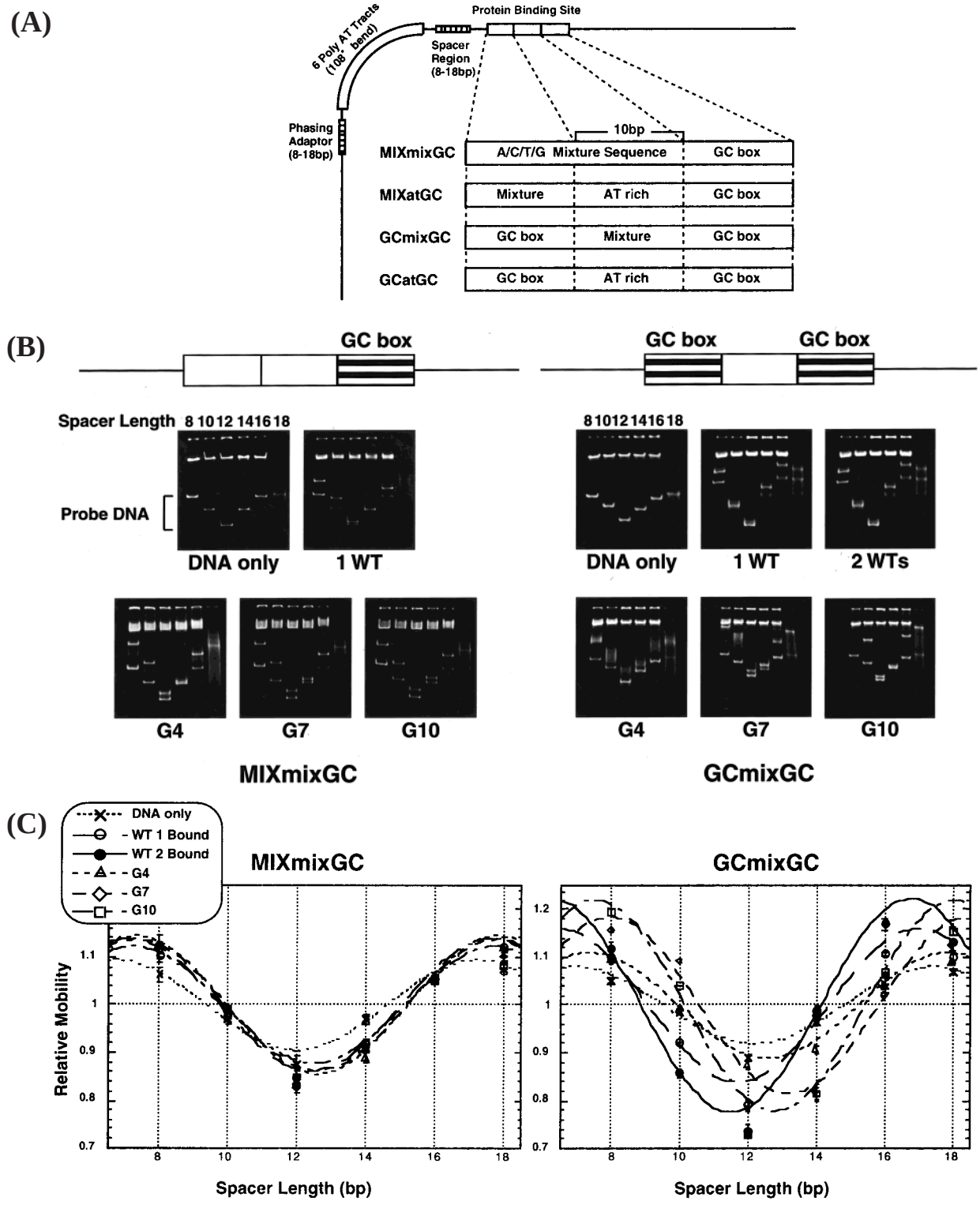

FIGURE 5: Sp1ZF6(Gly)*n*-induced bending analysis via phasing analysis. (A) Structure of the DNA fragment used for the phasing analysis and the sequences used for the protein binding site. A variable spacer region from 8 to 18 bp is placed between the protein binding sequence and the permanent bend resulting from six consecutive phased AT tracts. (B) The results of phasing analyses of wtSp1(530−623), Sp1ZF6(Gly)4, Sp1ZF6(Gly)7, and Sp1ZF6(Gly)10 for 'MIXmixGC' (left) and for 'GCmixGC' (right). (C) Plot of relative mobility versus spacer length. The protein binding sequences were 'MIXmixGC' (left) and 'GCmixGC' (right). The electrophoretic mobility of the phasing DNA fragments is plotted against the spacer length. Data were normalized to the average mobility over the entire spacer range.

phase from the case of all other proteins. Based on the results of footprinting, the binding of Sp1ZF6(Gly)4 to DNA seems to produce no dramatic effects on DNA structure. In the phasing analysis, a smearing was somewhat detected in the band of the Sp1ZF6(Gly)4−GCmixGC complex. Presumably, Sp1ZF6(Gly)4 unstably binds to the two distal GC boxes or has a DNA binding mode distinct from those of Sp1ZF6(Gly)7 and Sp1ZF6(Gly)10. If the DNA binding mode of the two independent Sp1(530−623) molecules is regarded as that of a tentative 6-zinc finger protein molecule with an extremely flexible and long linker, the difference in the DNA structural changes induced by the artificial 6-zinc

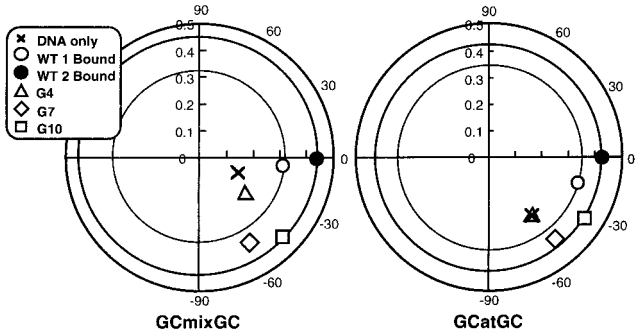

FIGURE 6: Relative end position of phasing DNA fragments containing 'GCmixGC' sequence in DNA bending. The complete binding of Sp1(530−623) is the standard position. The distance from the center to each plot and the deviation from the *x*-axis reflect the overall bending magnitude and the relative entire bending direction, respectively.

finger proteins, Sp1ZF6(Gly)*n*, can be attributed to their linker regions. Thus, our results indicate that Sp1ZF6(Gly)7 and Sp1ZF6(Gly)10 possess linkers of sufficient length to bind to both of the distal GC boxes. In addition, the linker length has an important effect on the entire DNA bending direction.

Previously, a 15-base-long TFO was demonstrated to cause DNA bending directed toward the DNA minor groove at the center of the intervening region between two TFO target sites (*5*−*7*). The present artificial 6-zinc finger proteins, Sp1ZF6-(Gly)7 and Sp1ZF6(Gly)10, bind to the DNA major groove like TFO but induce DNA bending directed toward the DNA major groove at the 4 bp site shifted to the 5′-side from the center of the intervening region between the two GC boxes. Although Sp1(530−623) binds to about 10 bp of the GC box sequence, the contribution of finger 1 to DNA binding is less than that of finger 2 or finger 3 (*45*). Even a finger 1 deletion mutant of Sp1 can bind to the GC box sequence. The 6-zinc finger proteins contain two GC box-binding units, but their binding affinities are not substantially the same. Accordingly, the asymmetrical binding results in the shifted bend center.

A designed 6-zinc finger protein in which two zinc finger units of Zif268 were connected with a slightly longer linker than a canonical type was demonstrated to have extremely high affinity for the target site (*23*). This finding suggests that the linker length between the two DNA-binding domains is associated with the binding affinity. Sp1ZF6(Gly)10 binds to the two distal GC boxes with a higher affinity than Sp1ZF6(Gly)7 (data not shown). Instead of disadvantage resulting from the shortage of linker length, Sp1ZF6(Gly)7 possibly induces a larger DNA conformational change. Therefore, the length of a peptide linker of the artificial 6-zinc finger proteins affects the DNA conformation, in particular the DNA bending direction. Although Sp1ZF6(Gly)4 seems to bind to only one GC box, the protein may bind momentarily to both of the two separated GC boxes with a dramatic DNA conformational change.

In conclusion, the overall bending directions induced by Sp1ZF6(Gly)7 and Sp1ZF6(Gly)10 are distinctly different. The direction of a local DNA distortion affects the total structure of a DNA−protein complex. It is known that some dimers of basic region−leucine zipper peptides bind to an AP-1 site and induce DNA bending. The direction and magnitude of their DNA bendings are dependent on the kind of dimer peptides (*32*, *46*). The difference in the direction of the local DNA bending is amplified, and the amplification affects the total structure of DNA. A topical DNA bending changes the situation of intrinsic DNA bending and protein binding modes. As a result, protein−DNA or protein−protein interaction would be transformed in the transcription initiation complex. Our results provide helpful information for the design of an optimized protein that controls gene expression by changing the DNA bending direction.

REFERENCES

1. Werner, M. H., Gronenborn, A. M., and Clore, G. M. (1996) *Science 271*, 778−784.
2. Parekh, B. S., and Hatfield, G. W. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 1173−1177.
3. Parvin, J. D., McCormick, R. J., Sharp, P. A., and Fisher, D. E. (1995) *Nature (London) 373*, 724−727.
4. Forget, D., Robert, F., Grondin, G., Burton, Z. F., Greenblatt, J., and Coulombe, B. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 7150−7155.
5. Akiyama, T., and Hogan, M. E. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 12122−12127.
6. Akiyama, T., and Hogan, M. E. (1997) *Biochemistry 36*, 2307−2315.
7. Liberles, D. A., and Dervan, P. B. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 9510−9514.
8. Strauss, J. K., and Maher, L. J., III (1994) *Science 266*, 1829−1834.
9. Strauss, J. K., Roberts, C., Nelson, M. G., Switzer, C., and Maher, L. J., III (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 9515−9520.
10. Strauss, J. K., Prakash, T. P., Roberts, C., Switzer, C., and Maher, L. J., III (1996) *Chem. Biol. 93*, 671−678.
11. Pavletich, N. P., and Pabo, C. O. (1991) *Science 252*, 809−817.
12. Pavletich, N. P., and Pabo, C. O. (1993) *Science 261*, 1701−1707.
13. Berg, J. M., and Shi, Y. (1996) *Science 271*, 1081−1085.
14. Elrod-Erickson, M., Rould, M. A., Nekludova, L., and Pabo, C. O. (1996) *Structure 4*, 1171−1180.
15. Elrod-Erickson, M., Benson, T. E., and Pabo, C. O. (1998) *Structure 6*, 451−464.
16. Scott, J. K., and Smith, G. P. (1990) *Science 249*, 386−390.
17. Reber, E. J., and Pabo, C. O. (1994) *Science 263*, 671−673.
18. Jamieson, A. C., Wang, H., and Kim, S. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 12834−12839.
19. Choo, Y., Castellanos, A., García-Hernández, B., Sánchez-García, I., and Klug, A. (1997) *J. Mol. Biol. 273*, 525−532.
20. Greisman, H. A., and Pabo, C. O. (1997) *Science 275*, 657−661.
21. Pomerantz, J. L., Sharp, P. A., and Pabo, C. O. (1995) *Science 267*, 93−96.
22. Pomerantz, J. L., Pabo, C. O., and Sharp, P. A. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 9752−9756.
23. Kim, J.-S., and Pabo, C. O. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 2812−2817.
24. Liu, Q., Segal, D. J., Ghiara, J. B., and Barbas, C. F., III (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 5525−5530.
25. Beerli, R. R., Segal, D. J., Dreier, B., and Barbas, C. F., III (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 14628−14633.
26. Kamiuchi, T., Abe, E., Imanishi, M., Kaji, T., Nagaoka, M., and Sugiura, Y. (1998) *Biochemistry 37*, 13827−13834.
27. Nagaoka, M., Hagihara, M., Kuwahara, J., and Sugiura, Y. (1994) *J. Am. Chem. Soc. 116*, 4085−4086.
28. Kim, Y.-G., Cha, J., and Chandrasegaran, S. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 1156−1160.
29. Nagaoka, M., and Sugiura, Y. (1996) *Biochemistry 35*, 8761−8768.
30. Kuwahara, J., and Coleman, J. E. (1990) *Biochemistry 29*, 8627−8631.
31. Tullius, T. D., and Dombroski, B. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5469−5473.

32. Kerppola, T. K., and Curran, T. (1991) *Cell 66*, 317−326.
33. Kadonaga, J. T., Carner, K. R., Masiarz, F. R., and Tjian, R. (1987) *Cell 51*, 1079−1090.
34. Zimm, B. H., and Levene, S. D. (1992) *Q. Rev. Biophys. 25*, 171−204.
35. Zinkel, S. S., and Crothers, D. M. (1987) *Nature (London) 328*, 178−181.
36. Crothers, D. M., Haran, T. E., and Nadeau, J. G. (1990) *J. Biol. Chem. 265*, 7093−7096.
37. Wu, J. C. Kozarich, J. W., and Stubbe, J. (1983) *J. Biol. Chem. 258*, 4694−4697.
38. Yang, J., and Carey, J. (1995) *Methods Enzymol. 259*, 452−468.
39. Wang, L., and Winans, S. C. (1995) *J. Mol. Biol. 253*, 32−38.
40. Mernagh, D. R., and Kneale, G. G. (1996) *Nucleic Acids Res. 24*, 4853−4858.
41. Dedon, P. C., Jiang, Z.-W., and Goldberg, I. H. (1992) *Biochemistry 31*, 1917−1927.
42. Sjøttem, E., Andersen, C., and Johansen, T. (1997) *J. Mol. Biol. 267*, 490−504.
43. Shi, Y., and Berg, J. M. (1996) *Biochemistry 35*, 3845−3848.
44. Kuwahara, J., Yonezawa, A., Futamura, M., and Sugiura, Y. (1993) *Biochemistry 32*, 5994−6001.
45. Yokono, M., Saegusa, N., Matsushita, K., and Sugiura, Y. (1998) *Biochemistry 37*, 6824−6832.
46. Paolella, D. N., Palmer, C. R., and Schepartz, A. (1994) *Science 264*, 1130−1133.

BI992989B