Pergamon

Bioorganic & Medicinal Chemistry 9 (2001) 2259–2267

BIOORGANIC & MEDICINAL CHEMISTRY

Effect of Arginine Mutation of Alanine-556 on DNA Recognition of Zinc Finger Protein Sp1

Keizo Matsushita and Yukio Sugiura*

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

Received 11 December 2000; accepted 19 March 2001

Abstract—Human transcription factor Sp1, which contains three Cys_2His_2-class zinc finger motives, specifically binds to the so-called GC box DNA. It has been indicated that finger 1 has a unique DNA-binding mode compared with fingers 2 and 3, or the Zif268 model. Therefore, we investigate the role of Ala at position 6 on the recognition helix, which is not responsible for guanine recognition and highly conserved among Sp1 family. Several Ala-556 mutations of Sp1 bind to DNA with different DNA-binding features. In particular, the Ala→Arg substitution alters the DNA-binding contribution of the three zinc fingers in Sp1. In this case, the DNA-binding specificity of each finger decreases in the order $2 > 1 > 3$. This result reveals that one amino acid in position 6 plays an important role not only for the selectivity to the putative finger 1 subsite, but also for the binding mode of the three fingers to each finger subsite. Probably, Ala-556 is indispensable to characterize the binding mode of the Sp1 zinc fingers, namely the diverse binding contribution of finger 1 and the rigid binding one of finger 3. In Sp1, the N-terminal finger 1 serves as a 'hinge finger'.

Introduction

Zinc finger is a ubiquitous eukaryotic DNA-binding motif used by transcription factors or hormone receptors. The zinc finger of Cys_2His_2 (C_2H_2) class, originally identified in the *Xenopus* transcription factor TFIIIA, has the consensus sequence (Tyr, Phe)-(X)-Cys-$(X)_{2,4}$-Cys-$(X)_3$-Phe-$(X)_5$-Leu-$(X)_2$-His-$(X)_{3\text{-}5}$-His-$(X)_{2\text{-}6}$.[1,2] On the basis of its simple finger structure and definitive interaction with DNA, DNA–protein recognition rule has been derived from study of this zinc finger class. In the Zif268–DNA complex, the first X-ray structure was determined for this family.[3–5] The result revealed that (1) the C_2H_2-type zinc finger contains a short two-stranded antiparallel β sheet and an α helix, (2) each helix of zinc fingers is inserted into the major groove of DNA, and (3) the residues in four positions of each helix (positions −1, 2, 3, and 6; named to key positions) contact three adjacent bases (a triplet). However, recent studies indicate that even the DNA recognition belong to the zinc finger of C_2H_2 class is not necessarily governed by such a simple rule, namely a one-to-one correspondence of the key amino acids to the DNA bases.[6–10]

Human transcription factor Sp1 isolated from HeLa cells[11,12] is a C_2H_2-type zinc finger protein called the Krüppel type which includes Sp2-4, EKLF, and BTEB.[13,14] Sp1 specifically binds to guanine-rich recognition elements (GC-boxes) within the 5′-flanking promoter sequences,[15,16] and enhances transcription from numerous viral and cellular genes such as the housekeeping gene, CRE BP-1 gene, and simian virus 40 gene.[15,17] For example, Sp1 regulates the maintenance of differentiated cells.[18] It has been suggested that the consensus sequence of Sp1-binding DNA is the dodecanucleotide 5′-(G/T)GGGCGG(G/A)(G/A)(C/T)-3′.[15,16] The model of three-zinc finger Sp1–DNA interaction had been analogously proposed from the Zif268–DNA interaction mode (Fig. 1).[17,19,20] However, some previous studies showed that the DNA binding mode of the N-terminal finger of Sp1, finger 1, differs from those of the other fingers of Sp1 and the Zif268 model.[15,16,19–22] The Sp1 finger 1 covers the 5 bp subsite at the 3′-portion of the GC box sequence and has moderate DNA specificity. This contribution to the DNA-binding affinity and the sequence specificity is smaller than those of fingers 2 and 3. In Zif268, the interactions at the edge of the complex (fingers 1 and 3) are weaker than similar interactions embedded in the core of the complex (finger 2).[23] In Sp1, on the other hand, the DNA binding contribution of the Sp1 finger ($3 > 2 > 1$) is obviously different from that of the Zif268 type. Lys-550 immediately preceding

*Corresponding author. Tel.: +81-774-38-3210; fax: +81-774-32-3038; e-mail: sugiura@scl.kyoto-u.ac.jp

0968-0896/01/$ - see front matter
PII: S0968-0896(01)00134-1

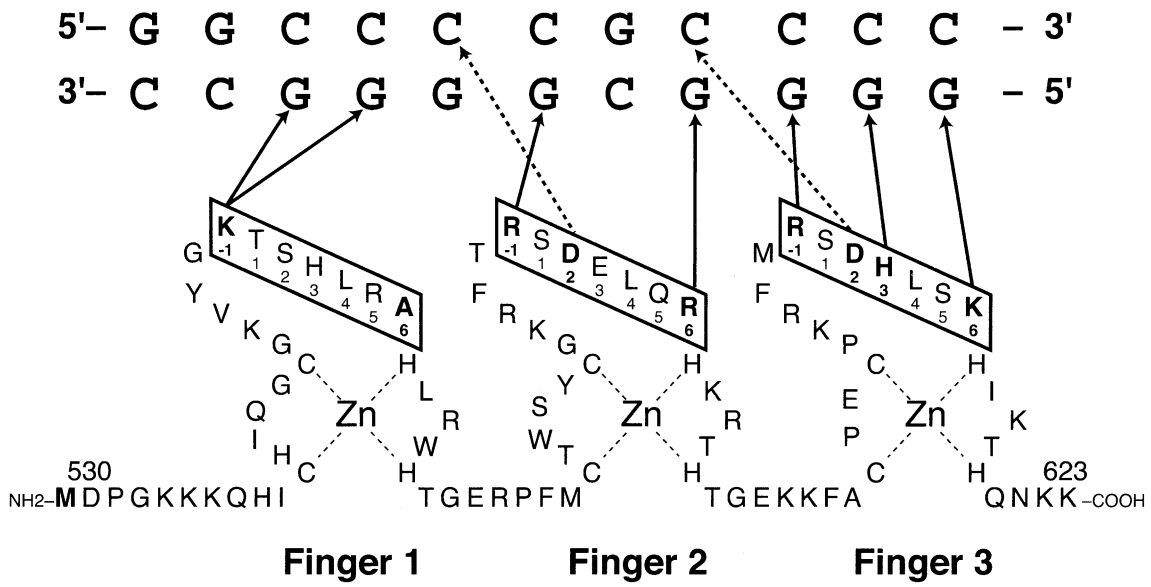


Figure 1. Amino acid sequence of Sp1 zinc finger region and its putative DNA binding mode with GC box DNA. Amino acids are denoted by one-letter codes. The numbers on the residues are the numbers of amino acid sequences in native Sp1. Recognition helices in each finger are boxed, and the numbers of the helical positions are shown in bold-type. Solid arrows depict the amino acid–base contacts.

finger 1 helix is essential for the recognition of two guanine bases, G(8) and G(9), whereas Thr-551, Ser-552, His-553 and Arg-555 do not directly participate in the DNA binding. From certain biological evidence and the solution structure of the first three zinc fingers of TFIIIA bound to the cognate DNA sequence,[24–30] the Sp1 finger 1 is similar to the TFIIIA finger 1 rather than to the Zif268 fingers in the DNA-binding mode. The TFIIIA finger 1 has the homology of the key amino acids at position −1 (Lys) and 6 (Ala) for the finger 1 of the Sp1 family. Our attention was focused on Ala, which both fingers contain as the key amino acid at position 6 on the recognition helix. In the typical zinc finger, this position is almost a basic amino acid residue such as Arg and Lys. In TFIIIA, however, the side chain of Ala at position 6 provides the contact surface on finger 1 and packs against the residues on finger 2.[28,29] In this study, some Ala-556-substituting Sp1 mutants have been prepared and their effects on the Sp1–DNA interaction were evaluated by several electrophoretic methods. The binding mode of the finger 3-deleted Ala-substituting mutant was also investigated. The amino acid at position 6 on a recognition helix probably plays an important role in the overall fingers–DNA interaction.

Results

Binding affinity of Ala-556 mutants for the GC box

Usually, Ala as a key amino acid in the zinc finger is unable to recognize guanine bases.[17,30–33] An Arg, mostly conserved at position 6 of the recognition helix, is employed to gain the energy of DNA binding rather than sequence selectivity.[34–36] To assess the role of Ala-556 in the Sp1 zinc finger–DNA complex formation, Ala was substituted by Arg, Ser, Thr, or Val (namely Sp1A556R, Sp1A556S, Sp1A556T, or Sp1A556V). Table 1 summarizes the binding affinities of these Sp1 mutants to the cognate GC box sequence analyzed by the gel mobility-shift assay. The present binding affinity, somewhat different from our previous data is presumably due to different experimental condition.[21] The free energy changes (ΔG) were also evaluated from K_d (Table 2). Due to the lower contribution of finger 1 compared to finger 2 or 3,[15,16,19–22] small differences among their affinities to the GC box were detected. However, the Ala→Ser, Thr, and Val mutation caused DNA binding with lower affinities than Sp1(530–623) and the only Ala→Arg change resulted in the opposite effect compared with the other substitutions, namely a higher binding affinity.

Binding selectivity of Ala-556 mutants for each finger DNA subsite

Fingers 2 and 3 of Sp1 recognize triplet DNA at the middle and 5′-portion of Sp1-binding site, respectively, whereas the Sp1 finger 1 is capable of recognizing the 5 bp DNA at the 3′-portion of Sp1-binding site.[21] Does

Table 1. Dissociation constants (K_d)[a] of several substituted Sp1 mutants for the GC box and mutated GC box oligonucleotides

	K_d (nM)[a]				
Binding site[b]	530-623	A556R	A556S	A556T	A556V
GC	39	29	63	100	59
F1-AT	130	170	200	270	150
F2-AT	740	550	820	1400	1500
F3-AT	5900	130	3100	5300	2300

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

Table 2. Free energy change for the GC box and mutated GC box oligonucleotides with several substituted Sp1 mutants

	ΔG (kJ/mol)[a]					Selectivity $\Delta\Delta G$ (kJ/mol)[c]				
Binding site[b]	530–623	A556R	A556S	A556T	A556V	530–623	A556R	A556S	A556T	A556V
GC	−41.54	−42.30	−40.40	−39.22	−40.56	—	—	—	—	—
F1-AT	−38.64	−37.94	−37.63	−36.88	−38.34	2.91	4.36	2.77	2.35	2.22
F2-AT	−34.39	−35.10	−34.15	−32.80	−32.61	7.16	7.20	6.25	6.42	7.95
F3-AT	−29.33	−38.64	−30.92	−29.58	−31.59	12.2	3.66	9.48	9.65	8.96

[a]ΔG was calculated from the equation $\Delta G = -RT \ln(1/K_d)$, where T is 293 K and R is 8.3144 J/mol.
[b]The nomenclature is described in the text (see Figs 1 and 2).
[c]$\Delta\Delta G$ was calculated from the relationships $\Delta\Delta G_{F1\text{-}AT} = \Delta G_{F1\text{-}AT} - \Delta G_{GC}$, $\Delta\Delta G_{F2\text{-}AT} = \Delta G_{F2\text{-}AT} - \Delta G_{GC}$, $\Delta\Delta G_{F3\text{-}AT} = \Delta G_{F3\text{-}AT} - \Delta G_{GC}$, respectively.

the above-mentioned mutation affect the recognition bases for each finger? Three kinds of mutated GC box, whose sequences in each finger subsite are replaced by the AT-rich sequence (F1-AT, F2-AT, and F3-AT), were prepared (Fig. 2). Each gel mobility-shift assay gave the following results (Tables 1 and 2): (1) All of the Sp1 mutants bound to F1-AT with a slightly higher dissociation constant (K_d = 150–270 nM), (2) most of these mutants bound with a slightly lower affinity to F2-AT and with higher affinity to F3-AT, except for Sp1A556R (820–1500 and 2300–5300 nM, respectively), and (3) Sp1A556R showed a somewhat strong affinity to F2-AT (550 nM) and a considerably stronger affinity to F3-AT (130 nM).

Detection of contacted guanine bases by methylation interference experiments

By using interference experiment of guanine methylation modified by dimethyl sulfate, we obtained the information about the guanine contacts for the DNA–protein complex. Figure 3A presents the methylation interference patterns of Sp1(530–623) and Sp1A556R for a 41 bp DNA fragment containing the native GC box. Figure 3B shows a histogram of the extent of interference. The pattern of Sp1(530–623) corresponds well to our previous results.[20,21,37] In the case of Sp1A556R concerning the putative finger 1-binding subsite, however, the methylation interference of the G(7) site was increased, and those of G(10′) and G(11′) sites were apparently reduced. To our surprise, the methylation interference of the G(4) and G(5′) sites in the finger 2-binding subsite, and G(1) in the finger 3-binding subsite, were also remarkably reduced. Sp1A556S, Sp1A556T, and Sp1A556V did not give any patterns of interference distinct from Sp1(530–623) (Fig. 3B, data not shown).

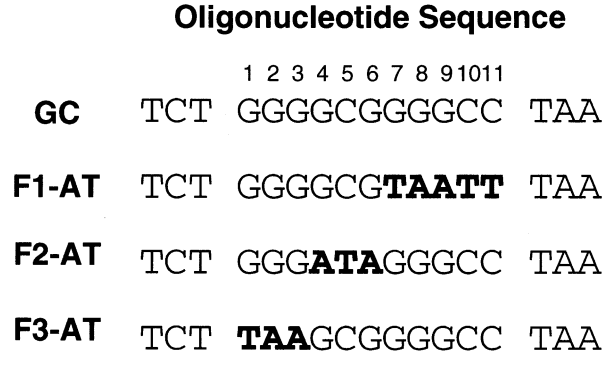


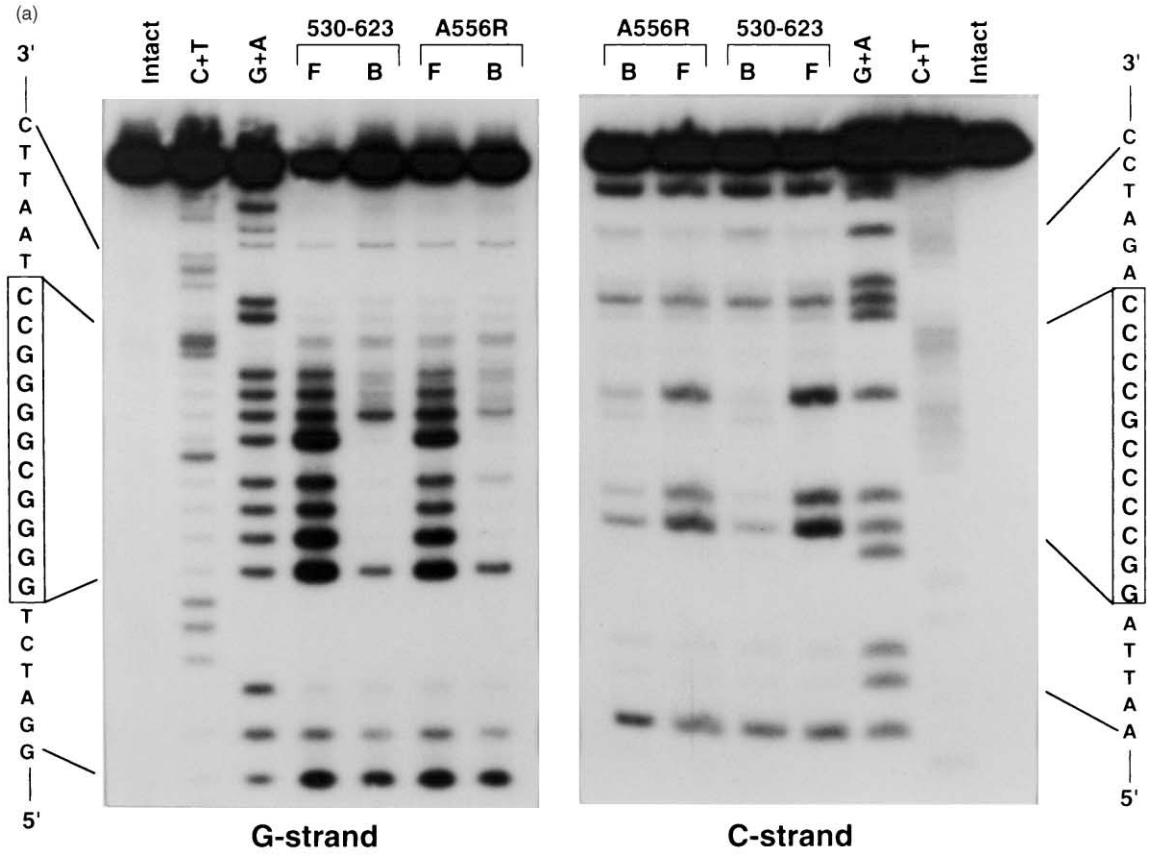
Figure 2. Native GC box and mutated GC box sequences. Substituted nucleotides are written in bold. The base numbers in the GC box are also shown. Nomenclatures 'F1', 'F2', and 'F3' refer to sequences mutated in the putative DNA binding subsites of each finger. 'AT' denotes AT-rich sequence sub-selected in random selection.

Detection of structural alteration of DNA by hydroxyl radical footprinting

The footprinting using the hydroxyl radical is useful for determining the protein–DNA interaction and polynucleotide conformation.[38] The hydroxyl radical cleaves DNA by abstracting a hydrogen atom exposed in the minor groove from a deoxyribose residue in the DNA backbone.[30,39] We used it to investigate the conformational change and minor groove width of the protein–bound DNA.[40] Figure 4A shows the DNA cleavage pattern by hydroxyl radical in the absent and present of Sp1(530–623) or Sp1A556R. Zinc finger proteins bind in the major groove of DNA, unwind DNA, and induce a distinctive form of B-DNA.[3–5,41–44] This cleavage reduction suggests that the protein binding narrows the minor groove of DNA. However, the extent of quenching was different between each DNA base of the two complexes. In the case of Sp1(530–623), strong cleavage quenching was observed at the putative finger 3-binding subsite, and moderate reduction at the finger 2 subsite. No quenching was detected at the finger 1 subsite. This result corresponds to the DNA-binding contribution of each finger. In the case of Sp1A556R, the cleavage of the finger 1- and 2-binding subsites was reduced, but the cleavage quenching of G(1) and T(−1) on the finger 3 subsite was slightly diminished compared with Sp1(530-623).

Binding specificity of finger deletion mutants

The two-finger peptides, in particular Sp1(zf12), remarkably reduced the binding affinity for the GC box sequence in comparison with Sp1(530–623).[21] We also obtained the dissociation constants (K_d) of these mutants (Table 3) and the free energy changes (ΔG) evaluated from K_d (Table 4). A gel shift assay of Sp1(zf23) to the native GC box presented two kinds of bound-band depending on its concentration (Matsushita et al., unpublished data). The band to F1-AT showed only a fast-migrated one at the same low concentration, suggesting that Sp1(zf23) can form the complex with the finger 1 subsite and/or its adjacency in spite of its lower affinity. When only considering the



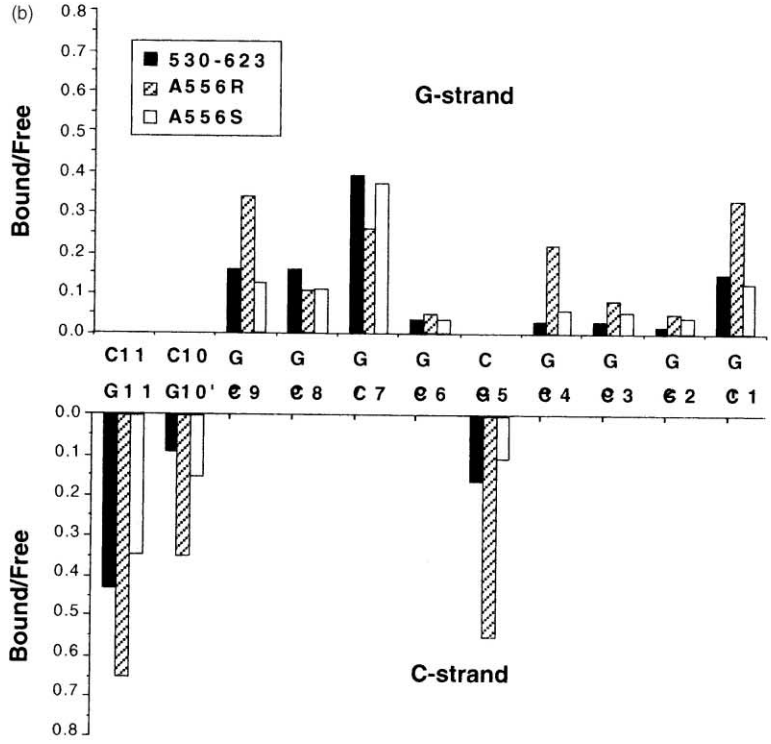


Figure 3. (a) Methylation interference experiments of Sp1(530–623) and Sp1A556R to the GC box DNA. The left and right panels show the results for the G- and C-strands, respectively. Nomenclatures 'F' and 'B' denote the free and peptide-bound DNA samples. Nomenclatures 'Intact', 'G + A', and 'C + T' indicate intact DNA, G + A, and C + T of the Maxam–Gilbert sequencing reaction, respectively. (b) A histogram showing the extent of methylation interference by Sp1(530–623), Sp1A556R, and Sp1A556S. Cleavage bands were visualized by autoradiography of the gel and their densitometric patterns were measured using NIH Image. The extent of interference was calculated as the ratio of the cutting probabilities for the two bands (B/F).

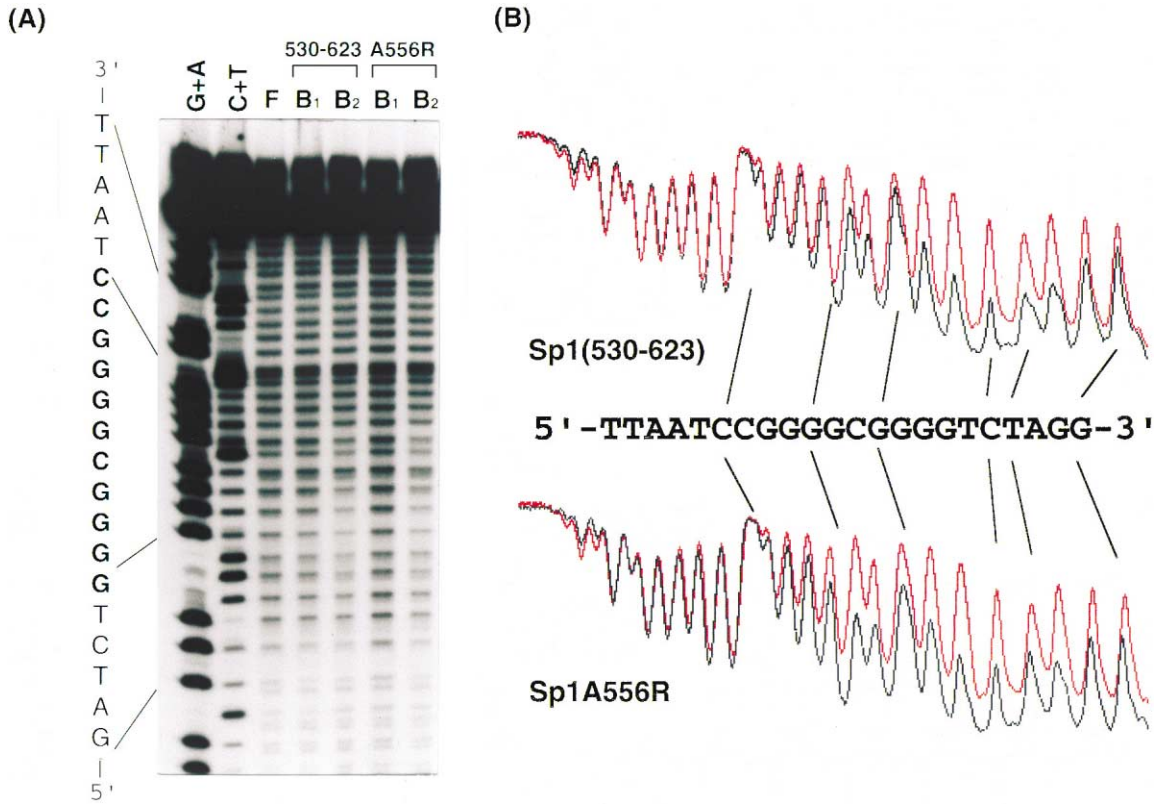


Figure 4. Hydroxyl radical cleavage analyses of Sp1(530–623) and Sp1A556R to the GC box DNA. (A) Autoradiograph of hydroxyl radical cleavage products. Nomenclatures 'F', 'B_1', and 'B_2' present free, 32 and 512 nM of peptide samples with the GC box DNA. Nomenclatures 'G + A' and 'C + T' indicate G + A and C + T of the Maxam–Gilbert sequencing reaction, respectively. (B) Densitometric traces of lane 'F' (gray) and lane 'B_2' (black). Upper and lower scans show cleavage alteration of Sp1(530–623) and Sp1A556R, respectively.

Table 3. Dissociation constants (K_d)[a] of several two-finger mutants for the GC box and mutated GC box oligonucleotides

	K_d (nM)[a]						
Binding site[b]	Sp1(zf12)	Sp1(zf23)	(zf12)AK	(zf12)AR	(zf12)AS	(zf12)AT	(zf12)AV
GC	980	160	600	6.50	6100	> 8000	ND[c]
F1-AT	4600	220	2900	ND[c]	> 8000	ND[c]	—
F2-AT	ND[c]	> 8000	> 8000	3500	ND[c]	ND[c]	—
F3-AT	> 8000	ND[c]	> 8000	13	6100	6900	—

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

first band, the dissociation constant of Sp1(zf23) for F1-AT (220 nM) was close to that for the native GC box (160 nM). For F2-AT and F3-AT, the dissociation constant of Sp1(zf23) was not estimated under this experimental condition. Under the competition assay, the affinity was 6.3- and 5.2-fold reduced, respectively (data not shown). Consequently, the deletion of finger 1 adapted to only the mutation for the finger 1 subsite and decreased the specificity for the finger 3 subsite. On the other hand, the dissociation constant of Sp1(zf12) for F3-AT was remarkably higher than that for the native GC box, but the specificity for the finger 1 subsite was not so high. The constant for F2-AT was not evaluated under this experimental condition.

Effects of Ala-556 mutations in Sp1(zf12)

Does proper specificity of Sp1(zf12) for the finger 3 subsite arise from the unique character of finger 1? This question was clarified by the introduction of the Ala-556 mutation to Sp1(zf12). The same substitutions as those to Sp1(530-623) were achieved to Sp1(zf12) [these mutant peptides were named as Sp1(zf12)A556R, Sp1(zf12)A556S, Sp1(zf12)A556T, and Sp1(zf12)A556V]. In addition, Sp1(zf12)A556K (Ala→Lys change) was also prepared. The binding affinities of these finger 1–2 mutants to various DNA fragments were investigated using the gel mobility-shift assay (Table 3). Ala→Ser, Thr, and Val mutants of Sp1(zf12) bound with an extremely lower affinity even to the native GC box.

Table 4. Free energy change for the GC box and mutated GC box oligonucleotides with several two-finger Sp1 mutants

	ΔG (kJ/mol)[a]					Selectivity $\Delta\Delta G$ (kJ/mol)[c]				
Binding site[b]	Sp1(zf12)	Sp1(zf23)	(zf12)AK	(zf12)AR	(zf12)AS	Sp1(zf12)	Sp1(zf23)	(zf12)AK	(zf12)AR	(zf12)AS
GC	−33.70	−38.14	−34.92	−45.91	−29.26	—	—	—	—	—
F1-AT	−29.93	−37.36	−31.04	ND[d]	> −28.58	3.77	0.78	3.89	ND[d]	> 0.68
F2-AT	ND[d]	> −28.58	−27.12	−30.64	ND[d]	ND[d]	> 9.56	7.80	15.3	ND[d]
F3-AT	−25.68	ND[d]	−28.36	−44.30	−29.24	8.02	ND[d]	6.56	1.62	0.014

[a]ΔG was calculated from the equation $\Delta G = -RT \ln(1/K_d)$, where T is 293 K and R is 8.3144 J/mol.
[b]The nomenclature is described in the text (see Figs 1 and 2).
[c]$\Delta\Delta G$ was calculated from the relationships $\Delta\Delta G_{F1\text{-}AT} = \Delta G_{F1\text{-}AT} - \Delta G_{GC}$, $\Delta\Delta G_{F2\text{-}AT} = \Delta G_{F2\text{-}AT} - \Delta G_{GC}$, $\Delta\Delta G_{F3\text{-}AT} = \Delta G_{F3\text{-}AT} - \Delta G_{GC}$, respectively.
[d]ND, not determined.

However, Sp1(zf12)A556R bound more tightly to the native GC than Sp1(zf12). From the result of DNase I footprinting (data not shown), this mutant can bind to the 5′-flanking independent site. A very low specificity of Sp1(zf12)A556R for the finger 3 subsite was observed and also a high specificity for the fingers 1 and 2 subsites. The deletion of finger 3 from Sp1A556R surely led to a loss in specificity for the finger 3 subsite. Figure 5 presents the methylation interference patterns of Sp1(zf12)A556R. In the case of the putative finger 1-binding subsite, all the methylation interference of G-strand [containing G(7)] were clearly observed, but those of G(10′) and G(11′) sites were not salient. As predicted, the finger 3-binding subsite was not interfered, and the G(4) in the finger 2-binding subsite was slightly interfered as well as Sp1A556R. Figure 5 also shows that Sp1(zf12)A556R can bind to the 5′-flanking site, 5′-GTGGATC-3′. However, the Ala→Lys mutant of Sp1(zf12) was similar to Sp1(zf12) in spite of the ability of the Lys-guanine hydrogen bond, though there was a reduction in specificity for the finger 2. The Lys side chain usually relates to the guanine base recognition. However, the long side chain of Lys is flexible and its donor can form fluctuating or bifurcated hydrogen bonds to a guanine or a thymine base.[45] Maybe, Sp1(zf12)A556K can not gain a large binding energy by formation of the flexible hydrogen bond. In fact, Lys (position 6 on the Sp1 finger 3 helix)-G(1) contact may be weak as indicated by the methylation interference analysis.

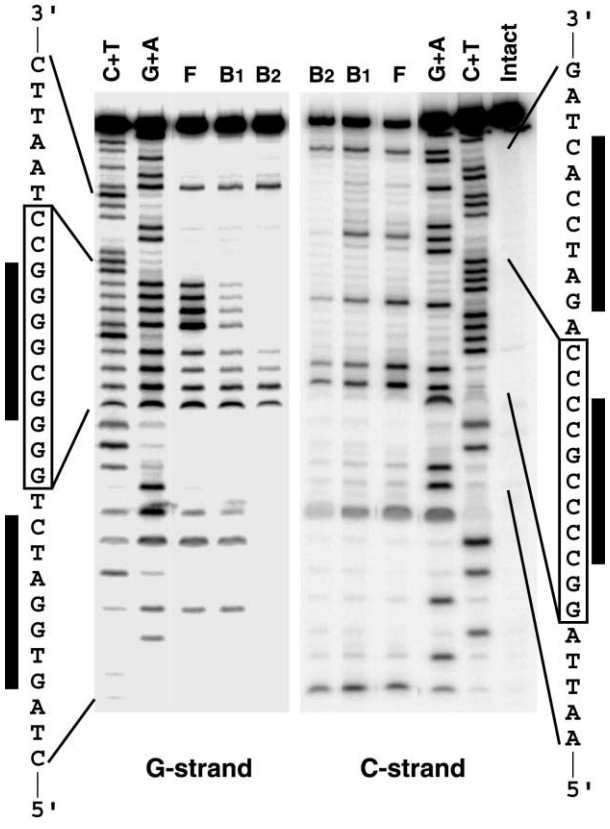


Figure 5. Methylation interference experiments of Sp1(zf12)A556R to the GC box DNA. The left and right panels show the results for the G- and C-strands, respectively. Nomenclatures 'F', 'B_1', and 'B_2' denote the free, one and two peptide-bound DNA samples, respectively. Nomenclatures 'Intact', 'G + A', and 'C + T' indicate intact DNA, G + A, and C + T of the Maxam–Gilbert sequencing reaction, respectively. The solid bars mean the peptide-bound positions.

Discussion

It is known that Sp1 reads the 5′-GGGGCGGGGCC-3′ sequence.[1,16] In particular, the N-terminal finger of Sp1, finger 1, has an inherently wide recognition range, which extends to five bases, the 5′-GGGCC-3′ region. Its contribution to the DNA binding affinity and the sequence specificity is smaller than that of fingers 2 and 3.[21] Lys-550 immediately preceding the finger 1 recognition helix is essential for the recognition of the two guanine bases, G(8) and G(9). The similarity of the Sp1 finger 1 with the TFIIIA finger 1 has been discussed.[21] Both fingers contain Ala at key position 6 of helix, which is not responsible for guanine recognition and highly conserved among Sp1 family. Herein, more detailed information of Sp1 finger 1 has been obtained. The present results reveal that the substitution for Ala-556 on the finger 1 helix plays an important role in the specificity of the putative finger 1-binding subsite. As shown in the result of the methylation interference, Sp1A556R and Sp1(zf12)A556R display lower specificity toward C(10) and C(11) at the finger 1 subsite than Sp1(530–623). The slight decrease in specificity for the 3′-portion is also indicated by the $\Delta\Delta G$ of Sp1A556R for CC/GG→AA/TT at the 3′-portion (data not shown). Moreover, Sp1(zf223), whose finger 1 is replaced by finger 2, does not recognize position 11 (Uno et al., unpublished data). The Arg residue at the helix in many zinc fingers has been conserved. Its two nitrogen donors can form a pair of hydrogen bonds

with two acceptors of the guanine base in the DNA major groove. Certainly, Ala→Arg substitution at position 6 seems to form the new contact to the finger 1 subsite and to change the binding subsite range of finger 1. In addition, the present results suggest that the unique DNA interaction mode of the Sp1 finger 1 depends on Ala-556. However, hydroxyl radical cleavage analysis presented no different cleavage patterns on C(10) and C(11). Probably, it is attributed to the low binding contribution of finger 1. The finger 1 of Sp1(530–623) may contact and/or dock to its DNA subsite with little or no structural changes in DNA. However, the Ala→Arg substitution induces a structural change in the finger 1 subsite.

Sp1A556R converts from the Sp1 recognition mode to the Zif268 mode. The difference in the free energy for the GC box between the three-finger and two-finger peptides determines the binding energy of the deleting finger. In the case of Sp1(530–623), finger 1 and 3 deletions lose 3.39 and 7.82 kJ/mol of DNA-binding energy, respectively. In the case of Sp1A556R, the finger 1 deletion induces the loss of binding energy (4.14 kJ/mol) similar to Sp1(530–623). However, the finger 3 deletion gains 3.60 kJ/mol of DNA binding energy. The result suggests that the Ala→Arg mutant of finger 1 causes the hindrance of finger 3 for DNA binding. This suggestion is supported by the difference of the methylation interference in the finger 1 and 2 binding sites between Sp1A556R and Sp1(zf12)A556R. Probably, in the Ala→Arg mutation, the interruption is attribute to the extreme loss of the DNA binding flexibility of finger 1. Table 2 shows the change in the binding specificity for each finger subsite. Corresponding to the result of the hydroxyl radical cleavage experiments, the DNA binding specificity (finger $3>2>1$) of the Sp1 zinc finger was altered compared to that of Zif268 or the designed zinc finger peptides. These common fingers showed the so-called end effect, similar to the DNA duplex formation.[23,46] The methylation interference experiment and hydroxyl radical analysis also revealed the changes in the zinc finger binding mode and local DNA structure at the finger 2 and 3 subsites by Ala→Arg. However, these alterations of conformation and/or orientation in Sp1(530–623) and Sp1A556R do not originate only from the hydrogen bond formation by the arginine side chain. Ala→Ser, Thr, and Val mutants counteracted the significant difference among each binding specificity to finger-binding subsite, despite the similar order of specificity ($3>2>1$). Furthermore, Sp1(zf12) has a high specificity for the finger 3 subsite, unlike Sp1(zf12)A556R for the finger 3-binding subsite. In addition, a detectable binding affinity was not obtained in the Ala→Ser, Thr, and Val substituted Sp1(zf12) mutants (Table 3, data not shown). These results indicate that Ala-556 is the favorable residue for stability of the zinc finger–DNA complex. The long-range motional correlation among the first three zinc finger domains of TFIIIA in the absence of DNA indicates that the average conformation of the three-finger is elongated in the internal linkers.[47] The inter-domain face of fingers 1–2 includes Asp at the turn and His at position 11 of the helix on finger 1, and Phe and Thr at the second strand on finger 2. In the presence of DNA, however, the internal linker regions are compactly folded. The contact surface on finger 1 is provided by the side chains of Trp and Ala at positions 2 and 6, which pack against Thr, Ser, and His at positions −2, −1, and 1 on finger 2.[28,29] These results show that the finger–finger interaction is directly or indirectly important for the finger orientation in the DNA complex. The alteration of the finger–finger interaction may cause the difference in the binding affinity or specificity of each finger to the binding subsite.[48] Our investigation suggests that Ala at position 6 on the finger 1 helix of Sp1 is also independent of the DNA binding mode of fingers 2 and 3. On the basis of high-resolution crystal structures of the variant Zif268–DNA complexes, the alterations of finger 1 had a relatively small effect on the structure and docking of fingers 2 and 3, in contrast with our report concerning Sp1.[5] If the consensus amino acid at position 6 on Zif268 finger 1, Arg, is substituted to Ala, remarkable change in the orientation and conformation of fingers 2 and 3 may occur.

As well as Sp1(zf23), Sp1(zf12)A556R binds to the GC sequences with a lower dissociation constant and has another binding sites out of the GC sequence (Fig. 5). The sequence selectivity of Sp1(zf12)A556R and Sp1(zf23) for the GC box is considerably weaker, namely less distinguishable, than Sp1(zf12) in spite of their higher affinities. This fact reveals that the sequence selectivity can prefer to a weak and broad recognition rather than strong and narrow one. Moreover, the deletion of finger 3 does not lead to a complete loss of specificity for the finger 3 subsite. Namely, finger 1 can assist the sequence selectivity and the high specificity for the finger 2 and 3 subsites. Probably, the N-terminal finger 1 of Sp1 plays different functional role from that of TFIIIA. TFIIIA comprises of nine zinc fingers, fingers 4–6 make less contribution to binding to the internal promoter element of the 5S rRNA gene, and finger 3 is necessary for TFIIIA–DNA binding.[27,30] On the while, Sp1 has three zinc fingers at the C-terminal region and the transcription activation domain at the N-terminal region. In general, a flexible hinge between functional domains containing protein and DNA binding is essential for various stable inter- and intra-molecular interactions (for example, NF-κB p50,[49] papillomavirus E2,[50] and the *lac* repressor[51]). Sp1 stabilizes the transcriptional pre-initiation complex, increasing the numbers of functional RNA poly II–promoter complexes.[52,53] The N-terminus of Sp1 contains Gln- and Ser/Thr-rich domains that associate with the TATA-box protein accessory factor $TAF_{II}110$[54,55] or cause tetrameric assembly inducing DNA looping.[56–58] Therefore, the lower contribution of the Sp1 finger 1 for the zinc finger–DNA complex may be responsible to the flexibility between the transcriptional activation domain and DNA binding domain. Such a finger is called a 'hinge finger', and then fingers 2 and 3 are enriched in the DNA binding activity. In fact, the peptide containing only the zinc finger domain can bind to the GC box DNA with almost the same affinity as full-length Sp1.[19,59] A small difference in the promoter sequence sometimes induces a drastic decrease in stability of the

transcriptional proteins–DNA complex.[60] In nature, Sp1 is widespread in various cell and tissue[14] and binds to the diverse GC box sequences. Moreover, almost the C-terminal zinc-finger of Sp1 family follows the N-terminal activation domain. Presumably, the 'hinge finger' gives ubiquity to Sp1. In order to clarify the role of Sp1 finger 1, we must await the definitive structure of Sp1–DNA complex by X-ray or NMR method.

Conclusion

The present work demonstrates that the Ala→Arg mutation at position 6 of the Sp1 finger 1 helix significantly affects the overall orientation and conformation of the Sp1–DNA complex. The three finger of Sp1(A556R) is similar to that of Zif268. Moreover, the Ala→Arg fingers 1–2 mutant adapts to the mutation for the finger 3-binding subsite, in contrast to the non-mutated one. The Sp1 finger 1 has a flexible binding activity, a weak but broad recognition, with the cognate DNA binding subsite. On the other hand, it is responsible to the recognition of fingers 2 and 3, namely enriches their contribution for DNA binding. The transcription activity of Sp1 may largely depend on this characteristic of finger 1 as the N-terminal 'hinge finger'.

Experimental

Preparation of zinc finger peptides from Sp1

The construction of the coding vector and the expression of the zinc finger peptides were proceeded as previously described.[21,37] The soluble form of the Sp1 peptides was purified by cation exchange and gel filtration chromatography. The former [HighS and UnoS (Bio-Rad)] was performed with phosphate buffer (pH 7.6) using NaCl gradient [0.13–2 M (or 1 M)], and the latter [Superdex 75 (Amersham Pharmacia Biotech)], with the TN buffer [10 mM Tris—HCl (pH 7.5) and 50 mM NaCl] containing 1 mM β-mercaptoethanol. Purified samples were stored at 4 °C. The finger 1- and 3-deleted mutants, Sp1(zf23) and Sp1(zf12), were prepared as previously described,[21] except for the purification with gel filtration chromatography. The Ala-556→X substituted series of Sp1(zf12), named Sp1(zf12)A556X, were generated from pUCSp1(530–595) and prepared as well as the other peptides.

Gel mobility-shift assay

The probe DNAs (41 bp *Hind*III-*Xba*I fragments) were prepared as previously described.[21] Binding reaction mixtures (final volume, 20 mL) contained 10 mM Tris–HCl (pH 8.0), 50 mM NaCl, 100 mM $ZnCl_2$, 1 mM β-mercaptoethanol, 0.05% Nonidet P-40, 5% glycerol, 0–4 or 0–8 μM Sp1 peptides, 25 ng/mL poly(dI-dC)·poly(dI-dC) double strand, and the 5′-end-labeled DNA fragment (~50 pM, 2000 cpm). After incubation at 20 °C for 30 min, the sample solutions were electrophoresed on a 12% nondenaturing polyacrylamide gel with TB buffer (88 mM Tris-base and 88 mM boric acid) at 20 °C. The bands were visualized by autoradiography and quantitated with the NIH Image (Version 1.58). The dissociation constants (K_d) of the Sp1 peptide–DNA fragment complexes were estimated as previously described,[21] and the free energy changes (ΔGs) were calculated by 293 K, except for the evaluation of Sp1(zf12)A556R.[61]

Methylation interference experiments

Methylation interference analyses were performed according to previous papers,[21,37] except for use of the reaction buffer containing 10 mM Tris–HCl (pH 8.0), 50 mM NaCl, 100 mM $ZnCl_2$, 1 mM β-mercaptoethanol, 0.05% Nonidet P-40, 5% glycerol, the 50–2000 nM Sp1 peptides, and 2.5 μg sonicated calf thymus DNA. The DNA probe was the 5′-end-labeled *Hind*III-*Xba*I fragment used above (~40 nM, 3×10^5 cpm). Estimation of obtained band patterns was carried out as previously described.[21]

Hydroxyl radical cleavage analyses

The hydroxyl radical cleavage experiments[39,62] were performed as follows. Each reaction buffer contained 10 mM Tris–HCl (pH 8.0), 50 mM NaCl, 1 mM β-mercaptoethanol, 20 μg/mL sonicated calf thymus DNA, the 5′-end-labeled *Hind*III-*Xba*I fragment, and 0–500 nM Sp1 peptides. After incubation at 20 °C for 30 min, the sample was digested at 20 °C for 2 min with hydroxyl radical produced by the three compounds, 100 μM $[Fe(EDTA)]^{2-}$, 0.003% hydrogen peroxide, and 1 mM sodium ascorbate. The cleavage products were precipitated by ethanol and sodium acetate, lyophilized, and then analyzed on a 10% denaturing polyacrylamide gel. Cleavage bands were visualized by autoradiography and their densitometric patterns were measured with the NIH Image (Version 1.58).

Acknowledgements

This study was supported in part by a Grant-in-Aid for Priority Project 'Biometals' and Scientific Research from the Ministry of Education, Science, Sports, and Culture, Japan.

References and Notes

1. Miller, J.; McLachlan, A. D.; Klug, A. *EMBO J.* **1985**, *4*, 1609.
2. Klug, A.; Rhodes, D. *Trends Biochem. Sci.* **1987**, *12*, 464.
3. Pavletich, N. P.; Pabo, C. O. *Science* **1991**, *252*, 809.
4. Elrod-Elickson, M.; Rould, M. A.; Nekludova, L.; Pabo, C. O. *Structure* **1996**, *4*, 1171.
5. Elrod-Elickson, M.; Benson, T. E.; Pabo, C. O. *Structure* **1998**, *15*, 451.
6. Jamieson, A. C.; Kim, S.-H.; Wells, J. A. *Biochemistry* **1994**, *33*, 5689.
7. Isalan, M.; Choo, Y.; Klug, A. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 5617.

8. Isalan, M.; Klug, A.; Choo, Y. *Biochemistry* **1998**, *37*, 12026.
9. Wolfe, S. A.; Greisman, H. A.; Ramm, E. I.; Pabo, C. O. *J. Mol. Biol.* **1999**, *285*, 1917.
10. Segal, D. J.; Dreier, B.; Beerli, R. R.; Barbas, C. F., III *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 2758.
11. Dynan, W. S.; Tjian, R. *Cell* **1983**, *32*, 669.
12. Kadonaga, J. T.; Carner, K. R.; Masiarz, F. R.; Tjian, R. *Cell* **1987**, *51*, 1079.
13. Turner, J.; Crossley, M. *Trends Biochem. Sci.* **1999**, *24*, 236.
14. Philipsen, S.; Suske, G. *Nucleic Acids Res.* **1999**, *27*, 2991.
15. Kadonaga, J. T.; Jones, K. A.; Tjian, R. *Trends Biochem. Sci.* **1986**, *11*, 20.
16. Bucher, P. *J. Mol. Biol.* **1990**, *212*, 563.
17. Berg, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 11109.
18. Marin, M.; Karis, A.; Visser, P.; Grosveld, F.; Philipsen, S. *Cell* **1997**, *89*, 619.
19. Kriwacki, R. W.; Schultz, S. C.; Steitz, T. A.; Caradonna, J. P. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9759.
20. Kuwahara, J.; Yonezawa, A.; Futamura, M.; Sugiura, Y. *Biochemistry* **1993**, *32*, 5994.
21. Yokono, M.; Saegusa, N.; Matsushita, K.; Sugiura, Y. *Biochemistry* **1998**, *37*, 6824.
22. Shi, Y.; Berg, J. M. *Chem. Biol.* **1995**, *2*, 83.
23. Choo, Y. *Nucleic Acids Res.* **1998**, *26*, 554.
24. Del Rio, S.; Setzer, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 168.
25. Del Rio, S.; Menezes, S. R.; Setzer, D. R. *J. Mol. Biol.* **1993**, *233*, 567.
26. Veldhoen, N.; You, Q.; Setzer, D. R.; Romaniuk, P. J. *Biochemistry* **1994**, *33*, 7568.
27. Zang, W.-Q.; Veldhoen, N.; Romaniuk, P. J. *Biochemistry* **1995**, *34*, 15545.
28. Foster, M. P.; Wuttke, D. S.; Radhakrishnan, I.; Case, D. A.; Gottesfeld, J. M.; Wright, P. E. *Nature Struct. Biol.* **1997**, *4*, 605.
29. Wuttke, D. S.; Foster, M. P.; Case, D. A.; Gottesfeld, J. M.; Wright, P. E. *J. Mol. Biol.* **1997**, *273*, 183.
30. Neely, L.; Trauger, J. W.; Baird, E. E.; Daevan, P. B.; Gottesfeld, J. M. *J. Mol. Biol.* **1997**, *274*, 439.
31. Suzuki, M.; Gerstein, M.; Yagi, N. *Nucleic Acids Res.* **1994**, *22*, 3397.
32. Klug, A.; Schwabe, J. W. R. *FASEB J.* **1995**, *9*, 597.
33. Choo, Y.; Klug, A. *Curr. Opin. Struct. Biol.* **1997**, *7*, 117.
34. Nardelli, J.; Gibson, T. J.; Vesque, C.; Charnay, P. *Nature (London)* **1991**, *349*, 175.
35. Jamieson, A. C.; Wang, H.; Kim, S.-H. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12834.
36. Elrod-Erickson, M.; Pabo, C. O. *J. Biol. Chem.* **1999**, *274*, 19281.
37. Nagaoka, M.; Sugiura, Y. *Biochemistry* **1996**, *35*, 8761.
38. Lilley, D. M. *Methods Enzymol.* **1992**, *212*, 133.
39. Price, M. A.; Tullius, T. D. *Methods Enzymol.* **1992**, *212*, 194.
40. Imanishi, M.; Hori, Y.; Nagaoka, M.; Sugiura, Y. *Biochemistry* **2000**, *39*, 4383.
41. Pavletich, N. P.; Pabo, C. O. *Science* **1993**, *261*, 1701.
42. Fairall, L.; Schwabe, J. W. R.; Chapman, L.; Finch, J. T.; Rhodes, D. *Nature (London)* **1993**, *366*, 483.
43. Houbaviy, H. B.; Usheva, A.; Shenk, T.; Burley, S. K. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13577.
44. Shi, Y.; Berg, J. M. *Biochemistry* **1996**, *35*, 3845.
45. Mandel-Gutfreund, Y.; Schueler, O.; Margalit, H. *J. Mol. Biol.* **1995**, *253*, 370.
46. Desjarlais, J. R.; Berg, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 2256.
47. Brüschweiler, R.; Liao, X.; Wright, P. E. *Science* **1995**, *268*, 886.
48. Ryan, R. F.; Darby, M. K. *Nucleic Acids Res.* **1999**, *26*, 703.
49. Ghost, G.; van Duyne, G.; Ghost, S.; Sigler, P. B. *Nature (London)* **1995**, *373*, 303.
50. Antson, A. A.; Burns, J. E.; Moroz, O. V.; Scott, D. J.; Sanders, C. M.; Bronstein, I. B.; Dodson, G. G.; Wilson, K. S.; Maitland, N. J. *Nature (London)* **2000**, *403*, 805.
51. Lewis, M.; Chang, G.; Horton, N. C.; Kercher, M. A.; Pace, H. C.; Schumacher, M. A.; Brennan, R. G.; Lu, P. *Science* **1996**, *271*, 1247.
52. Yean, D.; Gralla, J. *Nucleic Acids Res.* **1996**, *24*, 2723.
53. Narayan, S.; Wilson, S. H. *Biochemistry* **2000**, *39*, 818.
54. Hoey, T.; Gill, G.; Chen, J. L.; Dynlacht, B. D.; Tjian, R. *Cell* **1993**, *72*, 247.
55. Gill, G.; Pascal, E.; Tseng, Z. H.; Tjian, R. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 192.
56. Courey, A. J.; Tjian, R. *Cell* **1988**, *55*, 887.
57. Mastrangelo, I. A.; Courey, A. J.; Wall, J. S.; Jackson, S. P.; Hough, P. V. C. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 5670.
58. Su, W.; Jackson, S.; Tjian, R.; Echols, H. *Genes Dev* **1991**, *5*, 820.
59. Desjarlais, J. R.; Berg, J. M. *Proteins* **1992**, *12*, 101.
60. Wefald, F. C.; Devlin, B. H.; Williams, R. S. *Nature (London)* **1990**, *344*, 260.
61. DNase I footprinting was carried out according to our previous method.[63] Sp1(zf12)A556R was bound to 5′-GTGGAT-3′, the 5′-flanking site independent from the putative cognate site (data not shown). Sp1(zf12)A556R-F2-AT1 complex gave the single shift band under these experimental conditions, showing one complex form. Thus, the dissociation constant ($K_d^{out} = 220$ nM) of Sp1(zf12)A556R for this outer site was obtained. The fraction of the labeled DNA bound to two Sp1(zf12)A556R is represented by the equation $\Theta = I_{b2}/(I_f + I_{b1} + I_{b2})$, where I_{b1} and I_{b2} are the intensities of the faster and later migrating DNA-bound bands, respectively. The dissociation constant (K_d) of Sp1(zf12)A556R was evaluated by fitting $\Theta = [\text{peptide}]^2/(K_d^{out}\cdot K_d + (K_d^{out} + K_d)\cdot[\text{peptide}] + [-K_d^{out}\cdot K_d + (K_d^{out} + K_d)\cdot[\text{peptide}] + [\text{peptide}]^2)$.
62. Dixon, W. J.; Hayes, J. J.; Levin, J. R.; Weidner, M. F.; Dombroski, B. A.; Tullius, T. D. *Methods Enzymol.* **1991**, *208*, 380.
63. Kamiuchi, T.; Abe, E.; Imanishi, M.; Kaji, T.; Nagaoka, M.; Sugiura, Y. *Biochemistry* **1998**, *37*, 13827.