Unique DNA Binding Mode of the N-Terminal Zinc Finger of Transcription Factor Sp1[†]

Masanori Yokono, Nana Saegusa, Keizo Matsushita, and Yukio Sugiura*

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

Received November 10, 1997; Revised Manuscript Received March 5, 1998

ABSTRACT: Transcription factor Sp1 has three tandem repeats of a Cys_2His_2-type zinc finger motif and specifically binds to GC box DNA. Although the DNA binding mode of the three zinc fingers of Sp1 is predicted to be similar to that of Zif268, this model does not explain the DNA binding property of the N-terminal zinc finger (finger 1). To understand the DNA recognition mode of Sp1, we have performed detailed analyses for the contribution of finger 1 to the high-affinity binding to the GC box DNA and for the interaction mechanism between finger 1 and DNA. Results of electrophoretic analyses using finger-deleted mutants of Sp1 and GC box mutants in the finger-contacting subsite demonstrate that the contribution of finger 1 to the total DNA binding affinity is lower than that of the C-terminal finger 3 but is dispensable for the high-affinity binding. The DNA sequence selectivity of finger 1 at the 3′-portion of the GC box is lower than that of fingers 2 and 3 at the 5′-portion. Alanine scanning mutagenesis in the α-helix of finger 1 reveals that Lys−1 immediately preceding the helix is important for the recognition of the two guanine bases, but other putative key amino acids do not affect the DNA binding. These results demonstrate that (1) the contribution of finger 1 to the DNA binding affinity and the sequence selectivity of Sp1 is smaller than that of fingers 2 and 3 and (2) the interaction mechanism between finger 1 and DNA is different from the Zif268 model. DNA interaction of Sp1 finger 1 has also been discussed in connection with that of TFIIIA or WT1.

Human transcription factor Sp1 is a sequence-specific DNA binding protein isolated from HeLa (human) cells (*1*, *2*). Sp1 enhances transcription in a variety of viral and cellular genes by binding to GC-rich recognition elements (GC boxes) within the 5′-flanking promoter sequences (*3*, *4*). It has been proposed that the consensus sequence of Sp1 binding is the decanucleotide 5′-(G/T)GGGCGG(G/A)(G/A)(C/T)-3′ (*3*, *4*). Sp1 contains three contiguous repeats of a typical Cys_2His_2-type zinc finger motif as a DNA binding domain (*2*, *5*, *6*), and the peptide containing only this domain can bind to GC box DNA with almost the same affinity as full-length Sp1 (*7*, *8*). The crystal structure of the DNA binding domain of this family, Zif268, bound to its target DNA site, has been determined (*9*, *10*). In the complex, the three zinc fingers occupy the major groove of the DNA in series, each making base-specific contacts to overlapping 4 bp subsites (*10*). The protein interacts with both strands of the DNA, but the majority of the contacts is with the guanine-rich strand. Sequence specificity arises from 1:1 interactions between residues in four positions of each finger α-helix (positions −1, 2, 3, and 6) and the DNA bases, and also, the ability of sequence discrimination is similar among the three fingers (*11*). On the basis of the similarity of amino acid sequences and their target DNA sequences between Sp1 and Zif268, the model of three-zinc finger Sp1−DNA interaction has been proposed (*8*, *12*, *13*).

However, some previous studies of the Sp1−DNA interaction showed the different contribution of the three fingers of Sp1 to the GC box DNA binding. The mapping of the binding site for Sp1 zinc fingers using primer extension gel mobility-shift assay and chemical interference analyses revealed that the putative DNA subsites for fingers 2 and 3 at the 5′-portion of the GC box contribute more strongly to the total binding affinity than the subsite for finger 1 (*8*, *13*). Cis element analyses in Sp1-responsive promoters and the in vitro Sp1 binding site selection assay also demonstrated that the DNA sequence is highly conserved at the subsites of fingers 2 and 3, whereas there is a high degree of sequence diversity at the finger 1 DNA subsite (*3*, *4*, *14*, *15*). These findings indicate a binding nature of finger 1 different from that of the Zif268 model. However, detailed study of the DNA binding of finger 1 has never been performed.

To understand the DNA recognition mode of Sp1, we have investigated the contribution of finger 1 to the high-affinity binding to the GC box DNA with three approaches: finger deletion from the three-zinc finger peptide, introduction of mutations throughout the consensus GC box DNA, and alanine substitutions at critical residues in the recognition helix. These results demonstrate that the contribution of finger 1 to the DNA binding affinity and the sequence selectivity of Sp1 is smaller than that of fingers 2 and 3 and that the interaction mechanism between finger 1 and DNA is different from the Zif268 model.

[†] This study was supported in part by a Grant-in-Aid for Priority Project "Biometals" (08249103) and Scientific Research (B) (09557200) from the Ministry of Education, Science, Sport, and Culture, Japan.

* To whom correspondence should be addressed.

S0006-2960(97)02764-5 CCC: $15.00
Published on Web 04/18/1998

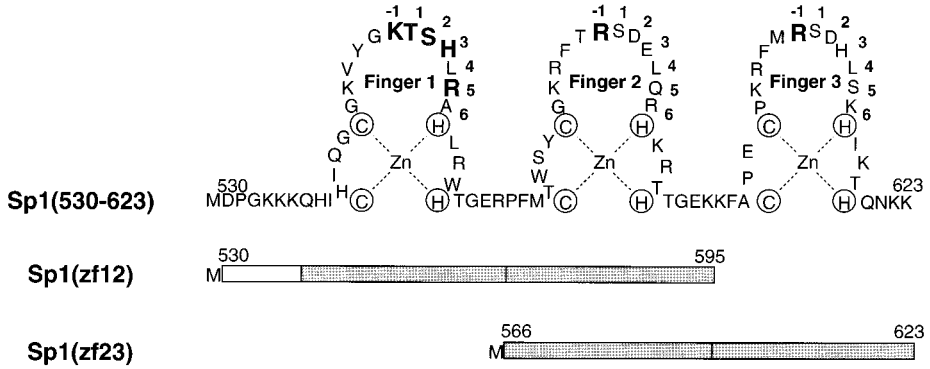


FIGURE 1: Amino acid sequences of various peptide mutants of Sp1. In Sp1(530−623), amino acids are written in one-letter codes. Invariant ligands, cysteines and histidines, are circled. The numbers on the residues are the numbers of the amino acid sequence in native Sp1. Coding regions in finger-deleted mutants are shown by solid bars. The residues substituted with alanine are bold in Sp1(530−623), and the numbers of the helical positions in each finger are also bold. The nomenclature is described in the text.

MATERIALS AND METHODS

Materials. All enzymes were purchased from New England Biolabs (Beverly, MA), except for restriction enzyme *Sal*I which was obtained from Takara Shuzo (Kyoto, Japan). [γ-^{32}P]ATP was supplied by DuPont, and dimethyl sulfate was obtained from Aldrich (Milwaukee, WI). All other chemicals were of commercial reagent grade.

Preparation of Zinc Finger Peptides from Sp1. The plasmid pBS-Sp1-fl was kindly provided by R. Tjian. The pUCSp1(530−623), which encodes the three-zinc finger region, was constructed as described previously (*16*). All mutant constructs were generated from pUCSp1(530−623), and site-directed mutagenesis was carried out according to the method of Kunkel et al. (*17*). The finger 3-deleted mutant, Sp1(zf12), was prepared by making a stop codon at lysine 596 on pUCSp1(530−623). The finger 1-deleted mutant, Sp1(zf23), was produced by changing arginine 565 to methionine with new *Nde*I site. The single point mutants were made by changing the target amino acid residues into the desired amino acid residues and named according to the following rules; the lysine to alanine mutant at position 550 was named Sp1(K550A). All sequences of mutated regions were confirmed by DNA sequence analysis using the BcaBEST dideoxy sequencing kit (Takara Shuzo). Figure 1 summarizes amino acid sequences of all mutants. The *Bam*HI−*Eco*RI fragments of Sp1 derivatives were cut off from pUCSp1 derivatives and inserted into the similarly digested plasmid pEV-3b (*16*), to construct the pEVSp1 derivatives. These zinc finger peptides were overexpressed in *Escherichia coli* strain BL21(DE3)pLysS as described previously (*18*). Purification procedures were performed at 4 °C. *E. coli* cells in which Sp1 peptides were overexpressed were resuspended and lysed in phosphate buffer [130 mM NaCl, 2.7 mM KCl, and 10 mM phosphate buffer (pH 7.0)]. After centrifugation, the supernatant containing the soluble form of Sp1 peptides was purified by cation-exchange chromatography (HighS, Bio-Rad; MonoS, Pharmacia) with phosphate buffer (pH 8.0 or 7.0) using a NaCl gradient (0.13 to 1 M). The isolated peptides were dialyzed with peptide storage buffer [10 mM Tris-HCl (pH 7.5), 50 mM NaCl, and 1 mM β-mercaptoethanol] and stored at 4 °C. The procedure yielded 1 mg of Sp1 peptides/g wet weight of cells.

Gel Mobility-Shift Assay. The 17 base pair (bp) single GC box oligonucleotide and mutated GC box oligonucleotides (Figure 2) were synthesized on an Applied Biosystems 391 DNA synthesizer. These oligonucleotides were annealed and inserted between a *Bam*HI site and an *Eco*RI site of pBluescript II KS+ (Toyobo, Osaka, Japan). These plasmids were renamed as the pBS-GC series. The *Hin*dIII−*Xba*I fragment (41 bp) was cut out from these plasmids, ^{32}P-labeled at the 5′-end with T4 polynucleotide kinase, and gel-purified. Binding reaction mixtures (final volume, 20 μL) contained 10 mM Tris-HCl (pH 8.0), 50 mM NaCl, 100 μM $ZnCl_2$, 1 mM β-mercaptoethanol, 0.05% Nonidet P-40, 5% glycerol, 0−5 μM Sp1 peptides, 25 ng/μL poly(dI-dC) (Pharmacia), and the 5′-end-labeled DNA fragment (~50 pM, 500 cpm). After incubation at 20 °C for 30 min, the sample solutions were electrophoresed on 12% nondenaturing polyacrylamide gels with Tris-borate buffer [88 mM Tris-HCl (pH 8.0) and 88 mM boric acid] at 20 °C. The bands were visualized by autoradiography and quantitated with NIH Image (Version 1.58). The dissociation constants (K_d) of the Sp1 peptide−DNA fragment complexes were estimated as follows. The fractions of labeled DNA bound to Sp1 peptides were calculated using the equation $\theta_b = I_b/(I_b + I_f)$, where I_b and I_f are the intensities of the peptide-bound DNA band and the free DNA band, respectively. The K_d was evaluated by fitting the experimentally obtained values of θ_b to the binding isotherm equation θ_b = [peptide]/([peptide] + K_d) (Kaleida Graph program, Abelbeck Software).

Oligonucleotide Sequence

	1 2 3 4 5 6 7 8 9 10 11
GC	TCT GGGGCGGGGCC TAA
F1-AT1	TCT GGGGCG**TAATT** TAA
F1-AT2	TCT GGGGCG**ATTAA** TAA
F1-AT3	TCT GGGGCGG**AATT** TAA
F1-AT4	TCT GGGGCGG**TTAA** TAA
F2-AT	TCT GGG**ATA**GGGCC TAA
F3-AT	TCT **TAA**GCGGGGCC TAA

FIGURE 2: GC box and mutated GC box sequences used in this work. Substituted nucleotides are bold. The base numbers in the GC box are also shown. Nomenclatures F1, F2, and F3 refer to sequences mutated in the putative DNA binding subsites of each finger. AT represents an AT-rich sequence.

DNase I Footprinting Analysis. DNase I footprinting was carried out according to the method of Brenowitz et al. (*19*). The plasmid pCPA 5 (*16*) was digested with *Xba*I and *Sal*I to generate a 148 bp fragment containing a single GC box,

Table 1: Dissociation Constants (K_d) for Dissociation of Sp1(530−623) and Sp1(zf23) for the GC Box and Mutated GC Box Oligonucleotides

binding[b] site	K_d (nM)[a] 530−623	zf23
GC	3.5 ± 0.5	310 ± 20
F1-AT1	37 ± 7	440 ± 120
F1-AT2	ND[c]	ND
F1-AT3	72 ± 22	1200 ± 100
F1-AT4	52 ± 12	2000 ± 100
F2-AT	200 ± 40	ND
F3-AT	3100 ± 900	ND

[a] Apparent dissociation constants are determined by titration using a gel mobility-shift assay as described in Materials and Methods. Values are averages of three or more independent determinations with standard deviations. [b] The nomenclature is described in the text (see Figures 1 and 2). [c] ND, not determined.

and the fragment was 5′-end-labeled as described above and gel-purified. Each reaction mixture (final volume, 20 μL) involved 26 mM Tris-HCl (pH 8.0), 30 mM NaCl, 5 mM $CaCl_2$, 10 mM $MgCl_2$, 0.6 mM β-mercaptoethanol, 0−5 μM Sp1 mutants, 1 μg of sonicated calf thymus DNA, and the 5′-end-labeled DNA fragment (∼4 nM, 20 000 cpm). After incubation at 20 °C for 30 min, the sample was digested with DNase I (final concentration, 0.75 unit/mL) at 20 °C for 2 min. After the digestion reaction, 20 μL of DNase I stop solution (0.1 M EDTA and 0.6 M sodium acetate) and 100 μL of ethanol were added to the sample solution. The cleavage products were analyzed on an 8% polyacrylamide/7 M urea sequencing gel. The bands were visualized by autoradiography.

Methylation Interference Analysis. Methylation interference experiments were performed as described previously (*13*, *16*), except for the use of reaction buffer containing 10 mM Tris-HCl (pH 8.0), 50 mM NaCl, 100 μM $ZnCl_2$, 1 mM β-mercaptoethanol, 0.05% Nonidet P-40, 5% glycerol, 50−2000 nM Sp1 peptides, 1 μg of sonicated calf thymus DNA, and the 5′-end-labeled *Hin*dIII−*Xba*I fragment of pBS-GC (∼40 nM, 3×10^5 cpm). Densitometric measurements were obtained with NIH Image (Version 1.58). On the basis of the obtained cutting probabilities *F* and *B*, where *F* represents the intensity of the typical band in the free lane and *B* in the bound lane, the extent of interference was estimated for each base as the ratio *B*/*F*. Band intensities between free and bound lanes were corrected by comparing nucleotides outside the binding site.

RESULTS

Effect of Finger 1 Deletion on GC Box DNA Binding. Two-finger mutant peptides Sp1(zf12) and Sp1(zf23) were generated from three-zinc finger peptide Sp1(530−623) (Figure 1). Sp1(zf12) encodes amino acid residues 530−595 and contains two N-terminal fingers (fingers 1 and 2). Sp1(zf23) encodes residues 566−623 and includes two C-terminal fingers (fingers 2 and 3). The properties of binding of these peptides to the consensus GC box sequence were characterized by gel mobility-shift assay, DNase I footprinting, and methylation interference analyses. The GC box sequence (Figure 2) was derived from mouse dihydrofolate reductase promoter (I and III) (*3*, *4*). Table 1 showed the dissociation constants (K_d) for dissociation of Sp1(530−623) and Sp1(zf23) from the GC box (GC). The dissociation constants for dissociation of Sp1(530−623) and Sp1(zf23) from GC were 3.5 and 310 nM, respectively. In contrast, Sp1(zf12) did not bind the GC box under this experimental condition. The deletion of finger 1 causes 89-fold reduction in DNA binding affinity, and the deletion of finger 3 reduces the affinity much more than that of finger 1. Figure 3 displays the DNase I footprinting patterns of these three peptides for a 148 bp DNA fragment containing the GC box. In both the guanine-rich strand (G-strand) and the cytosine-rich strand (C-strand), Sp1(530−623) protected all the residues of the GC box at a low peptide concentration (∼1 μM) (lanes 2−5 in both strands). Sp1(zf23) also protected the residues at the 5′-portion of the GC box, but a higher concentration of the peptide (>5 μM) was required for full protection as compared with that of Sp1(530−623) (lanes 10−13 in both strands). The binding of Sp1(zf12) was not detected (lanes 6−9 in both strands). Figure 4 presents methylation interference patterns of Sp1(530−623) and Sp1(zf23) for a 41 bp DNA fragment containing the GC box. The extent of interference is shown by a histogram (Figure 4B). As described previously (*16*), the methylations of eight guanines in the G-strand [G(1), G(2), G(3), G(4), G(6), G(7), G(8), and G(9)] and of three guanines in the C-strand [G(5′), G(10′), and G(11′)] interfered with the GC box binding of Sp1(530−623) (A, lanes 3, 4, 7, and 8), although the extent of interference was different for each guanine base. Namely, strong interference was observed at G(2), G(3), G(4), and G(6) in the G-strand and at G(5′) in the C-strand, while weak interference was detected at G(1), G(7), G(8), and G(9) in the G-strand and at G(10′) and G(11′) in the C-strand. In the case of Sp1(zf23), methylation interference was observed only at six guanines in the G-strand [G(1), G(2), G(3), G(4), and G(6)] and C-strand [G(5′)] (A, lanes 5, 6, 9, and 10) which correspond to the putative DNA binding subsites of fingers 2 and 3 (*8*, *12*, *13*). These results reveal that the peptide containing only fingers 2 and 3 can specifically bind to the GC box DNA and that finger 1 interacts with the 5 bp DNA subsite at the 3′-portion of the GC box, namely, G(7), G(8), and G(9) in the G-strand and G(10′) and G(11′) in the C-strand.

Evaluation of the Sequence-Specific DNA Binding of Finger 1. To evaluate the DNA sequence selectivity of finger 1, six mutated GC box sequences were prepared (Figure 2). The design of these DNA mutants is based on the previous Sp1−DNA interaction model (*8*, *12*, *13*) and on our results of methylation interference analyses. According to the DNA base contacts of Zif268 (*10*, *20*), the base pair at position 7 would be overlapped between the binding subsites of fingers 1 and 2. We thus prepared two sets of mutants. F1-AT1 and -2 are the DNAs with 5 bp mutated at the putative DNA subsite for finger 1 (positions 7−11). F1-AT3 and -4 are the DNAs with 4 bp mutated at positions 8−11 with the G(7) base being conserved. F2-AT and F3-AT are the oligonucleotides mutated at the 3 bp finger 2 subsite (positions 4−6) and finger 3 subsite (positions 1−3), respectively. The affinities of binding of Sp1(530−623) and Sp1(zf23) to these mutated GC boxes were analyzed by a gel mobility-shift assay. Table 1 summarizes the dissociation constants (K_d) of these peptide−DNA complexes. Sp1(530−623) bound to the three finger 1 subsite mutants, F1-AT1, F1-AT3, and F1-AT4, with 11−

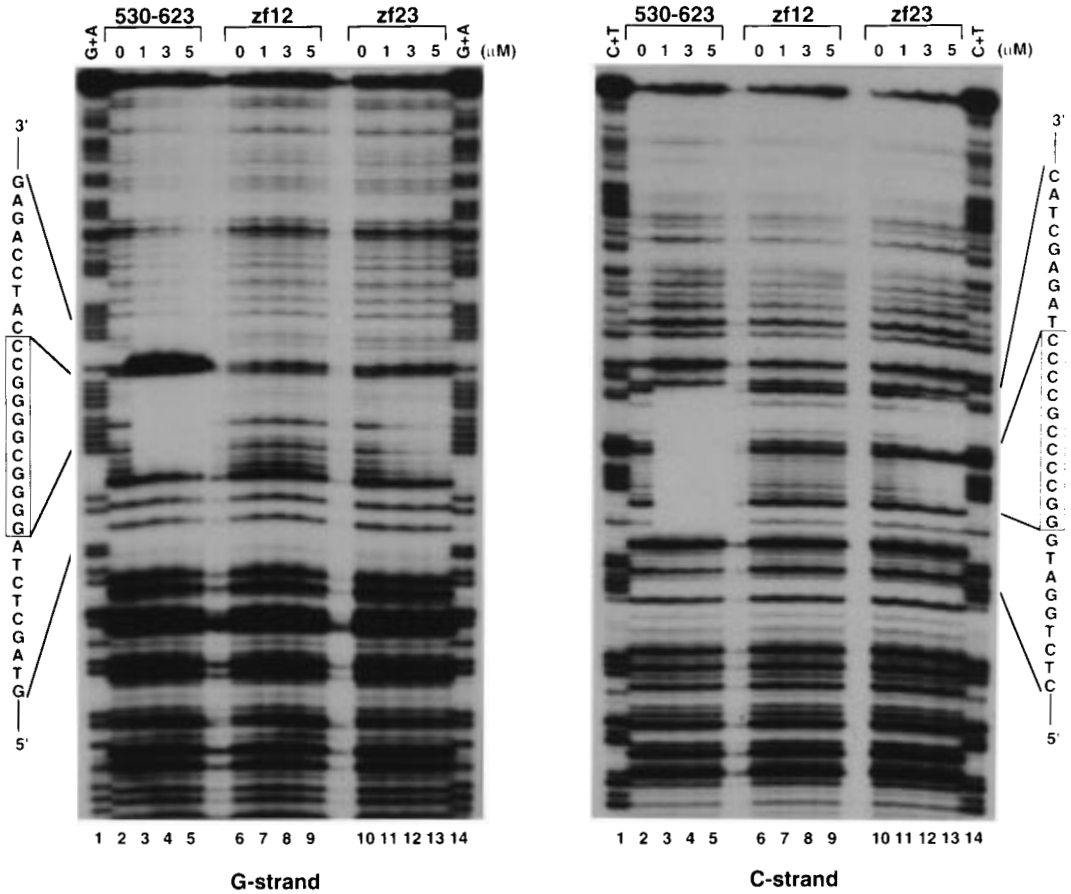
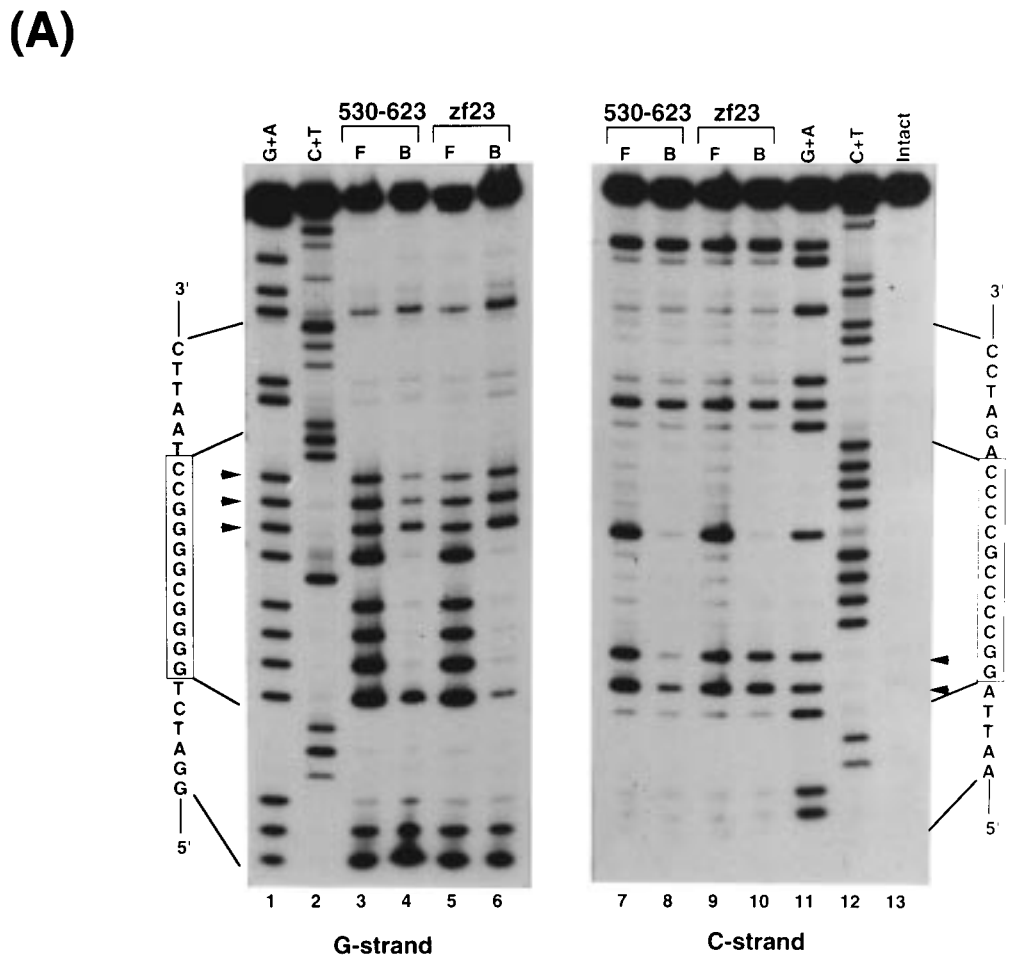


FIGURE 3: DNase I footprinting analyses of binding of Sp1(530−623), Sp1(zf12), and Sp1(zf23) to the GC box DNA. The left and right panels show the results for the G- and C-strands, respectively. The samples (20 μL) contained peptides at the following concentrations: 0 μM (lanes 2, 6, and 10), 1 μM (lanes 3, 7, and 11), 3 μM (lanes 4, 8, and 12), and 5 μM (lanes 5, 9, and 13). Lanes 1 and 14 in the G- and C-strands indicate G and A, and C and T, of the Maxam−Gilbert sequencing reactions, respectively.

21-fold lower affinities than GC, while it bound to F2-AT and F3-AT with 57- and 890-fold lower affinities, respectively. The affinity of Sp1(530−623) to F1-AT2 was much lower than that to the other finger 1 subsite mutants. Sp1(zf23) bound to F1-AT1, F1-AT3, and F1-AT4 with 1.4- to 6.5-fold lower affinities than GC. However, its bindings to F1-AT2, F2-AT, and F3-AT were not detected under this condition. Accordingly, DNA binding of finger 1 is more tolerant to the base pair substitutions than that of finger 2 or 3.

Essential Amino Acid Residues for Sequence-Specific Recognition by Finger 1. The DNA binding mode of three-zinc finger Sp1 has been modeled on the crystallographic data of the Zif268−DNA complex (*8*, *12*, *13*). In this model, key amino acid residues at the N terminus of the α-helices take part in the DNA base recognition by making contact between their side chains and DNA base edges. Indeed, previous mutagenic experiments of Sp1 indicated that the key amino acid residues in finger 2 are important in the GC box DNA binding (*7*, *21*−*23*). However, the amino acid−DNA base interaction for finger 1 proposed previously has never been examined. To obtain precise information on the interaction between amino acid side chains in finger 1 and DNA base edges, amino acid residues at the N terminus of the α-helix in finger 1 (Lys−1, Thr1, Ser2, His3, and Arg5; the numbers with the residues illustrate the positions in the α-helices of zinc fingers) were changed to Ala by site-directed mutagenesis (*24*) (Figure 1). These amino acids correspond to helical positions that make contacts with DNA base edges in four C_2H_2-type zinc finger proteins, Zif268 (*9*, *10*), GLI (*25*), Tramtrack (*26*), and YY1 (*27*). Ala6 in finger 1 was not mutated because of its inability to recognize guanine bases (*12*). As a control, the key amino acid residues, Arg−1 in finger 2 and Arg−1 in finger 3, were also changed to Ala. DNA binding properties of these peptides were assessed by gel mobility-shift assay and methylation interference analysis. Table 2 shows the dissociation constants (K_d) of these point-mutated peptides for GC and GC box mutants. Finger 2 mutant Sp1(R580A) and finger 3 mutant Sp1(R608A) bound to GC with 160- and 1100-fold lower affinities than Sp1(530−623). These results indicate that each Arg−1 in fingers 2 and 3 corresponding to key residues in the Zif268 binding model contributes greatly to the GC box DNA binding of Sp1. Finger 1 mutant Sp1(K550A) bound to GC with a 6.9-fold lower affinity than Sp1(530−623), while four other finger 1 mutants [Sp1(T551A), Sp1(S552A), Sp1(H553A), and Sp1(R555A)] bound with similar affinities to Sp1(530−623) (0.77−1.3-fold lower affinity). The affinities of all five finger 1 mutants for F1-AT DNAs were almost the same as that of Sp1(530−623). These results indicate that Lys−1 contributes to the GC box DNA binding affinity and sequence selectivity for the finger 1 subsite. Figure 5 displays methylation interference patterns of Sp1(K550A) and Sp1(H553A) in comparison with that of Sp1(530−623). The extent of the interference is shown by a histogram (Figure 5B). In the case of



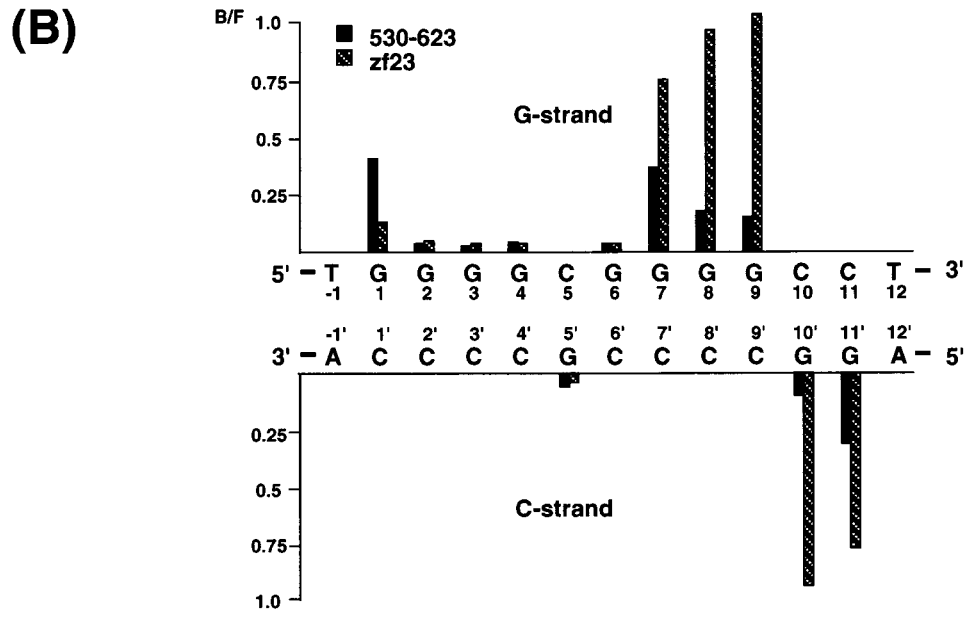


FIGURE 4: Methylation interference analyses of binding of Sp1(530−623) and Sp1(zf23) to the GC box DNA. (A) The left (lanes 1−6) and right (lanes 7−13) panels show the results for the G- and C-strands, respectively. Lanes 3−10 represent free (F) and peptide-bound (B) DNA samples. Lanes 1 and 11 contain G and A of the Maxam−Gilbert sequencing reaction and lanes 2 and 12 C and T; lane 13 contains intact DNA. Arrowheads in the G- and C-strands indicate the guanine bases that the extent of interference is different between Sp1(530−623) and Sp1(zf23). (B) A histogram showing the extent of methylation interference by Sp1(530−623) and Sp1(zf23). An autoradiogram of the gel was scanned with a densitometer, and the extent of interference was calculated as the ratio of the cutting probabilities for the two bands (*B*/*F*).

Table 2: Dissociation Constants (K_d) of Alanine-Substituted Sp1 Mutants for the GC Box and Mutated GC Box Oligonucleotides

	K_d (nM)[a]							
		finger 1					finger 2	finger 3
binding[b] site	530−623	K550A	T551A	S552A	H553A	R555A	R580A	R608A
GC	3.5 ± 0.5	24 ± 1	2.7 ± 0.7	4.3 ± 0.2	4.4 ± 0.8	4.0 ± 0.5	560 ± 10	3800 ± 500
F1-AT1	37 ± 7	47 ± 10	42 ± 3	51 ± 11	59 ± 5	55 ± 1	−	−
F1-AT4	52 ± 12	57 ± 9	47 ± 1	60 ± 15	72 ± 4	83 ± 7	−	−

[a] Apparent dissociation constants are determined by titration using a gel mobility-shift assay as described in Materials and Methods. Values are averages of three or more independent determinations with standard deviations. [b] The nomenclature is described in the text (see Figures 1 and 2).

Sp1(K550A), the methylation interference of G(8) and G(9) sites was significantly reduced. Sp1(H553A) as well as Sp1(T551A), Sp1(S552A), and Sp1(R555A) gave the same interference patterns as Sp1(530−623) in both of the strands (data not shown). These results suggest that Lys−1 at the N terminus of the α-helix in finger 1 plays a dominant role in recognition of G(8) and G(9) of the G-strand in the finger 1 DNA subsite.

DISCUSSION

Different Contributions of Individual Zinc Fingers to the DNA Binding of Sp1. These results reveal that all three of the zinc fingers of Sp1 participate in the high-affinity binding to the GC box DNA. However, the contributions of individual fingers to the DNA binding affinity and the sequence selectivity of Sp1 seem to be different. In the experiments using finger-deleted mutants, the finger 1-deleted mutant Sp1(zf23) binds specifically to the GC box DNA (GC) and the finger 3-deleted mutant Sp1(zf12) does not bind under the same conditions, indicating a smaller contribution from finger 1 than from finger 3 to the binding affinity. The comparison of K_d values for dissociation of Sp1(530−623) from the GC box and GC box mutants reveals that the affinity decreases in the order of finger 3 subsite > finger 2 subsite > finger 1 subsite DNA mutants and that finger 1 has a more relaxed sequence selectivity than fingers 2 and 3. From the experimental results using alanine mutants at the possible key residues for the DNA recognition, Lys−1 in finger 1 contributes to the GC box DNA binding but other residues at the N terminus of the recognition helix in finger 1 do not affect the binding. The binding affinities of three mutants at helical position −1 of each finger suggest that the contributions of these proposed critical residues to the total binding affinity and selectivity are reduced in the order of Arg−1 in finger 3 > Arg−1 in finger 2 > Lys−1 in finger 1. Therefore, the specific interaction between finger 1 and its DNA subsite appears to be less important than the DNA interactions of fingers 2 and 3, but is indispensable for the high-affinity binding to the GC box sequences. The finding is consistent with the previous biochemical results (*8*, *13*−*15*).

The reduction in the affinity of binding of Sp1(zf23) to GC (310 nM) is greater than those of Sp1(530−623) binding to F1-AT1, -3, and -4 (37, 72, and 52 nM, respectively). DNA substitutions possibly preserve the nonspecific interaction between finger 1 and DNA. Finger deletion would cause not only the loss of nonspecific interaction but also changes in the environment around its adjacent finger 2, leading to destabilization of the finger 2−DNA interaction. These possibilities explain that the decrease in affinity by finger deletion is larger than that by DNA substitutions. Interestingly, the affinity of Sp1(zf23) is considerably affected by the nucleotide composition at the noncontacting finger 1 DNA subsite (1.4−6.5-fold reduction in affinity). Presumably, nucleotide substitutions result in subtle changes in the DNA conformation. Several reports demonstrate that the DNA binding of the Sp1 zinc finger region induces DNA bending (*16*, *28*), unwinding (*29*), and local distortion (*13*) at the GC box region. Thus, the base substitutions within the finger 1 subsite may disturb the appropriate DNA conformational change for Sp1 binding which leads to a reduction in affinities of Sp1(zf23). On the other hand, the affinities of Sp1(530−623) are similar among F1-AT1, F1-AT3, and F1-AT4. It is likely that the nonspecific interaction between finger 1 and its DNA subsite can overcome the loss of binding energy derived from the unfavorable DNA structure.

The key amino acid residues in the α-helices and three-dimensional structure of Sp1 fingers 2 and 3 are similar to those of fingers 1 and 2 of Zif268 (*12*, *30*). Because each finger of Zif268 seems to equally contribute to the DNA binding (*11*), the contribution to the specific DNA binding is presumed to be similar for fingers 2 and 3 of Sp1. However, our results clearly indicate that finger 3 is more important than finger 2 for the DNA binding of Sp1. Therefore, it is difficult to predict the DNA binding property of each zinc finger only from the homology of amino acid sequences.

DNA Binding Mode of Sp1. From the similarity between fingers 2 and 3 of Sp1 and fingers 1 and 2 of Zif268 (*12*, *30*), the DNA interaction mode of Sp1 fingers 2 and 3 is deduced (Figure 6). In this recognition mode, there are seven amino acid−base contacts: in finger 2, Arg−1−G(6), Asp2−C(7′), Arg6−G(4); and in finger 3, Arg−1−G(3), Asp2−C(4′), His3−G(2), and Lys6−G(1). Thr6 in Zif268 finger 2 does not make contacts with any DNA base. But it is possible that Lys6 in Sp1 finger 3 can make contact with G(1) because the Lys6−guanine contacts are seen in other zinc finger protein−DNA complexes (*25*, *27*). Our results confirmed that Arg−1 residues in fingers 2 and 3 take part in DNA binding (*7*, *21*−*23*). Furthermore, Sp1(530−623) does not bind to F1-AT2 containing T(7′) in the C-strand but binds to other F1-ATs containing C(7′) or A(7′). This observation is consistent with the finding that Asp2 in each helix of Zif268 could recognize A or C in the C-rich strand at the 5′-portion of its N-terminal finger subsite (*20*). The great reduction in affinity for F1-AT2 is possibly attributed not only to the loss of contact but also to the DNA conformational change by the base substitutions.

As for the DNA interaction mode of finger 1, these results suggest that finger 1 covers 5 bp DNA, 5′-GGGCC-3′, and Lys−1 in finger 1 contacts the two bases, G(8) and G(9), in

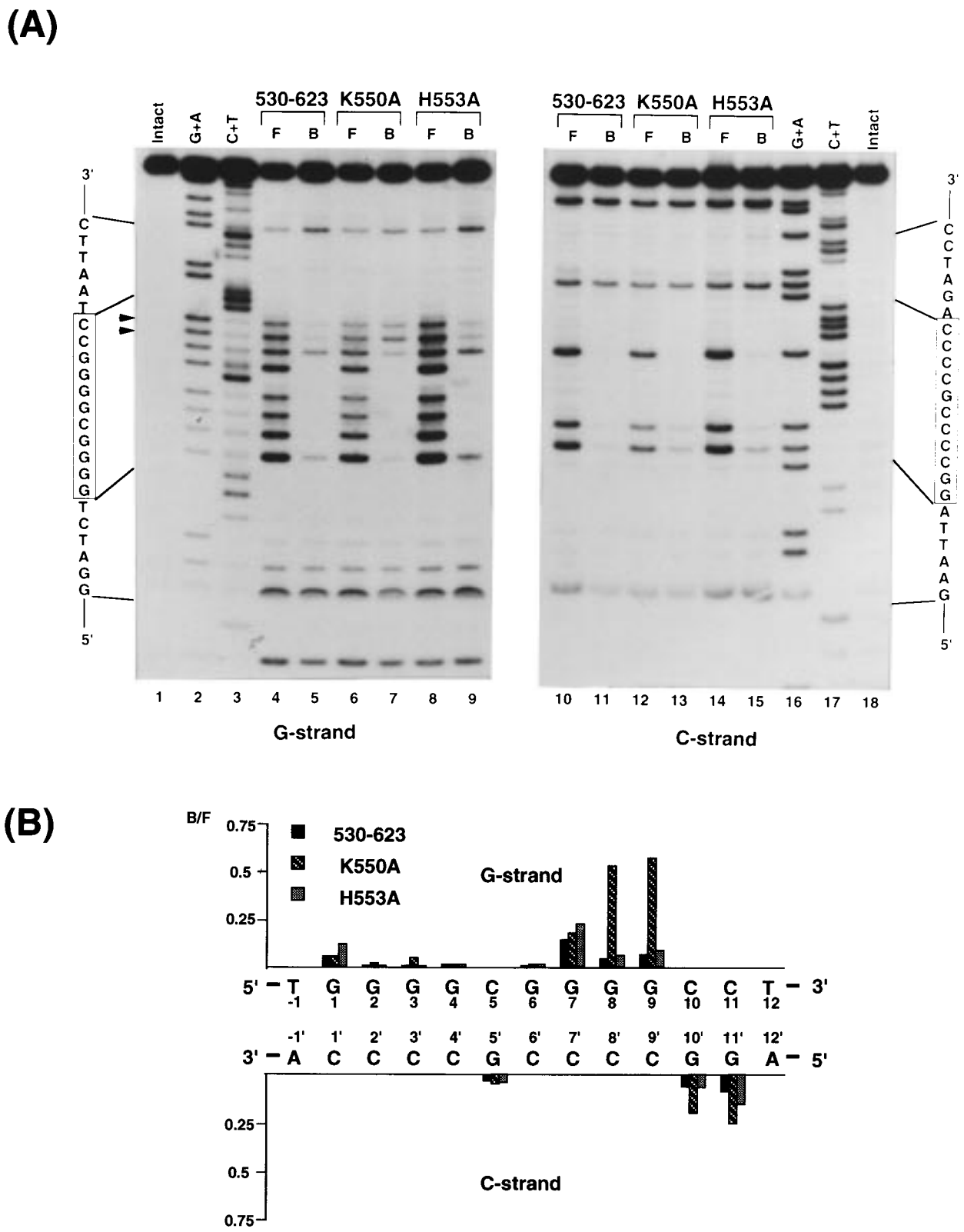


FIGURE 5: Methylation interference analyses of Sp1(530−623), Sp1(K550A), and Sp1(H553A). (A) The left (lanes 1−9) and right (lanes 10−18) panels show the results for the G- and C-strands, respectively. Lanes 4−15 represent free (F) and peptide-bound (B) DNA samples. Lanes 2 and 16 contain G and A of the Maxam−Gilbert sequencing reaction, lanes 3 and 17 C and T, and lanes 1 and 18 intact DNA. Arrowheads in the G-strand indicate that the extent of interference is different among these peptides. (B) A histogram showing the extent of methylation interference by Sp1(530−623), Sp1(K550A), and Sp1(H553A). An autoradiogram of the gel was scanned with a densitometer, and the extent of interference was calculated as the ratio of the cutting probabilities for the two bands (*B*/*F*).

the G-strand. The DNA binding is unaffected by the substitution of His3 which can make contact with G(8) in the Zif268 model. In addition, the base contacts with Ser2 and Arg5 seen in other zinc finger−DNA complexes (*25*−*27*) are not detected in our study. From these findings, the base recognition mode of finger 1 is predicted (Figure 6). The relaxed sequence selectivity at positions 8 and 9 may arise from the long and flexible nature of the Lys−1 side

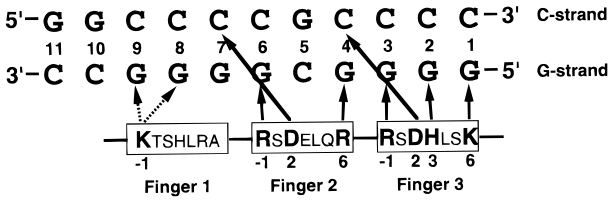


FIGURE 6: Mode of putative interaction of Sp1 with GC box DNA. Amino acid residues at the N terminus of the α-helix in each finger are depicted by their one-letter codes, and the putative residues participating in base recognition are bold with the numbers of the helical positions. Solid arrows depict the amino acid−base contacts assumed by the DNA binding mode of Zif268 (*9*, *10*), and dotted arrows depict the contacts indicated by these results.

chain. The base pair at position 7 could not be distinguished by the residues of finger 1 but could be by Asp2 of finger 2. The base pairs at position 10 and 11 are important for the DNA binding of Sp1 because the base substitution of CC to AA at these positions causes a 3.3-fold reduction in the binding affinity of Sp1 (M. Yokono et al., unpublished results), and the cis element analyses of GC box sequences showed the moderate base preference for C or T at position 10 (*3*, *4*). The importance of these bases is also indicated by the finding that the reduction of affinity by Lys−1 substitution is lower than that of the DNA substitution within the finger 1 subsite. But the structural basis for the base preference at positions 10 and 11 remains to be determined. In any event, the mechanism of interaction between finger 1 and DNA is different from the Zif268 model.

A relatively high degree of sequence diversity has been observed in other zinc finger proteins. Transcription factor IIIA (TFIIIA) involves nine zinc fingers and binds to the internal control region (ICR) of the 5S RNA gene (*5*). The peptide containing the first three fingers 1−3 can bind to the C-block element of the ICR (*31*, *32*). Certain biochemical experiments revealed that TFIIIA finger 1 makes a smaller contribution to the specific DNA binding than fingers 2 and 3 (*33*−*37*). A recent solution structure of the TFIIIA N-terminal three fingers−DNA complex (*38*, *39*) demonstrated that the orientations of the α-helices of fingers 2 and 3 in the DNA major groove are almost similar to that of Zif268, and that the helix of finger 1 makes more extensive contacts across the major groove. The long side chains of Lys−1 and Lys3 in finger 1 can extend into the major groove to make multiple contacts with DNA. These contacts distinct from Zif268 would bring about relaxed sequence selectivity for its DNA subsite. It is likely that Sp1 finger 1 also docks in the major groove in a geometry like that of TFIIIA finger 1 and tolerates some degree of sequence diversity at its subsite. The Wilms' tumor suppressor, WT1, is a transcriptional regulatory protein with four zinc fingers (*40*). The three C-terminal fingers (fingers 2−4) are extensively homologous to Zif268, and the three N-terminal fingers (fingers 1−3) are similar to Sp1. Previous studies showed that WT1 finger 1 also makes a small contribution to the specific DNA binding of WT1 (*41*−*43*). The first WT1 finger may bind to DNA in a manner similar to that of Sp1 finger 1. However, the role of WT1 finger 1 would be more complicated because four fingers participate in the DNA binding simultaneously and WT1 has two splicing isoforms of the zinc finger region (*44*).

In conclusion, this study demonstrates that all of the three zinc fingers of Sp1 participate in the high-affinity binding to the GC box DNA but that the contribution of finger 1 to the DNA binding affinity and the sequence selectivity of Sp1 is smaller than that of fingers 2 and 3. The mechanism of interaction between finger 1 and DNA is different from the Zif268 model because finger 1 covers 5 bp DNA, Lys−1 makes contacts with two guanine bases, and other putative key amino acids in the Zif268 model do not affect DNA binding. In addition, it is possible that the orientation of Sp1 finger 1 in the DNA major groove is similar to that of TFIIIA finger 1 rather than to that of Zif268 fingers. The unique interaction between finger 1 and the GC box DNA may be indispensable for binding of Sp1 to diverse GC box sequences bound in many promoter regions.

REFERENCES

1. Dynan, W. S., and Tjian, R. (1983) *Cell 32*, 669−680.
2. Kadonaga, J. T., Carner, K. R., Masiarz, F. R., and Tjian, R. (1987) *Cell 51*, 1079−1090.
3. Kadonaga, J. T., Jones, K. A., and Tjian, R. (1986) *Trends Biochem. Sci. 11*, 20−23.
4. Bucher, P. (1990) *J. Mol. Biol. 212*, 563−578.
5. Miller, J., McLachlan, A. D., and Klug, A. (1985) *EMBO J. 4*, 1609−1614.
6. Brown, R. S., Sander, C., and Argos, P. (1985) *FEBS Lett. 186*, 271−274.
7. Desjarlais, J. R., and Berg, J. M. (1992) *Proteins 12*, 101−104.
8. Kriwacki, R. W., Schultz, S. C., Steitz, T. A., and Caradonna, J. P. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 9759−9763.
9. Pavletich, N. P., and Pabo, C. O. (1991) *Science 252*, 809−817.
10. Elrod-Elickson, M., Rould, M. A., Nekludova, L., and Pabo, C. O. (1996) *Structure 4*, 1171−1180.
11. Swirnoff, A. H., and Milbrandt, J. (1995) *Mol. Cell. Biol. 15*, 2275−2287.
12. Berg, J. M. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 11109−11110.
13. Kuwahara, J., Yonezawa, A., Futamura, M., and Sugiura, Y. (1993) *Biochemistry 32*, 5994−6001.
14. Thiesen, H.-J., and Bach, C. (1990) *Nucleic Acids Res. 18*, 3203−3209.
15. Shi, Y., and Berg, J. M. (1995) *Chem. Biol. 2*, 83−89.
16. Nagaoka, M., and Sugiura, Y. (1996) *Biochemistry 35*, 8761−8768.
17. Kunkel, T. A., Bebenek, K., and McClary, J. (1991) *Methods Enzymol. 204*, 125−139.
18. Kuwahara, J., and Coleman, J. E. (1990) *Biochemistry 29*, 8627−8631.
19. Brenowitz, M., Senear, D. F., Shea, M. A., and Ackers, G. K. (1986) *Methods Enzymol. 130*, 132−181.
20. Isalan, M., Choo, Y., and Klug, A. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 5617−5621.
21. Thiesen, H.-J., and Bach, C. (1991) *FEBS Lett. 283*, 23−26.
22. Desjarlais, J. R., and Berg, J. M. (1992) *Proteins 13*, 272.
23. Desjarlais, J. R., and Berg, J. M. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 7345−7349.
24. Cunningham, B. C., and Wells, J. A. (1989) *Science 244*, 1081−1085.
25. Pavletich, N. P., and Pabo, C. O. (1993) *Science 261*, 1701−1707.
26. Fairall, L., Schwabe, J. W. R., Chapman, L., Finch, J. T., and Rhodes, D. (1993) *Nature 366*, 483−487.
27. Houbaviy, H. B., Usheva, A., Shenk, T., and Burley, S. K. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 13577−13582.
28. Sjøttem, E., Andersen, C., and Johansen, T. (1997) *J. Mol. Biol. 267*, 490−504.
29. Shi, Y., and Berg, J. M. (1996) *Biochemistry 35*, 3845−3848.
30. Narayan, V. A., Kriwacki, R. W., and Caradonna, J. P. (1997) *J. Biol. Chem. 272*, 7801−7809.

31. Christensen, J. H., Hansen, P. K., Lillelund, O., and Thϕgersen, H. C. (1991) *FEBS Lett. 281*, 181−184.
32. Liao, X., Clemens, K. R., Tennant, L., Wright, P. E., and Gottesfeld, J. M. (1992) *J. Mol. Biol. 223*, 857−871.
33. You, Q., Veldhoen, N., Baudin, F., and Romaniuk, P. J. (1991) *Biochemistry 30*, 2495−2500.
34. Del Rio, S., and Setzer, D. R. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 168−172.
35. Del Rio, S., Menezes, S. R., and Setzer, D. R. (1993) *J. Mol. Biol. 233*, 567−579.
36. Veldhoen, N., You, Q., Setzer, D. R., and Romaniuk, P. J. (1994) *Biochemistry 33*, 7568−7575.
37. Zang, W.-Q., Veldhoen, N., and Romaniuk, P. J. (1995) *Biochemistry 34*, 15545−15552.
38. Foster, M. P., Wuttke, D. S., Radhakrishnan, I., Case, D. A., Gottesfeld, J. M., and Wright, P. E. (1997) *Nat. Struct. Biol. 4*, 605−608.
39. Wuttke, D. S., Foster, M. P., Case, D. A., Gottesfeld, J. M., and Wright, P. E. (1997) *J. Mol. Biol. 273*, 183−206.
40. Call, K. M., Glaser, T., Ito, C. Y., Buckler, A. J., Pelletier, J., Haber, D. A., Rose, E. A., Kral, A., Yegar, H., Lewis, W. H., Jones, C., and Housman, D. E. (1990) *Cell 60*, 509−520.
41. Drummond, I. A., Rupprecht, H. D., Rohwer-nutter, P., Lopez-guisa, J. M., Madden, S. L., Rauscher, F. J., III, and Sukhatme, V. P. (1994) *Mol. Cell. Biol. 14*, 3800−3809.
42. Nakagama, H., Heinrich, G., Pelletier, J., and Housman, D. E. (1995) *Mol. Cell. Biol. 15*, 1489−1498.
43. Hamilton, T. B., Barilla, K. C., and Romaniuk, P. J. (1995) *Nucleic Acids Res. 23*, 277−284.
44. Haber, D. A., Bucker, A. J., Glaser, T., Call, K. M., Pelletier, J., Sohn, R. L., Douglass, E. C., and Housman, D. E. (1990) *Cell 61*, 1257−1269.

BI9727646