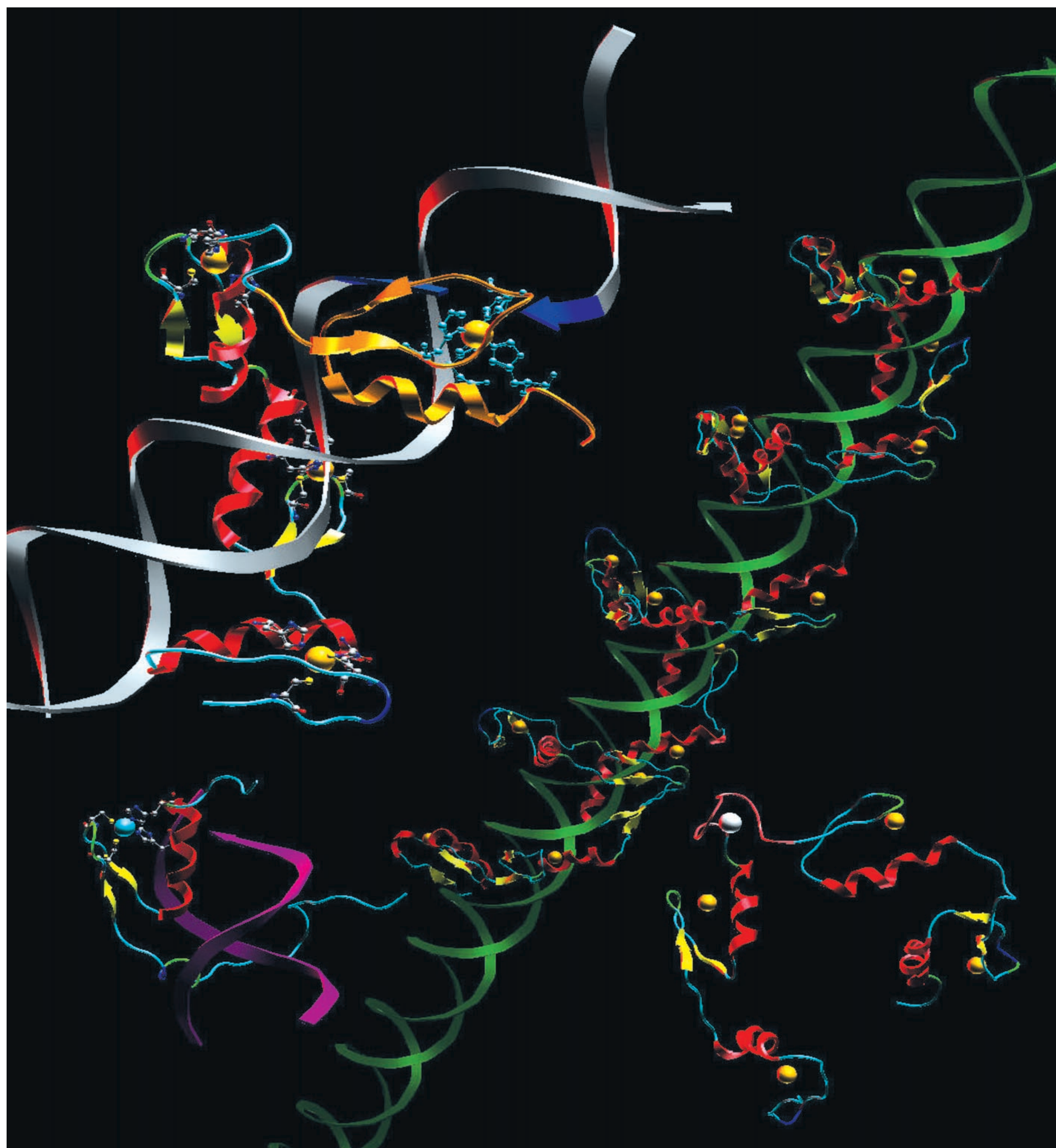
CHEMISTRY
A EUROPEAN JOURNAL

DOI: 10.1002/chem.200701320

New Redesigned Zinc-Finger Proteins: Design Strategy and Its Application

Shigeru Negi,*[a] Miki Imanishi,[b] Makoto Matsumoto,[a] and Yukio Sugiura*[a]



Abstract: The design of DNA-binding proteins for the specific control of the gene expression is one of the big challenges for several research laboratories in the post-genomic era. An artificial transcription factor with the desired DNA binding specificity could work as a powerful tool and drug to regulate the target gene. The zinc-finger proteins, which typically contain many fingers linked in a tandem fashion, are some of the most intensively studied DNA-binding proteins. In particular, the Cys_2His_2-type zinc finger is one of the most common DNA-binding motifs in eukaryotes. A simple mode of DNA recognition by the Cys_2His_2-type zinc-finger domain provides an ideal framework for designing proteins with new functions. Our laboratory has utilized several design strategies to create new zinc-finger peptides/proteins by redesigning the Cys_2His_2-type zinc-finger motif. This review focuses on the aspects of design strategies, mainly from our recent results, for the creation of artificial zinc-finger proteins, and discusses the possible application of zinc-finger technology for gene regulation and gene therapy.

Keywords: gene technology · protein design · protein modifications · zinc fingers

Introduction

From the human genome sequencing project, it has become apparent that the zinc-finger proteins constitute 2–3% of the entire human genome, and are the most common DNA-binding motifs found in human transcription factors.[1] The zinc-finger proteins are related to a wide range of biological functions such as development, differentiation, and suppression of tumors.[2] Each zinc-finger domain consists of twenty to thirty amino acid residues that have a special secondary structure stabilized by zinc ion bound to the Cys (C) and His (H) residues of the finger. The zinc fingers have been divided into several classes according to the number and type of amino acids involved in the Zn^{II} coordination, such as C_2H_2, C_2HC, C_4 ribbon, C_4 GATA family, C_6, C_8, C_3HC_4 ring finger, and H_2C_2.[3–6] Of these, the classical C_2H_2-type zinc fingers are the most common DNA-binding motif and often described as $CX_{2-4}CX_{12}HX_{2-6}H$, which represents the distances between the zinc-binding residues. This motif was first identified in the transcription factor TFIIIA derived from *Xenopus laevis* and also found in the transcription factors related to the RNA pol I.[7,8] The classical C_2H_2-type zinc fingers fold into a ββα compact structure with the aid of the Zn^{II} ion. The Zn^{II} ion pins together the β-hairpin in the N-terminal sequence and α-helix in the C-terminal sequence through a tetrahedral coordination with the two conserved cysteines located in the β-hairpin and two histidine residues located in the α-helix as shown in Figure 1 (top).[7,9]

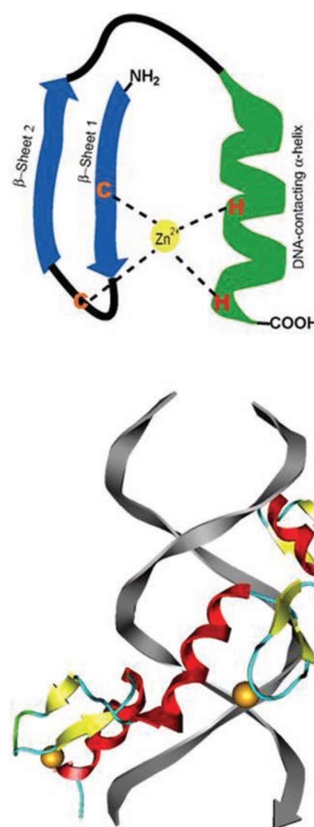


Figure 1. Top: Schematic representation of typical C_2H_2-type zinc-finger motif displaying ββα fold. Bottom: X-ray structure of the Zif268-DNA complex.[12]

The Zn^{II} ion is an essential structural factor for the formation of the ββα secondary structure of the zinc-finger domain, but is not directly involved in the DNA-binding activity. This folding structure is also stabilized by hydrophobic interactions.[9–11] The X-ray crystal structure of the C_2H_2-type zinc-finger protein Zif268–DNA complex clearly showed that each finger can typically interact with the major groove of DNA, and the zinc-finger proteins are arranged around the DNA strand in such a way that the α-helix of each finger contacts the DNA, forming an almost continuous stretch of α-helices around the DNA (Figure 1,

[a] Prof. S. Negi, M. Matsumoto, Prof. Y. Sugiura
Faculty of Pharmaceutical Sciences
Doshisha Women's University
Koudo, Kyotanabe-Shi, 610-0395 (Japan)
Fax: (+81) 774-658652
E-mail: snegi@dwc.doshisha.ac.jp
ysugiura@dwc.doshisha.ac.jp
[b] Prof. M. Imanishi
Institute for Chemical Research, Kyoto University
Uji 611-0011 (Japan)

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

bottom).[12] The characteristic DNA-binding mode of C_2H_2-type zinc-finger proteins can be described as follows:

1) One finger domain typically recognizes 3 or 4 contiguous base pairs (bp) of the DNA sequence. The specific three amino acid residues at positions −1, 3, and 6 on the α-helix (called the "recognition helix") are responsible for the contact with the contiguous DNA bases. In addition, the amino acid residue at position 2 is also involved in the contact with a DNA base on the other strand of the DNA.
2) The C_2H_2-type zinc-finger proteins most commonly contain the fingers (from 2 to more than 30 fingers) as connected in tandem by the short consensus linker, that is, the TGEK(R)P sequence,[12,13] except for the GAGA factor, which only contains a single C_2H_2-type zinc-finger domain.[14] Two or three tandem zinc fingers are necessary and sufficient for the specific-DNA binding without participation of any other domains.
3) The zinc-finger domain can bind to the cognate DNA as a monomer. Therefore, they can recognize the sequence of non-palindromic sequences unlike other nucleic acid recognition motifs, such as the basic leucine zipper and prokaryotic helix-turn-helix.[15]

Thus the C_2H_2-type zinc-finger domains with unique structure and molecular recognition abilities provide an attractive framework for the design of the artificial zinc-finger proteins with new DNA-binding properties and functions by mixing, matching, and linking zinc fingers and/or fusing the zinc-finger domain with various kinds of effecter domains. Over the past few years, our laboratory has been creating new DNA-binding proteins by redesigning the natural C_2H_2-type zinc-finger domain of the human transcription factor Sp1, which is the most intensely studied.[16] We now describe our design strategies to create artificial zinc-finger proteins and review the recent trends in zinc-finger technology for biomedical applications.

Fundamental Structural Analysis of Zinc-Finger Peptides toward Protein Redesign

Metalloenzymes are an important member of the protein family that can catalyze some of the most difficult reactions with the fine control of the reactivity and selectivity.[17–19] The design of metalloproteins can significantly contribute to our understanding of the fundamental principles in chemistry and biology, and provide useful information for biotechnological and biomedical applications. While much progress has been made in the study of native metalloproteins,[20–22] little is known about how to design a metalloprotein with the desired structure and activity. This problem has been approached through de novo protein design[23–28] as well as through the redesign of naturally occurring proteins.[29–31] The de novo design of a large protein is still a big challenge, because of our limited fundamental understanding of the relationship between the structures and functions in large proteins. However, the increasing success in the design of peptides and small proteins with metal sites gives great promise for the future.[23,32,33] In contrast to the de novo design, the redesign of naturally occurring proteins has an advantage in starting with a stable folded protein. Of the various approaches studied to mimic the catalytic properties of an enzyme, the catalytic antibody is one of the successful accomplishments.[34,35]

We firmly believe that the C_2H_2-type zinc-finger domains are a promising framework for the redesign and creation of artificial metal-binding proteins, because they have a unique structure and accurate molecular recognition abilities (metal binding and DNA recognition). From such a viewpoint, we have carried out several kinds of structural and functional characterizations of the zinc-finger proteins.

The human transcription factor Sp1 is a sequence-specific DNA binding protein isolated from HeLa cells.[36–39] Sp1 has three tandem repeats of a C_2H_2-type zinc-finger motif and specifically binds to the GC box DNA (Figure 2). Although the DNA binding mode of the three zinc fingers of Sp1 was predicted to be similar to that of Zif268, this model could not completely explain the DNA binding property of the N-terminal zinc finger. To investigate the DNA recognition mode of Sp1, a detailed analysis of the recognition mechanism with the finger-deleted mutants of Sp1 was performed.[40] Our electrophoretic analyses and alanine-scanning mutagenesis in the α-helix of finger 1 demonstrated

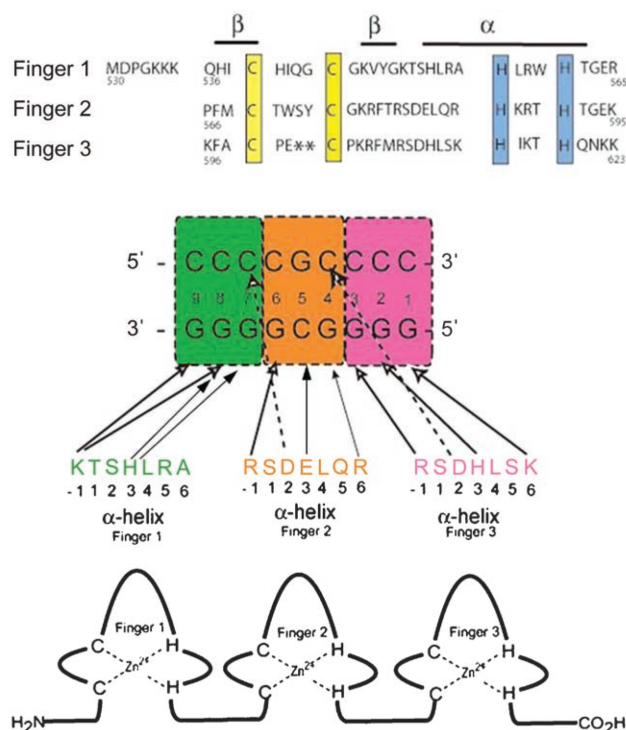


Figure 2. Top: Primary sequence of Sp1(530–623). The number indicates the positions of the amino acids in native Sp1. The zinc-coordinating amino acids (cystein (C) and histidine (H)) are highlighted. Bottom: Putative base recognition mode of Sp1(530–623).[40,41]

© 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

that all of the zinc-finger motifs 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 sequence selectivity of Sp1 is lower than that of fingers 2 or 3, indicating that the manner of interaction between finger 1 and DNA is different from the Zif268-DNA complex model. Recently, we determined the solution structure of the DNA-binding domain of the Sp1 by solution NMR techniques.[41] The NMR results of the Sp1–DNA complex supported our above-mentioned findings. The NMR structural analysis also showed that the DNA recognition mode of fingers 2 and 3 is similar to those of Zif268, in which each finger can recognize three or four base pairs strictly by using residues at positions −1, 2, 3, and 6 in the recognition helix. On the other hand, finger 1 can use only two residues at positions −1 and 3 in the helix region for DNA recognition, and has a more relaxed sequence and site specificity than the other C_2H_2-type zinc fingers.

Next, we produced the four mutant peptides in which Gly or Ala was substituted for one of the four zinc ligands corresponding to the second zinc-finger domain of Sp1, because little is known about the contribution of the individual zinc ligands and the relationship between the zinc binding and protein folding in the zinc-finger domains.[42] Their metal binding and folding properties were characterized by a series of spectroscopic methods and molecular dynamics simulations (Insight II software). This work revealed that even the zinc-finger peptide mutants with only three metal coordination residues can bind metal ion, and that His and the hydrophobic core formed between the α-helix and β-sheet play an essential role in the α-helix induction. Of special interest is that each ligand does not equally contribute to the α-helix formation and coordination geometry in the zinc-finger peptide. Moreover, this study sheds light on the possibility that small metallopeptides with the desired coordination geometry can be synthesized by selecting the appropriate amino acid sequence. As described later, indeed, we created the first catalytic zinc finger with a hydrolytic-ability-like nuclease by redesigning the zinc-binding site in the finger domain.

We designed an artificial protein, “Antenafinger” (Ant-F),[43] and investigated its physicochemical properties using thermodynamic measurements.[44] Ant-F was produced by introducing the consensus sequence of the C_2H_2-type zinc fingers into a non-metalloprotein scaffold. The important process is how to find a scaffold protein the structure of which is well-defined and is not lost by introducing the consensus residues among the enormous natural proteins found in the PDB database. The bioinformatic tool (GenomeNet motif) was utilized to obtain the desired scaffold protein. Based on the bioinformatics analysis, we finally selected the suitable scaffold protein, the Antennapedia homeodomain mutant (Ant-wt), which is a DNA-binding protein involved in the development control of *Drosophila* and consists of four helical regions. The reason why this protein was chosen is that it is small (54 amino acid residues) and does not contain any metal ion or disulfide bond. Therefore, the Ant-F was constructed from Ant-wt with only four point mutations to two Cys and two His residues corresponding to the consensus sequence of the C_2H_2-type zinc fingers. This protein was easily synthesized by a simple Fmoc solid-phase technique. Ant-F made its structural change by forming a tetrahedral metal-binding site similar to those of the C_2H_2-type zinc fingers by adding Zn^{II}, while Ant-F gave secondary structures similar to that of Ant-wt in the absence of Zn^{II} (Figure 3). The CD spectral analysis clearly indicated that

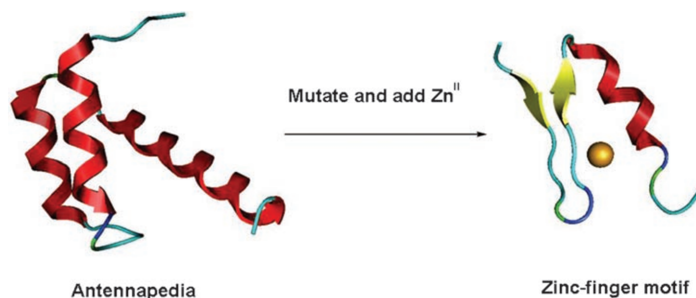


Figure 3. Conversion of antennapedia homedomain to zinc-finger-like domain by Zn^{II}-triggered structural alteration.

Zn^{II}-binding induces a structural change in Ant-F and decreases the α-helical content. In addition, Ant-F regulated its DNA binding activity through the Zn^{II}-triggered structural alteration. These results provide valuable knowledge about protein dynamics and a new concept for metalloprotein design.

Recently, we found that the DNA-binding ability of the GAGA zinc finger depends on the nature of the amino acid in the β-hairpin.[45] The GAGA factor of *Drosophila melanogaster* has a unique structure, namely specific binding to the target DNA using only a single finger domain (Figure 4).[14] The other naturally occurring zinc-finger proteins require more than two zinc-finger units linked in a tandem fashion for specific DNA recognition. Thus, the design of a DNA-binding functional protein with the GAGA zinc finger is not limited to the tandem assembly of multiple zinc fingers. The minimal DNA-binding domain of the GAGA transcription factor (GAGA-DBD) is 63 residues in length and specifically binds to the 5′-GAGAGAG −3′ sequence with a dissociation constant of 5 nM. Omichin-

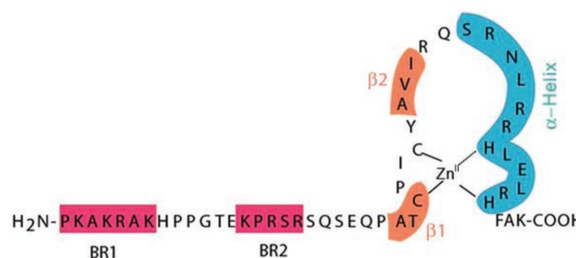


Figure 4. The primary peptide sequence of designed minimal GAGA DNA-binding domain showing different secondary structure elements and the two basic regions.

© 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

ski et al. investigated the solution structure of a specific GAGA-factor/DNA complex by an NMR structural analysis, indicating that the complex structure as well as the DNA-binding mode is very similar to other classical C_2H_2-type zinc fingers.[46] However, there is a striking difference in the amino acid preference located in the β-hairpin region, when compared to other structurally characterized zinc fingers. The sequence alignment of the zinc fingers reveals that the amino acid preference in the short β-hairpin of GAGA-DBD is different from the other zinc-finger proteins. To clarify its sequence difference and the substitution effect of aromatic amino acids in the β-hairpin on the DNA binding, we studied the folding and DNA-binding properties of three GAGA-DBD mutants designed by substitution of the consensus phenylalanine, an aromatic amino acid, at the key position in the β-hairpin. The metal binding and the overall fold of the mutant peptides are very similar to those of the wild-type as shown by the UV and CD spectroscopy. However, the gel-mobility shift assay and isothermal calorimetric studies revealed that substitution of the consensus phenylalanine abolished the specific DNA-binding activity of all the GAGA-DBD mutants. Although it is still difficult to exactly interpret the reason why the mutants showed no DNA-binding, it is likely that they formed a slightly different conformation from that of the native GAGA peptide. To obtain additional structural information about the GAGA-DBD mutants, we utilized the energy minimization calculation with the standard protocol available in the MOE program. The resultant energy-minimized structures indicated that the substitution of Phe at either of the β-hairpin relaxed the compactness of the minimal hydrophobic core necessary for the compact ββα fold of the zinc-finger domain (Figure 5). Presumably, the loss of the DNA-binding ability of the GAGA mutants is correlated to the changes in the hydrophobic core, which eventually affects the DNA-binding surface. This result suggests that the nature of the amino acids in the β-hairpin region plays an important role in the DNA-binding of the GAGA factor protein. The study also provides useful information about the structure-based artificial zinc-finger design.

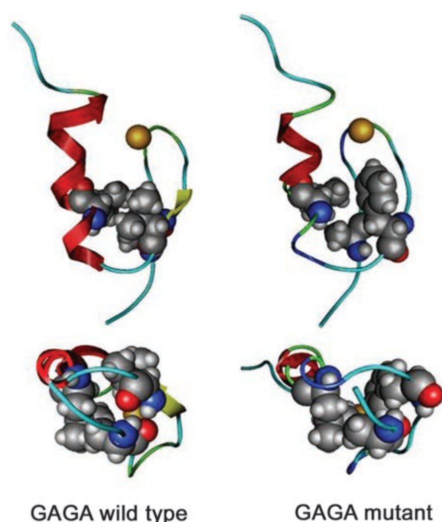


Figure 5. Energy-minimized structures of GAGA wild-type and mutant. The picture was produced on the basis of solution NMR structure of GAGA/DNA complex (PDB code: 1YUJ) using MOE program (Chemical Computing Group, Montreal).

Manipulation of Recognition Code of Zinc-Finger Proteins by Swapping the Secondary Elements

As mentioned above, the α-helix in zinc-finger domain bind to the cognate DNA duplex in the major groove (major groove binding), and the side chains at positions −1, 2, 3, and 6 selectively contact the three or four successive bases. From this viewpoint, it is possible to change the recognition bases by mutating the amino acid residues involved in the DNA recognition. In fact, libraries of zinc-finger variants have been made from designed strategies and selection methods.[47–59] The phage display is one of the most popular selection methods that has been used to produce the libraries of zinc-finger variants. In this method, the key DNA-contacting amino acid residues on the α-helix are randomized, while all other regions of the structure are left unchanged. A sequential selection strategy was developed by Pabo's group.[54,58] Barbas's group also used a similar phage display method to obtain the new sequence specific zinc-finger proteins and succeeded in creating the zinc-finger domains that specifically recognize many of the 64 triplets (ca. 75% of all triplets combinations).[55–57,60] They demonstrated that the selected and optimized finger domains are assembled in tandem to form three- or six-zinc-finger proteins that can bind to the contiguous DNA triplet with a sophisticated specificity and high affinity.

As another approach, a "nondegenerate recognition code table" is a useful method for the rational design of artificial zinc-finger proteins. It became apparent that certain amino acids in DNA-recognition positions exhibit preferences for certain DNA bases by collecting a lots of information concerning DNA-binding properties by either natural or engineered zinc fingers. Zinc-finger proteins have been created purely by rational design based on recognition code and used in the mammalian system,[61,62] the plant reporter system, and plant viruses.[63]

On the other hand, the secondary structure-swapping method is also useful for the designing of zinc-finger proteins with new DNA-binding specificities. Recently, we succeeded in changing the DNA-recognition pattern and DNA-binding affinities of the zinc-finger proteins by α-helix or β-sheet exchanges between different fingers. For the α-helix swapping experiment, we utilized two different kinds of zinc-finger proteins, the Sp1 and CF2-II zinc-finger proteins. The Sp1 zinc-finger protein has three C_2H_2-type zinc-finger domains (Sp1(zf123)) and binds to the GC box, 5′-GGG GCG GGGC-3′.[36–39] On the other hand, three of the six C-terminal C_2H_2-type zinc fingers (finger 4–6) of the *Drosophila* transcription factor CF2-II can bind to the AT-rich sequence, 5′-GATTATATA-3′.[64] A new zinc-finger pep-

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

tide, Sp1HM, was created by α-helix exchange between the Sp1 and CF2-II zinc fingers.[65] In the chimeric Sp1HM, the original GC specific α-helices of the Sp1 zinc-finger protein were exchanged for the AT specific α-helices of CF2-II as shown in Figure 6.[65] The folding property of Sp1HM was investigated by CD measurement. By adding Zn^{II}, Sp1HM induced negative Cotton effects around 206 and 222 nm, indicating that peptide has an ordered secondary structure, namely α-helices. The DNA-binding ability of Sp1HM was examined by gel-mobility shift assays for the GC-rich sequence (5′-GGG GCG GGGC-3′) and AT-rich sequence (5′-GAT TAT ATA-3′), and then the dissociation constants for each cognate DNA were determined. Sp1HM binds to the AT-rich sequence with 3.2 and 1330 nM (K_d values) in the absence and presence of the competitor DNA (calf thymus DNA), respectively. On the other hand, the K_d values of Sp1(zf123) for the GC-rich sequence in the absence and presence of the competitor DNA were 4.0 and 46.1 nM, respectively. Interestingly, no evident binding complex for Sp1HM-GC and Sp1(zf123)-AT was detected under our experimental conditions. These results indicate that Sp1HM binds to the AT-rich sequence with a high affinity and specificity.

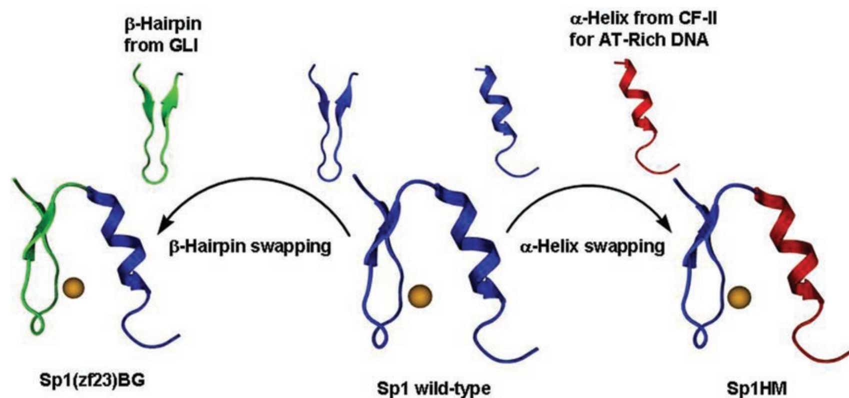


Figure 6. Swapping of the α-helix and β-hairpin regions alters the DNA binding properties of Sp1.

Recent our work has revealed that swapping of the β-hairpin between two fingers can alter the DNA-binding properties of the zinc-finger proteins.[66] In general, the β-hairpin regions play only a structural role in the protein folding and do not engage in the DNA recognition, because the DNA-contact amino acid residues are located in the α-helix regions. Therefore, the recent design strategies of the zinc-finger proteins have mainly focused on the α-helix region. By swapping the β-hairpin regions between the Sp1 and GLI zinc fingers, we investigated the role of the β-hairpin regions of the C_2H_2-type zinc-finger peptides on the DNA-binding events.[66] The Sp1 and GLI zinc fingers were selected because each zinc finger shows distinct base and phosphate interactions, and their structural and functional properties have been studied.[40,41,67] Sp1(zf23)BG and GLI-(zf45)BS were created by exchanging the β-hairpin regions between Sp1(zf23) and GLI(zf45) in Figure 6. Both engineered Sp1(zf23)BG and GLI(zf45)BS showed the typical CD feature similar to that of the C_2H_2-type zinc-finger peptides. However, the value at 222 nm of GLI(zf45)BS ($[\theta]_{222} = -9403\,^{\circ}\,cm^2\,dmol^{-1}$) was greater than that of Sp1-(zf23)BG ($[\theta]_{222} = -3029\,^{\circ}\,cm^2\,dmol^{-1}$), suggesting a structural difference in the histidine–histidine spacing.[67] To evaluate the effect of the β-hairpin region on the DNA-binding affinity, we performed the gel-mobility shift assays. It was surprising to find that the Sp1 mutant, Sp1(zf23)BG, with the GLI β-hairpin had a higher DNA-binding affinity for the GC box than that of the wild-type Sp1. On the other hand, the GLI mutant with the Sp1 β-hairpin, GLI(zf45)BS, completely lost its DNA-binding ability for targeting the GLI sequence (5′-TTGGG T GGTC-3′) despite the existence of the DNA-recognizing helical region for the GLI sequence. To determine the DNA-binding mode of these mutants, a DNase I footprinting analysis was carried out. From this analysis, the DNA-binding mode of the Sp1-type mutants is evidently affected by the substitution of the β-hairpin regions, and the β-hairpin region appears to participate in the DNA binding of the non-Zif268-type zinc fingers such as the GLI zinc finger.

These results imply that our rational design strategy is also effective and convenient for creating zinc-finger proteins with a high affinity and/or novel-binding specificities comparable to those of a zinc finger selected by the phage display technique.

Creation of Multiple Zinc-Finger Proteins

Long-sequence-recognizing fingers: Considering the size of the human genome (3×10^9 bp), it is necessary for the sequence selective DNA binding protein to recognize more than 16 bp ($4^{16} = 4\times10^9$) in order to target a unique site. In general, recognition of a long DNA sequence is a rare function in the natural DNA-binding proteins. For example, the naturally occurring nine-zinc-finger protein TFIIIA binds to a shorter DNA sequence, because it uses only a few selected fingers for DNA binding.[13,69] Since one finger unit can typically recognizes the contiguous three or four base pairs of a DNA sequence, it would be possible, in principle, to create artificial zinc-finger proteins that can recognize more than 16 bp by connecting several independent zinc-finger domains.[70,71] There is much interest in engineering sequence-specific binding proteins that bind to the longer regions of DNA, because such polydactyl zinc-finger proteins should

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

be broadly applicable as genomic-specific transcription switches in gene therapy strategies and the development of new transgenic plants and animals.[72–75] As described above, the zinc-finger motifs are normally connected by the well-conserved TGEK(R)P five amino acid linker (Krüppel-type linker). This linker functions in controlling the orientation and spacing of the adjacent fingers.[76] To examine the possibility of artificial zinc-finger proteins to recognize more than 16 bp, we have created new six-, nine-, and fifteen-finger peptides (Sp1ZF6, Sp1ZF9, and Sp1ZF15) which have two, three, and five Sp1(zf123)s, respectively, connected by the consensus linker[71,77] (Figure 7). Of special interest is the fact that the resulting multiple fingers can bind to the expected contiguous long DNA sequence, 2GC (18 bp), 3GC (27 bp), and 5GC (45 bp), respectively. This result clearly shows that the length of the cognate DNA binding sequence increases in proportion to the number of zinc-finger motifs. As for the binding affinity, the multiple zinc-finger proteins do not show much cooperativity when compared to the original three-zinc-finger unit. The artificial transcription factors based on these multi-zinc-finger proteins have great potential for the genomic-specific transcription manipulation and the regulation of a vast number of cellular processes.

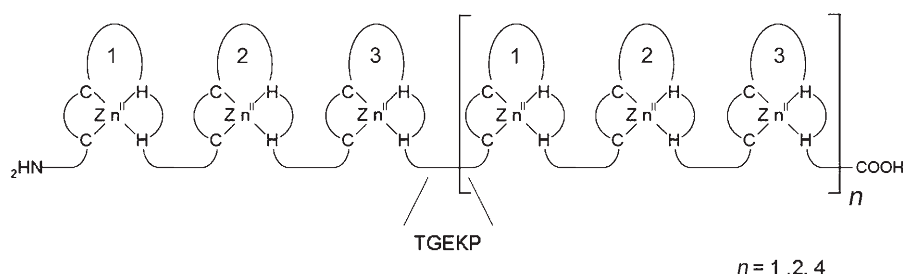


Figure 7. A schematic representation of multiple zinc-finger proteins based on Sp1(zf123). Sp1ZF6, Sp1ZF9 and Sp1ZF15 have two, three, and five Sp1(zf123)s connected by the consensus linker (TGEKP), respectively.

Artificial zinc-finger proteins with expanded and functional linkers: After the design of multi-zinc-finger proteins that possess a long-DNA-sequence recognition ability, we tried to create a new DNA bending finger to regulate the gene expression. Because DNA structural changes, such as bending, are some of the most important factors for gene regulation, the artificial zinc-finger proteins that induce DNA conformation changes are interesting as a transcriptional regulator of a specific gene. We created the six-zinc-finger proteins, Sp1ZF6(Gly)$_7$ and Sp1ZF6-(Gly)$_{10}$, by connecting two three-zinc-finger units of Sp1 with flexible poly-glycine linkers[78] (Figure 8). As the target DNA, two GC-box sequences separated by one helical turn of DNA (10 bp) were used. The phasing assays clearly showed that both Sp1ZF6(Gly)$_7$ and Sp1ZF6(Gly)$_{10}$ can induce DNA bending at the intervening region between two distal binding sites. In particular, Sp1ZF6(Gly)$_7$ caused the most significant directional change in the DNA bending, indicating that the linker length between two three-finger units has a crucial effect on the entire DNA bending direction. To investigate the effect of the linker type on the DNA-binding properties in more detail, we designed a six-zinc finger with charged linkers, Sp1ZF6(GE)$_4$ and Sp1ZF6(GR)$_4$[79] (Figure 8). These proteins are also able to cause DNA structural changes similar to the Sp1ZF6(Gly)$_{10}$ protein. However, on the basis of the surface plasmon resonance experiment, the kinetic aspects of their DNA binding were different in each case. In particular, the dissociation rate of Sp1ZF6(Gly)$_{10}$ was the fastest and that of Sp1ZF(GR)$_4$ was the slowest. The kinetic stability of the bent DNA–protein complexes is very important for the assembly of many molecules in various biological systems. These results indicate that the length and net charge of the linker region between two three-zinc-finger domain units affects the DNA structural changes and that the property of the linker regions produces different DNA-binding kinetics.

Although many artificial zinc fingers, including our results, have been created to recognize a long contiguous DNA sequence, the recognition for a discontiguous DNA target has never been determined because of the limitation of the zinc-finger framework. Therefore, binding to a discontiguous target with a non-recognized DNA sequence is meaningful. Pabo et al. synthesized dimeric zinc-finger peptides and succeeded in recognizing a discontiguous target

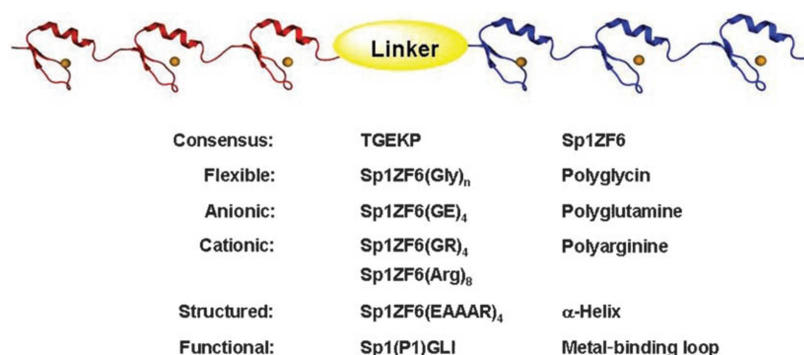


Figure 8. Artificial six-zinc-finger proteins with various linkers.

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

DNA sequence.[80,81] We have also used Sp1ZF6(Gly)$_{10}$ with a flexible long glycine linker in similar investigations;[82] Sp1ZF6(Gly)$_{10}$ can bind to the discontiguous GC-box sequence separated by 10 bp. However, it also binds to the contiguous GC-box sequence with almost the same affinity. To obtain the selectivity to discontiguous recognition sequence with a particular length of the non-recognized gap is challenging. As the next step, we tried to create a multi-zinc-finger protein that binds to the discontiguous target sequence and not to the contiguous ones. Eight arginine residues were introduced to connect the three-zinc-finger units of Sp1, and Sp1ZF6(Arg)$_8$ was created[82] (Figure 8). Because the side chain of an arginine residue is bulky and positively charged, we expected that the steric hindrance and/or electrostatic repulsion might inhibit DNA binding to the contiguous GC-box sequences. The gel-mobility shift assay and footprinting analysis clearly showed that Sp1ZF6(Arg)$_8$ specifically bound to the discontinuous GC-box by a 10 bp gap. In contrast, the binding affinity of Sp1ZF6(Arg)$_8$ to the contiguous GC-box site was extremely low compared to that to a discontiguous sequence. The bulky and cationic polyarginine linker seems to contribute to the selective binding to the discontiguous GC-box site.

Recently, we created a new six-zinc-finger protein, Sp1ZF6 (EAAAR)$_4$, in which two three-zinc-finger units are connected by an α-helical peptide, (EAAAR)$_4$.[83] The peptides (EAAAR)$_n$ ($n>2$) are known for forming an α-helix structure stabilized by salt bridges,[84] and the length of the (EAAAR)$_4$ helix is estimated to be about 30 Å, corresponding to that of the 10 bp DNA. We expected that a six-zinc-finger peptide with an α-helix linker would have DNA-binding preferences dependent on the relative positions of the discontigous recognition site, especially in the same phase (10 bp gap; 2GC(10)) rather than in the opposite phase (5 bp gap; 2GC(5)) of the DNA helix (Figure 9). The gel shift assays indicated that Sp1ZF(EAAAR)$_4$ preferably binds to the 10 bp-gapped target rather than the 5 bp-gapped one. The CD results clearly revealed that the α-helix content of the (EAAAR)$_4$ linker is higher in the complex with the 10 bp-gapped target than in the complex with the 5 bp-gapped one. Theses results provided basic and useful information about the linker design of new six-zinc-finger peptides that could possibly discriminate any two separated recognition sites.

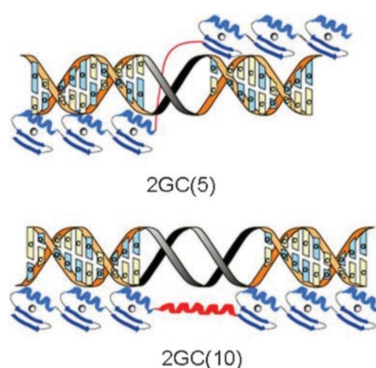


Figure 9. Schematic representations of Sp1ZF6(EAAAR)$_4$-DNA complexes. The 2GC(5) and the 2GC(10) sequences contain two discontinuous GC-box sequences with nonbound gaps of five and ten base pairs, respectively.

Moreover, we investigated the effects of the length and position of an extended linker on the sequence-selective DNA recognition of the zinc-finger peptide.[85] Hence, we constructed zinc-finger peptides with extended linkers, created by the insertion of spacer amino acid residues into the canonical linker. These results showed that insertion of three amino acids (GGS) is enough for recognition of a cognate DNA with a small gap, and that an extended linker is a promising tool for expansion of the natural recognition code of the zinc-finger proteins.

The transcription factor TFIIIA, as described above, contains nine contiguous zinc-finger motifs; six N-terminal fingers of TFIIIA bind to the 31 bp of the 5S rRNA gene, but finger 4 does not make any specific base contacts with DNA. TFIIIA has four types of non-consensus linker sequences, -NIKCV-, -TQQLP-, -AG-, and -QDL-. The unique linker sequence of TFIIIA seems to significantly affect its special DNA recognition mode. We tired to create a new nine zinc-finger proteins, Sp1ZF9T, by introducing the TFIIIA-type linker to Sp1ZF9.[86,87] Several electrophoresis analyses indicated that the DNA-binding mode of Sp1ZF9T is evidently different from that of Sp1ZF9, and that the position of the TFIIIA-type linker in Sp1ZF9 is also important for the DNA recognition control by multi-zinc-finger proteins.

Barbas's group has also reported that the regulation of two endogenous genes by using artificial zinc-finger proteins containing 12 fingers with expanded linker regions.[88] They tired to change the level of ErbB2 and ErbB3 receptors, which are members of the ErbB family of tyrosine kinase receptors in normal and cancer-derived cell line. They could successfully co-regulate ErbB2 and ErbB3 by a 12-finger artificial transcription factor and demonstrate the potential of designer zinc fingers to the regulation of multiple genes.

Furthermore, we examined the possibility to introduce the DNA cleavage function to the linker region for creation of a sequence specific multiple zinc-finger type endonuclease.[89] An analogue of the metal binding loop was used to connect the two zinc-finger domains derived from Sp1 and GLI. The loop is known to strongly bind to a lanthanide ion, and exhibits a DNA cleavage activity in the presence of cerium(IV) ions.[90–93] The zinc-finger-based artificial nuclease, Sp1(P1G)GLI, was demonstrated to cleave DNA at the region between two binding sites for Sp1 and GLI in the presence of Ce^{IV} (Figure 8). These results indicate that a new sequence specific endonuclease can be created by introducing a metal-binding loop as a functional linker of two zinc-finger domains.

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Mirror-Image Recognition in DNA/Zinc-Finger-Protein Interactions

Most biomolecules are "chiral" or handed, that is to say, they exist in left and right-handed mirror forms. However, nature only uses one hand, that is, it is "homochiral". One of the biggest puzzles in biophysics and biochemistry is the question of why life on earth is based on left-handed (L) amino acids and right-handed (D) sugars. It is now known that biochemical interactions of biological macromolecules are inherently chiral, and the chiral recognition of biomolecules plays a central role in various biological interactions.[94,95] Kent et al. experimentally demonstrated the mirror-image relationship in the enzyme–substrate system.[96] They chemically synthesized an enzyme enantiomer with a reciprocal chiral specificity on peptide substrates; the L-enzyme cleaved only the L-substrate, whereas the D-enzyme cleaved only the corresponding D-substrate.

The biomolecular chirality can probably be applied in biotechnology and biomedicinal chemistry. A recently developed method, the so-called "mirror-image phage display", allows the use of the phage display to ultimately identify peptides that specifically bind to a given target and consist solely of D-amino acids.[97–102] For example, Willbold et al. applied the mirror-image phage display to identify D-peptides that specifically bind to the amyloid peptide Aβ(1–42), and showed its medicinal application.[101] Aβ is the major component of the amyloid plaques found in the brains of people suffering from Alzheimer's disease (AD). Upon screening, they found a suitable peptide (D-pep) which can bind Aβ(1–42) with high afiinity and selectivity. In addition, amyloid plaques were specifically stained with a fluorescence-labeled derivative of D-pep. These results imply that D-pep may be a very suitable probe for the detection of amyloid plaques in humans and AD therapy.

In an attempt to experimentally demonstrate the mirror-image recognition in transcription proteins and DNA interaction, we recently designed a small DNA-binding peptide based on the zinc-finger domain of the GAGA transcription factor.[103] The zinc-finger domain is attractive for studying mirror-image protein-DNA recognition because of its simplicity, DNA-binding ability, and molecular assembly.[2] Both the L- and D-GAGA peptides and DNAs were synthesized by using the Fmoc-standard protocol (Figure 10). CD measurements suggested that the conformation of the GAGA peptide enantiomers as well as the DNA enantiomers have a completely mirror-image relationship. The gel-mobility shift assay indicated that the synthetic enantiomers of the peptide show reciprocal chiral-specific interactions with DNA; the natural L-peptide specifically binds with the natural D-DNA substrate, and the unnatural D-peptide specifically binds with the unnatural L-DNA substrate. These data provided valuable first experimental evidence for the mirror-image recognition in the protein–DNA interaction and generalizes the chiral nature of biomolecular interactions. Moreover, D-proteins have potentially practical applications; the protein enantiomer may be used in chiral catalysis for the production of both enantiomers of fine chemicals and therapeutic biomolecular reagents because of its resistance to naturally occurring proteases.

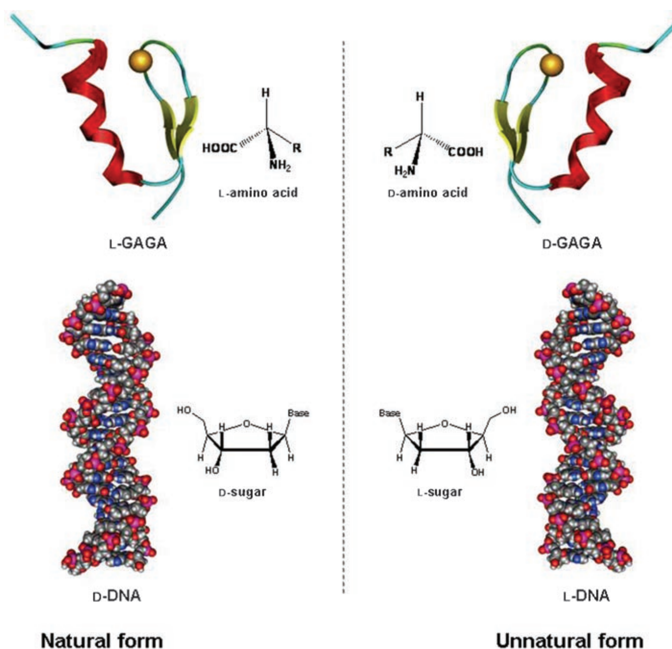


Figure 10. Schematic representations of structures of the D- and L-enantiomers of GAGA DNA binding domains.

Creation of New "Zinc Fingerzyme" with Specific Catalytic Ability

Zinc-finger proteins possess two molecular recognition abilities, that is, metal- and DNA-binding abilities, but no function to catalyze some kinds of biological reactions as an enzyme would. The redesign of the metal-binding site in the zinc-finger motifs can become one of the most powerful and effective strategies to introduce the catalytic function to the zinc-finger protein itself. The zinc ion in the zinc-finger motifs is a typical structural metal ion, only stabilizing the compact structure required for specific DNA recognition (Figure 11). On the other hand, a catalytic zinc ion directly participates in the catalytic reaction. In contrast to the structural zinc-ion sites, a unique feature of the catalytic zinc-ion site is the existence of vacant site(s), that is, the zinc coordination geometry contains at least one water molecule in addition to three or four protein ligands (Figure 11).[104,105] Actually, the zinc-bound water is a critical component for a catalytic site. Several rational design studies for metal(zinc)-binding sites have been reported.[106–111] To create a single finger as a new metallopeptide with catalytic zinc ion site, a series of zinc-finger peptides of 30 amino acids in length were designed on the basis of the second finger of Sp1, such as the CCHG- and GCHH-types, in which one zinc-coordinating residue is substituted into a noncoordinating one.[42,112] On the basis of a spectroscopic experiment with

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

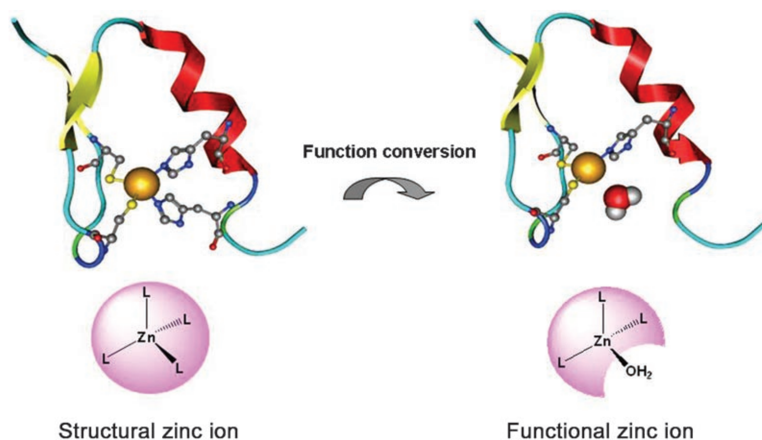
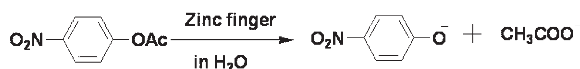


Figure 11. Function conversion from "structural zinc ion" to "functional zinc ion" by redesign of zinc ion binding site in zinc-finger domain.

proteins still retain the general ββα fold of the wild type in the presence of Zn^{II}.[115] The gel-mobility shift assay and DNase I footprinting analysis revealed the binding of Zn^{II}–H_4Sp1 to the GC-box site. In addition, an electronic spectral study of the Co^{II}–zf(HHHH) complex indicated a coordination geometry containing two vacant sites.[112] Along with the above-mentioned hydrolytic experiment, the hydrolysis abilities of His-substituted mutants were investigated. As expected, among the mutants, the zf(HHHH) showed a much better hydrolytic activity than the

the Co^{II}-substituted complex, the unsaturated zinc site seems to be occupied by one or two water molecules, and this information encourages us to examine the hydrolytic ability of the unsaturated sites in these zinc-finger mutants. The hydrolytic ability of the zinc-finger mutant was investigated by using 4-nitrophenyl acetate (NA) as the substrate (Scheme 1). All zinc complexes of the mutant peptides dis-



Scheme 1. Hydrolytic cleavage by engineered zinc fingers.

played an appreciable hydrolytic activity. The hydrolytic activity increased when the number of histidine residues was increased at the coordination site. In general, a zinc complex with a higher Lewis acidity exhibits a higher hydrolytic activity.[113]

Recently, Andreini et al. estimated the number of putative human zinc-binding proteins and classified all human zinc proteins according to their functions and zinc-binding patterns, based on bioinformatic analyses.[114] Nearly all (97 %) of the four-ligand metal-binding patterns (MBP) contain at least one Cys ligand, with 40 % of the MBPs containing four Cys, and 27 % of the CCHH-type, which correspond to the MBP of the most common human zinc-finger protein. Of special interest is the fact that any H_4-type zinc-binding protein (including zinc-finger proteins) has never been found in nature. To answer the question why such a protein does not exist in nature, we tried to create the first H_4-type zinc-finger protein (zf(HHHH) and H_4Sp1), engineered by cysteine to histidine mutations of the Sp1 zinc finger (Figure 12). We also expected that the H_4-type zinc-finger complex can exhibit a higher hydrolytic activity compared to other mutant peptides, because of its higher Lewis acidity. A conformational analysis by CD and 1H 2D-NMR spectroscopic measurements showed that the H_4-type zinc-finger

other mutants. The catalytic activity was evidently higher than that of the zinc–cyclen (cyclen = 1,4,7,10-tetraazacyclododecane) complex, which is one of the best artificial model complexes for hydrolytic zinc enzymes.[116] It is also interesting to note that this mutant catalyzes the hydrolysis of amino acid esters (Boc-glutamine 4-nitorophenyl ester, D- or L-Gln-ONp) with a high enantioselectivity.

The molecules that can hydrolyze the DNA phosphoester at a specific position are valuable tools in biotechnology, because of facilitating DNA manipulation in a variety of applications. Since the zinc-finger motifs are known to bind to the major groove of the DNA duplex in a sequence-specific fashion, we expected that the H_4-type zinc-finger proteins would function as a new type of nuclease with high sequence specificity for DNA duplexes. Therefore, we examined the DNA hydrolytic ability of the H_4-type zinc-finger proteins.[117] Both zf(HHHH) and H_4Sp1 showed a hydrolytic ability for the plasmid DNA. We also successfully demonstrated the selective hydrolysis of the DNA duplex at the target GC box by the tandem three-zinc-finger mutant, H_4Sp1. Interestingly, at a high ionic strength, the H_4Sp1

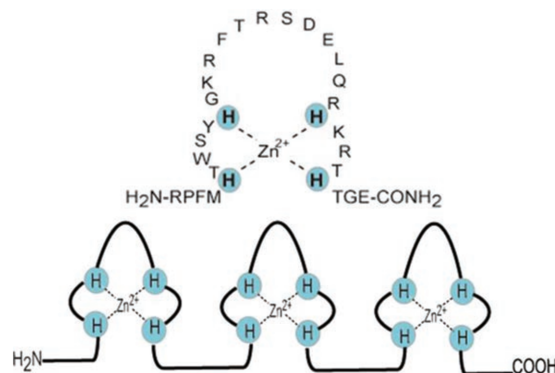


Figure 12. Artificial H_4-type zinc-finger proteins. Top: zf(HHHH) designed from the second finger of Sp1. Bottom: the three-tandem zinc-finger protein H_4Sp1.

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

mutant could convert the plasmid DNA from form I to III. This is the first example of the catalytic function of the zinc-finger protein itself by converting the function of the structural zinc ion into that of the catalytic zinc ion. These results clearly demonstrate that the redesign of a structural metal site in proteins is promising for the creation of new catalytically active metalloproteins. In addition, our results probably suggest the reason why an H_4-type zinc-finger protein was screened out in the natural and evolutional selection process.

Zinc-Finger Technology and Its Applications

Artificial gene-silencing methods by targeted DNA methylation are epigenetically very important, because nature can turn off transcription of a gene by adding a methyl group to the gene[118] and aberrant methylation (gene hypermethylation) is also related to the tumor cell process.[119,120] Several kinds of methyltransferases (M. SssI and M. HhaI) and their variants have been fused to zinc-finger domains and tested both in vitro and in vivo.[121–126] Li et al. demonstrated that targeted DNA methylation is effective in repressing the Herpes Simplex Virus type I infection in a cell culture.[125] These different groups have independently created the engineered zinc-finger proteins that inhibit replication of the integrated HIV-1[127,128] and high-risk human papillomavirus type 18.[129] Smith and Ford succeeded in the specific methylation of a chromosomal locus with an integrated zinc-finger targeting sequence, and they showed that the methylation pattern is inherited through successive cell divisions.[124] Barbas III and Nomura have designed a semisynthetic methyltransferase to mimic this behavior at a specific DNA site in bacteria.[126] In this system, DNA binding of artificial zinc-finger proteins brings together two fragments of the methyltransferase (M. Hhal), which then assemble into the active enzyme. They succeeded in transferring a methyl group to a cytosine in the region (CpG) between the two DNA binding sites using this artificial enzyme. In the near future, inherited gene-silencing of endogeneous genes would be possible by zinc-finger methyltransferases. With respect to the above-mentioned contents of the DNA methylation, Ghosh et al. reported a methodology termed the mCpG-sequence-enabled reassembly (mCpG-SEER) of proteins utilizing a split green fluorescent protein tethered to modular zinc-finger DNA-binding domains.[130] They clearly demonstrated the successful detection of the site specific CpG island methylation in dsDNA with their rationally designed mCpG-SEER system in vitro. They also mentioned that the mCpG-SEER system is a potentially useful method for the direct detection of the CpG methylation site and is applicable to the epigenome and cancer research.

Not only chromosomal DNA, but mitochondrial DNA could also be an important target of artificial zinc-finger proteins. In general, natural zinc fingers tend to localize in the nucleus even without typical nuclear localization signals. To avoid the nuclear import of zinc-finger proteins, Minczuk et al. developed a system that involves incorporation of a nuclear export signal (NES) into the fusion protein in addition to the mitochondrial targeting sequence (MTS).[131] The fusion protein was found to be localized in the mitochondria, and could bind to the appropriate DNA sequence. Moreover, they showed the sequence-specific modification of mitochondrial DNA using a chimeric zinc-finger methylase. These results provide information for designing zinc-finger proteins targeting various organelles.

There is also considerable interest in proteins that bind to the telomeric DNA sequence and G-quadruplexes with a high specificity. Such proteins would be useful for regulating the telomere length and be applicable to cancer therapeutic tools. Recently, Balasubramanian's group first reported an engineered zinc-finger protein (Gq1) with a high binding specificity to the intramolecular G-quadruplex by the human telomeric sequence 5′-$(GGTTAG)_5$-3′, and the inhibitory activity of the enzyme telomerase in vitro.[132] Since the control of the telomerase activity is one of the ultimate goals in anticancer research, further development of this work and application to medical supply are strongly expected.

Sera's group described a facile method to control the transcription of endogenous genes in a quantitative and specific manner by treatment of designed artificial transcription factors that are fused to cell-penetrating peptides.[133] This is the first application of artificial transcription factors with cell-penetrating peptides. They chose the VEGF-A gene for the target gene and succeeded in penetrating their artificial transcription factors into the nucleus in HEK293 cell, where they can both up- and down-regulate the endogenous VEGF-A gene.

Finally, we review the recent zinc-finger nuclease technology toward gene therapy. A zinc-finger nuclease is one of the most promising and attractive tools for gene therapy. Chandrasegaran's group fused a zinc-finger domain with the DNA cleavage domain of the Type IIS restriction enzyme, *Fok*I, and showed that the fusion proteins cleave the DNA sequence at a position close to the zinc-finger-binding sequence (Figure 13).[134,135] The double-strand break promoted by the artificial zinc-finger nucleases was shown to increase the rate of homologous recombination.[136,137] Urnov et al. (Sangamo BioScience Inc.) succeeded in the endogenous human gene correction by using artificial zinc-finger nucleases.[138] They used zinc-finger nucleases to target an X-linked, severe combined immune deficiency (SCID) mutation in the *IL2Rγ* gene. Without any selection, about 20% of the mutated chromosomes in the cell line were corrected by homologous recombination induced by double strand cleavage of the zinc-finger nucleases. Moreover, it was reported that zinc-finger nucleases can drive the addition of not only short fragments, but also an 8-kb sequence containing genes into an endogenous locus.[139] However, there is still a big problem in that zinc-finger nucleases often induce a considerable cytotoxicity likely arising from cleavage at off-target sites.[137,140] In general, two zinc-finger nuclease units are assembled as heterodimers at the cleavage site, be-

© 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

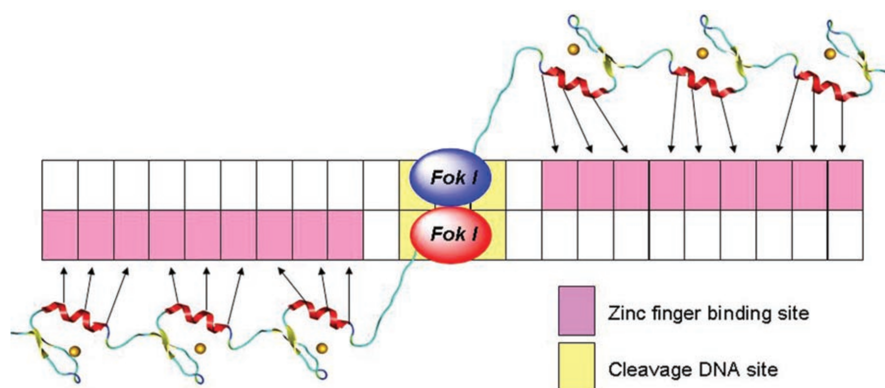


Figure 13. DNA recognition and cleavage by zinc-finger nucleases. The three-zinc finger is linked to the *Fok* I DNA cleavage domain through a flexible linker.

cause the catalytic domain in the zinc-finger nucleases must dimerize to become the active form. On the other hand, current zinc-finger nuclease architectures form cleavage-component homodimers that can decrease its safety or efficacy through off-target cleavage. Very recently, two different groups reported the promising solution of the side-reaction problem by using the similar structure-based redesign of the dimerized interface region.[141,142] Each group succeeded in engineering two variant zinc-finger nucleases that show effective DNA cleavage only when paired as a heterodimer by redesign of the dimarization interface amino acids. These findings should be useful in developing the therapeutic application of the zinc-finger nuclease technology.

Summary and Outlook

The first C_2H_2-type zinc-finger protein, TFIIIA, was discovered in 1985.[7] Before the human genome analysis ends, zinc-finger proteins were considered to be extremely unusual proteins in nature. However, after finishing the human genome project, zinc-finger proteins have determined to constitute one of the most common DNA binding motifs in eukaryotic transcription factors.[1] This review has described some of our strategies to create new zinc-finger proteins by redesigning the naturally occurring DNA-binding domains of the transcription factors. Especially, the C_2H_2-type zinc-finger domain of the Sp1 transcription factor provided an attractive framework for the redesign of artificial zinc-finger proteins with specific functions. The DNA binding properties of the designed zinc-finger proteins were customized by changing the secondary structure elements, namely, the α-helix and β-sheet, present in the domain. We successfully recognized the long DNA sequence using artificial multiple zinc-finger proteins that have several zinc-finger domains connected by a consensus linker. Artificial zinc-finger proteins with a designed non-consensus and functional linker can expand the natural recognition mode of the zinc-finger proteins. Catalytically active metalloproteins could be created by manipulating the zinc-coordination site; replacement of Cys with His produced an unnatural H_4-type zinc finger, which showed the specific hydrolytic cleavage of DNA. During these past two decades, zinc-finger research based on functional structural studies and their applications have been significantly developed. In particular, the zinc-finger technology toward desired endogeneous gene regulation is quickly progressing and its application to medical treatment has definitely been realized. These demonstrations suggest that designer zinc-finger proteins have great potential for applications in biomedicine and biotechnology. In the near future, as zinc-finger research advances, it may become possible to construct complicated synthetic cell networks and control better higher level vital functions at will.

Acknowledgements

This work was supported by the Ministry of Education, Culture, Sports, Science, and Technology, Japan (grants: 14370755, 16659028 and 17390028), and Suntory Institute for Bioorganic Research. We also thank the past and present members of Professor Sugiura's laboratory for their great contributions to this research.

[1] "Initial sequencing and analysis of the human genome": International Human Genome Sequencing Consortium, *Nature* **2001**, *409*, 860–921.
[2] S. Iuchi, N. Kuldell, *Zinc Finger Proteins: From Atomic Contact to Cellular Function*, Landes Biosciences, Georgetown, TX, **2004**.
[3] A. Klug, J. W. Schwabe, *FASEB J.* **1995**, *9*, 597–604.
[4] J. P. Mackay, M. Crossley, *Trends Biochem. Sci.* **1998**, *23*, 1–4.
[5] I. Sánchez-García, T. H. Rabbits, *Trends Genet.* **1994**, *10*, 315–320.
[6] J. M. Matthews, M. Sunde, *IUBMB Life* **2002**, *54*, 351–355.
[7] J. Miller, A. D. McLachlan, A. Klug, *EMBO J.* **1985**, *4*, 1609–1614.
[8] J. S. Hanas, D. J. Hazuda, D. F. Bogenhagen, F.-Y. Wu, C.-W. Wu, *J. Biol. Chem.* **1983**, *258*, 14120–14125.
[9] M. S. Lee, G. P. L. Gippert, K. V. Soman, D. A. Case, P. E. Wright, *Science* **1989**, *245*, 635–637.
[10] R. E. Klevit, J. R. Herriott, S. J. Horvath, *Proteins* **1990**, *7*, 215–226.
[11] J. G. Omichinski, G. M. Clore, E. Appella, K. Sakaguchi, A. M. Gronenborn, *Biochemistry* **1990**, *29*, 9324–9334.
[12] N. P. Pavletich, C. O. Pabo, *Science* **1991**, *252*, 809–817.
[13] R. T. Nolte, R. M. Conlin, S. C. Harrison, R. S. Brown, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 2938–2943.
[14] P. V. Pedone, R. Ghirlando, G. M. Clore, A. M. Gronenborn, G. Felsenfeld, J. G. Omichinski, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 2822–2826.
[15] M. Elrod-Erickson, M. A. Rould, L. Nekludova, C. Pabo, *Structure* **1996**, *4*, 1171–1180.
[16] M. Dhanasekaran, S. Negi, Y. Sugiura, *Acc. Chem. Res.* **2006**, *39*, 45–52.
[17] R. N. Armstrong, *Biochemistry* **2000**, *39*, 13625–13632.
[18] S. Diekmann, J. Weston, E. Anders, W. Boland, B. Schönecker, T. Hettmann, J. von Langen, S. Erhardt, M. Mauksch, M. Bräuer, C.

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Beckmann, M. Rost, P. Sperling, E. Heinz, *J. Biotechnol.* **2002**, *90*, 73–94.
[19] K. D. Karlin, *Science* **1993**, *261*, 701–708.
[20] D. C. Rees, *Annu. Rev. Biochem.* **2002**, *71*, 221–246.
[21] T. V. O'Halloran, *Science* **1993**, *261*, 715–725.
[22] H. B. Gray, B. G. Malmström, *Biochemistry* **1989**, *28*, 7499–7505.
[23] W. F. DeGrado, C. M. Summa, V. Pavone, F. Nastri, A. Lombardi, *Annu. Rev. Biochem.* **1999**, *68*, 779–819.
[24] R. B. Hill, D. P. Raleigh, A. Lombardi, W. F. DeGrado, *Acc. Chem. Res.* **2000**, *33*, 745–754.
[25] L. Baltzer, H. Nilsson, J. Nillsson, *Chem. Rev.* **2001**, *101*, 3153–3163.
[26] D. A. Moffet, M. H. Hecht, *Chem. Rev.* **2001**, *101*, 3191–3203.
[27] G. L. Butterfoss, B. Kuhlman, *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 49–65.
[28] C. M. Goodman, S. Choi, S. Shandler, W. F. DeGrado, *Nat. Chem. Biol.* **2007**, *3*, 252–262.
[29] K. L. Harris, S. Lim, S. J. Franklin, *Inorg. Chem.* **2006**, *45*, 10002–10012.
[30] H. W. Hellinga, *Curr. Opin. Biotechnol.* **1996**, *7*, 437–441.
[31] Y. Lu, S. M. Berry, T .D. Pfister, *Chem. Rev.* **2001**, *101*, 3047–3080.
[32] L. Baltzer, J. Nilsson, *Curr. Opin. Biotechnol.* **2001**, *12*, 355–360.
[33] G. Xing, V. J. De Rose, *Curr. Opin. Chem. Biol.* **2001**, *5*, 196–200.
[34] P. Wentworth, K. D. Janda, *Curr. Opin. Chem. Biol.* **1998**, *2*, 138–144.
[35] B. Golinelli-Pimpaneau, *Curr. Opin. Struct. Biol.* **2000**, *10*, 697–708.
[36] W. S. Dynan, R. Tjian, *Cell* **1983**, *32*, 669–680.
[37] D. Gidoni, W. S. Dynan, R. Tjian, *Nature* **1984**, *312*, 409–413.
[38] J. T. Kadonaga, K. A. Jones, R. Tjian, *Trends Biochem. Sci.* **1986**, *11*, 20–23.
[39] J. T. Kadonaga, K. R. Carner, F. R. Masiarz, R. Tjian, *Cell* **1987**, *51*, 1079–1090.
[40] M. Yokono, N. Saegusa, K. Matsushita, Y. Sugiura, *Biochemistry* **1998**, *37*, 6824–6832.
[41] S. Oka, Y. Shiraishi, T. Yoshida, T. Ohkubo, Y. Sugiura, Y. Kobayashi, *Biochemistry* **2004**, *43*, 16027–16035.
[42] A. Nomura, Y. Sugiura, *Inorg. Chem.* **2002**, *41*, 3693–3698.
[43] Y. Hori, Y. Sugiura, *J. Am. Chem. Soc.* **2002**, *124*, 9362–9363.
[44] Y. Hori, Y. Sugiura, *Biochemistry* **2004**, *43*, 3068–3074.
[45] M. Dhanasekaran, S. Negi, M. Imanishi, Y. Sugiura, *Biochemistry* **2007**, *46*, 7506–7513.
[46] J. G. Omichinski, P. V. Pedone, G. Felsenfeld, A. M. Gronenborn, G. M. Clore, *Nat. Struct. Biol.* **1997**, *4*, 122–132.
[47] J. R. Desjarlais, J. M. Berg, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 7345–7349.
[48] J. Nardelli, T. Gibson, P. Charnay, *Nucleic Acids Res.* **1992**, *20*, 4137–4144.
[49] W. E. Taylor, H. K. Suruki, A. H. Lin, P. Naraghi-Arani, R. Y. Igarashi, M. Younessian, P. Katkus, N. V. Vo, *Biochemistry* **1995**, *34*, 3222–3230.
[50] E. J. Rebar, C. O. Pabo, *Science* **1994**, *263*, 671–673.
[51] Y. Choo, A. Klug, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 11163–11167.
[52] A. C. Jamieson, S.-H. Kim, J. A. Wells, *Biochemistry* **1994**, *33*, 5689–5695.
[53] H. Wu, W. Yang, C. F. Barbas III, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 344–348.
[54] H. A. Greisman, C. O. Pabo, *Science* **1997**, *275*, 657–661.
[55] D. J. Segal, B. Dreier, R. R. Beerli, C. F. III Barbas, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 2758–2763.
[56] B. Dreier, D. J. Segal, C. F. Barbas, *J. Mol. Biol.* **2000**, *303*, 489–502.
[57] B. Dreier, R. R. Beerli, D. J. Segal, J. D. Flippin, C. F. Barbas III, *J. Biol. Chem.* **2001**, *276*, 29466–29478.
[58] S. A. Wolfe, R. A. Grant, M. Elrod-Erickson, C. O. Pabo, *Structure* **2001**, *9*, 717–723.
[59] T. Sera, C. Uranga, *Biochemistry* **2002**, *41*, 7074–7081.
[60] J .G. Mandell, C. F. Barbas III, *Nucleic Acids Res.* **2006**, *34*, W516–W523.
[61] N. Corbi, V. Libri, M. Fanciulli, J. M. Tinsley, K. E. Davies, C. Passananti, *Gene. Ther.* **2000**, *7*, 1076–1083.
[62] V. Libri, A. Onori, M. Fanciulli, C. Passananti, N. Corbi, *FEBS Lett.* **2004**, *560*, 75–80.
[63] T. Sera, *J. Virol.* **2005**, *79*, 5428–5436.
[64] T. Hsu, J. A. Gogos, S. A. Kirsh, F. C. Kafatos, *Science* **1992**, *257*, 1946–1950.
[65] M. Nagaoka, Y. Doi, J. Kuwahara, Y. Sugiura, *J. Am. Chem. Soc.* **2002**, *124*, 6526–6527.
[66] Y. Shiraishi, M. Imanishi, T. Morisaki, Y. Sugiura, *Biochemistry* **2005**, *44*, 2523–2528.
[67] N. P. Pavletich, C. O. Pabo, *Science* **1993**, *261*, 1701–1707.
[68] Y. Shiraishi, M. Imanishi, Y. Sugiura, *Biochemistry* **2004**, *43*, 6352–6359.
[69] K. R. Clemens, X. Liao, V. Wolf, P. E. Wright, J. M. Gottesfeld, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 10822–10826.
[70] Q. Liu, D. J. Segal, J. B. Ghiara, C. F. III Barbas, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 5525–5530.
[71] T. Kamiuchi, E. Abe, M. Imanishi, T. Kaji, M. Nagaoka, Y. Sugiura, *Biochemistry* **1998**, *37*, 13827–13834.
[72] M. Moore, Y. Choo, A. Klug, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 1437–1441.
[73] D. J. Segal, C. F. III Barbas, *Curr. Opin. Biotechnol.* **2001**, *12*, 632–637.
[74] D. K. Lee, W. Seol, J. S. Kim, *Curr. Top. Med. Chem.* **2003**, *3*, 645–657.
[75] D. J. Segal, J. T. Stege, C. F. III Barbas, *Curr. Opin. Plant Biol.* **2003**, *6*, 163–168.
[76] J. H. Laity, H. J. Dyson, P. E. Wright, *J. Mol. Biol.* **2000**, *295*, 719–727.
[77] T. Hirata, W. Nomura, M. Imanishi, Y. Sugiura, *Bioorg. Med. Chem. Lett.* **2005**, *15*, 2197–2201.
[78] M. Imanishi, Y. Hori, M. Nagaoka, Y. Sugiura, *Biochemistry* **2000**, *39*, 4383–4390.
[79] M. Imanishi, Y. Sugiura, *Biochemistry* **2002**, *41*, 1328–1334.
[80] J. L Pomerantz, S. A. Wolfe, C. O. Pabo, *Biochemistry* **1998**, *37*, 965–970.
[81] S. A. Wolfe, R. A. Grant, C. O. Pabo, *Biochemistry* **2003**, *42*, 13401–13409.
[82] M. Imanishi, W. Yan, T. Morisaki, Y. Sugiura, *Biochem. Biophys. Res. Commun.* **2005**, *333*, 167–173.
[83] W. Yan, M. Imanishi, S. Futaki, Y. Sugiura, *Biochemistry* **2007**, *46*, 8517–8524.
[84] G. Merutka, W. Shalongo, E. Stellwagen, *Biochemistry* **1991**, *30*, 4245–4248.
[85] W. Nomura, Y. Sugiura, *Biochemistry* **2003**, *42*, 14805–14813.
[86] M. Nagaoka, W. Nomura, Y. Shiraishi, Y. Sugiura, *Biochem. Biophys. Res. Commun.* **2001**, *282*, 1001–1007.
[87] W. Nomura, M. Nagaoka, Y. Shiraishi, Y. Sugiura, *Biochem. Biophys. Res. Commun.* **2003**, *300*, 87–92.
[88] C. V. Lund, M. Popkov, L. Magnenat, C. F. Barbas III, *Mol. Cell Biol.* **2005**, *25*, 9082–9091.
[89] T. Nakatsukasa, Y. Shiraishi, S. Negi, M. Imanishi, S. Futaki, Y. Sugiura, *Biochem. Biophys. Res. Commun.* **2005**, *330*, 247–252.
[90] J. Wöjcik, J. Göral, K. Pawlowski, A. Bierzynski, *Biochemistry* **1997**, *36*, 680–687.
[91] S. J. Franklin, *Curr. Opin. Chem. Biol.* **2001**, *5*, 201–208.
[92] J. T. Welch, W. R. Kearney, S. J. Franklin, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3725–3730.
[93] R. T. Kovacic, J. T. Welch, S. J. Franklin, *J. Am. Chem. Soc.* **2003**, *125*, 6656–6662.
[94] J. L. Bada, *Nature* **1995**, *374*, 594–595.
[95] R. Bentley, *Arch. Biochem. Biophys.* **2003**, *414*, 1–12.
[96] R. C. Milton, S. C. Milton, S. B. Kent, *Science* **1992**, *256*, 1445–1448.
[97] T. N. Schumacher, L. M. Mayr, D. L. Minor, Jr., M. A. Milhollen, M. W. Burgess, P. S. Kim, *Science* **1996**, *271*, 1854–1857.

© 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[98] K. P. Williams, X. H. Liu, T. N. Schumacher, H. Y. Lin, D. A. Ausiello, P. S. Kim, D. P. Bartel, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 11285–11290.
[99] D. M. Eckert, V. N. Malashkevich, L. H. Hong, P. A. Carr, P. S. Kim, *Cell* **1999**, *99*, 103–115.
[100] I. A. Kozlov, S. Mao, Y. Xu, X. Huang, L. Lee, P. S. Sears, C. Gao, A. R. Coyle, K. D. Janda, C. H. Wong, *ChemBioChem* **2001**, *2*, 741–746.
[101] K. Wiesehan, K. Buder, R. P. Linke, S. Patt, M. Stoldt, E. Unger, B. Schmitt, E. Bucci, D. Willbold, *ChemBioChem* **2003**, *4*, 748–753.
[102] K. Wiesehan, D. Willbold, *ChemBioChem* **2003**, *4*, 811–815.
[103] S. Negi, M. Dhanasekaran, T. Hirata, H. Urata, Y. Sugiura, *Chirality* **2006**, *18*, 254–258.
[104] B. L. Vallee, D. S. Auld, *Matrix Suppl.* **1992**, *1*, 5–19.
[105] B. L. Vallee, D. S. Auld, *Faraday Discuss.* **1992**, *93*, 47–65.
[106] T. M. Handel, S. A. Williams, W. F. DeGrado, *Science* **1993**, *261*, 879–885.
[107] B. L. Iverson, S. A. Iverson, V. A. Roberts, E. D. Getzoff, J. A. Tainer, S. J. Benkovic, R. A. Lerner, *Science* **1990**, *249*, 659–662.
[108] H. N. Müller, A. Skerra, *Biochemistry* **1994**, *33*, 14126–14135.
[109] L. Regan, N. D. Clarke, *Biochemistry* **1990**, *29*, 10878–10883.
[110] L. D. Costanzo, H. Wada, S. Geremia, L. Randaccio, V. Pavane, W. F. DeGrado, A. Lombardi, *J. Am. Chem. Soc.* **2001**, *123*, 12749–12757.
[111] A. Lombardi, C. M. Summa, S. Geremia, L. Randaccio, V. Pavone, F. W. DeGrado, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 6298–6305.
[112] A. Nomura, Y. Sugiura, *Inorg. Chem.* **2004**, *43*, 1708–1713.
[113] T. Itoh, Y. Fujii, T. Tada, Y. Yoshikawa, H. Hisada, *Bull. Chem. Soc. Jpn.* **1996**, *69*, 1265–1274.
[114] C. Andreini, L. Banci, I. Bertini, A. Rosato, *J. Proteome. Re*s. **2006**, *5*, 196–201.
[115] Y. Hori, K. Suzuki, Y. Okuno, M. Nagaoka, S. Futaki, Y. Sugiura, *J. Am. Chem. Soc.* **2000**, *122*, 7648–7653.
[116] T. Koike, M. Takamura, E. Kimura, *J. Am. Chem. Soc.* **1994**, *116*, 8443–8449.
[117] A. Nomura, Y. Sugiura, *J. Am. Chem. Soc.* **2004**, *126*, 15374–15375.
[118] R. Jaenisch, A. Bird, *Nat. Genet.* **2003**, *33*, 245–254.
[119] J. F. Costello, M. C. Frühwald, D. J. Smiraglia, L. J. Rush, G. P. Robertson, X. Gao, F. A. Wright, J. D. Feramisco, P. Peltomäki, J. C. Lang, D. E. Schuller, L. Yu, C. D. Bloomfield, M. A. Caligiuri, A. Yates, R. Nishikawa, H.-J. S. Huang, N. J. Petrelli, X. Zhang, M. S. O'Dorisio, W. A. Held, W. K. Cavenee, C. Plass, *Nat. Genet.* **2000**, *24*, 132–138.
[120] M. Esteller, P. G. Corn, S. B. Baylin, J .G. Herman, *Cancer Res.* **2001**, *61*, 3225–3229.
[121] G. L. Xu, T. H. Bestor, *Nat. Genet.* **1997**, *17*, 376–378.
[122] A. R. McNamara, P. J. Hurd, A. E. Smith, K. G. Ford, *Nucleic Acids Res.* **2002**, *30*, 3818–3830.
[123] C. D. Carvin, R. D. Parr, M. P. Kladde, *Nucleic Acids Res.* **2003**, *31*, 6493–6501.
[124] A. E. Smith, K. G. Ford, *Nucleic Acids Res.* **2007**, *35,* 740–754.
[125] F. Li, M. Papworth, M. Minczuk, C. Rohde, Y. Zhang, S. Ragozin, A. Jeltsch, *Nucleic Acids Res.* **2007**, *35*, 100–112.
[126] W. Nomura, C. F. Barbas III, *J. Am. Chem. Soc.* **2007**, *129*, 8676–8677.
[127] L. Reynolds, C. Ullman, M. Moore, M. Isalan, M. J. West, P. Clapham, A. Klug, Y. Choo, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1615–1620.
[128] D. J. Segal, J. Gonçalves, S. Eberhardy, C. H. Swan, B. E. Torbett, X. Li, C. F. Barbas III, *J. Biol. Chem.* **2004**, *279*, 14509–14519.
[129] T. Mino, T. Hatone, N. Matsumoto, T. Mori, Y. Mineta, Y. Aoyama, T. Sera, *J. Virol.* **2006**, *80*, 5405–5412.
[130] C. I. Stains, J. L. Furman, D. J. Segal, I. Ghosh, *J. Am. Chem. Soc.* **2006**, *128*, 9761–9765.
[131] M. Minczuk, M. A. Papworth, P. Kolasinska, M. P. Murphy, A. Klug, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 19689–19694.
[132] S. D. Patel, M. Isalan, G. Gavory, S. Ladame, Y. Choo, S. Balasubramanian, *Biochemistry* **2004**, *43*, 13452–13458.
[133] K. Tachikawa, O. Schröder, G. Frey, S. P. Briggs, T. Sera, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 15225–15230.
[134] Y. G. Kim, J. Cha, S. Chandrasegaran, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 1156–1160.
[135] J. Smith, J. M. Berg, S. Chandrasegaran, *Nucleic Acids Res.* **1999**, *27*, 674–681.
[136] M. Bibikavo, K. Beumer, J. K. Trautman, D. Carroll, *Science* **2003**, *300*,764.
[137] M. H. Porteus, D. Baltimore, *Science* **2003**, *300*, 763.
[138] F. D. Urnov, J. C. Miller, Y-L. Lee, C. M. Beausejour, J. M. Rock, S. Augustus, A. Jamieson, M. H. Porteus, P. D. Gregory, M. C. Holmes, *Nature* **2005**, *435*, 646–651.
[139] E. A. Moehle, J. M. Rock, Y-L. Lee, Y. Jouvenot, R. C. DeKelver, P. D. Gregory, F. D. Urnov, M. C. Holmes, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 3055–3060.
[140] S. Alwin, M. B. Gere, E. Guhl, K. Effertz, C. F. Barbas III, D. J. Segal, M. D. Weitzman, T. Cathomen, *Mol. Ther.* **2005**, *12*, 610–617.
[141] J. C. Miller, M. C. Holmes, J. Wang, D. Y. Guschin, Y.-L. Lee, I. Rupniewski, C. M. Beausejour, A. J. Waite, N. S. Wang, K. A. Kim, P. D. Gregory, C. O. Pabo, E. J. Rebar, *Nat. Biotechnol.* **2007**, *25*, 778–785.
[142] M. Szczepek, V. Brondani, J. Büchel, L. Serrano, D. J. Segal, T. Cathomen, *Nat. Biotechnol.* **2007**, *25*, 786–793.

Published online: January 31, 2008

© 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim