

C₂H₂ Zinc Fingers As DNA Binding Domains

Shiro Iuchi

Abstract

A great number of C₂H₂ zinc finger proteins selectively bind to specific DNA sequences and play a critical role in controlling transcription of genes. The specific binding is achieved by zinc finger domains with $\beta\beta\alpha$ structure that is formed by tetrahedral binding of Zn²⁺ ion to the canonical cysteine and histidine residues. Two to three tandem zinc fingers are necessary and sufficient for the specific binding without participation of any other domains. Zinc fingers bind in the major groove of the DNA, wrapping around the strands, with specificity conferred by side chains of several amino acid on the α helices. Some zinc finger proteins undergo homodimerization by hydrophobic interactions or by finger-finger binding and reinforce the specific binding to DNA. Conserved linkers between tandem fingers are necessary for stabilizing the DNA complex. Regulatory mechanisms of zinc finger binding to DNA are emerging. Some cellular factors are found to acetylate and phosphorylate zinc fingers and the linkers of a few proteins. These modifications alter the binding activity of the zinc finger proteins and hence control expression of their target genes. Other factors can methylate promoter regions of genes. This modification alters affinity of zinc finger proteins for the DNA segments and hence controls expression of their target genes.

Introduction

The C₂H₂ zinc finger consists of twenty to thirty amino acid residues that have a special secondary structure stabilized by zinc tetrahedral binding to two cysteine and two histidine residues.¹⁻⁶ Proteins with these zinc fingers are called C₂H₂ zinc finger proteins. The C₂H₂ zinc finger protein family is the largest group of all zinc finger protein families (Table 1) and the second largest group of all protein classes after the envelope glycoprotein GP120 family. The proteins are present in prokaryotes as well as eukaryotes and are abundant in mammals. More than 700 human genes, or greater than 2% of the total human genes, encode C₂H₂ zinc finger proteins.⁷ Not surprisingly, C₂H₂ zinc finger proteins participate in a variety of cellular activities including development, differentiation, and tumor suppression. Among C₂H₂ zinc finger proteins, many bind to DNA duplexes in a finger-mediated specific manner and participate in controlling expression of the target genes. C₂H₂ zinc fingers, which are often described as X₂CX₂₋₄CX₁₂HX₂₋₈H to show the intervals between the zinc binding residues, contain two β strands and one α helix. All the primary, secondary and tertiary structures are important for binding to DNA duplexes. In this chapter the general features of C₂H₂ zinc fingers-DNA binding are described.

Folding of C₂H₂ Zinc Fingers

Requirement of zinc ion for transcription factors to bind their cognate DNA was found in *Xenopus* TFIIIA (Transcription Factor IIIA) first.⁸ The requirement was due to the ion coordination with a small peptide domain, named zinc finger, that contains two canonical cysteine and histidine residues.^{1,4-6} Now, thousands of proteins are known to have C₂H₂ zinc fingers and the majority of the fingers are thought to bind to DNA. Zinc finger proteins can take three states: unfolded, folded and DNA-bound forms (Fig. 1 and for details of folded zinc finger domain see ref. 1 and 3 as well as Fig. 1 of Chapter 8). Unfolded zinc fingers do not bind to target DNA, but folded fingers bind to the cognate DNA duplexes. The protein molecules of the DNA complex are usually associated with other transcription factors that bind to different domains on the zinc finger protein, or to different parts of the zinc fingers. The C₂H₂ zinc finger motif contains all the information necessary for its folding but folds properly only when Zn²⁺ binds to the canonical residues, two cysteines and two histidines.^{9,10} The change in Gibbs free energy (ΔG) of the folding is enthalpy-driven. The amount, about -8.8 kcal per mole (Table 2),¹¹ indicates that the folded fingers are very stable. C₂H₂ zinc fingers contain three conserved hydrophobic amino acids at position -12, -3 and +4 in addition to the two canonical cysteines and histidines (see bolded amino acids in the sequence below where the first amino acid residue of the α helix is designated as position 1). It has been shown that these seven amino acid residues are necessary and sufficient to fold peptides properly by using a designed-synthetic peptide, K(-13)-YACAACAAFAAKAALAAHAAAHA-K13.⁹ This peptide binds Co²⁺, a substitute for Zn²⁺, to the cysteines with a higher affinity than to the histidines, and the proper folding occurs only when the ratio of the ion to the peptide is one or higher. Similar binding of Zn²⁺ to the zinc finger motif has also been observed in the synthetic peptide of the zif268 third finger [F(-12)-ACDICGRKFARSDERKRHTKIHLRQ-K15].^{10,12} The authors of this work have proposed that the zinc finger folding begins with binding of Zn²⁺ to the canonical cysteines and then establishes the tetrahedral structure involving the histidines. They have also suggested that the α helix emerges at the S(1)-DERKRHTKI-H11 sequence as the metal ion binding proceeds.

Tandem C₂H₂ Zinc Fingers

C₂H₂ zinc finger proteins often contain the fingers as tandem repeats connected by short oligopeptides, called linkers. Based on the number and repeat pattern of the fingers, C₂H₂ zinc

Table 1. Number of zinc finger proteins reported as of May 2003

Type	Number of Proteins	Average Domain's Length (AA)
C ₂ H ₂	23,989	23
CCHC (C ₂ HC)	5,215	17
RING	2,010	41
LIM	1,246	58
C ₂ C ₂ (2x)*	1,243	65
CCCH (C ₂ CH)	1,019	26
TAZ	51	81

Data are taken from Pfam 9.0 at Washington University in St. Louis. *The number is for proteins containing two C₂C₂ fingers.

finger proteins can be divided into four classes (Fig. 2), (A) single C₂H₂, (B) triple C₂H₂, (C) multiple-adjacent C₂H₂, and (D) separated-paired C₂H₂ zinc finger proteins.¹³ This classification is useful to predict how the zinc fingers of the proteins exert their binding activity. The single zinc finger differs from the other zinc fingers in that it requires an additional, non-zinc finger domain to establish the binding to the target DNA.¹⁴ The other classes, that is the triple, multiple-adjacent and separated-paired C₂H₂ fingers, bind to the specific DNA sequence without the aid of other domains. Both the triple C₂H₂ and the multiple-adjacent C₂H₂ zinc fingers bind to the cognate DNA at the three consecutive fingers.¹³ Another experiment has shown that a finger peptide with four-tandem identical repeats binds to the target DNA sequence at the three consecutive fingers only.¹⁵ Furthermore, the separated-paired C₂H₂ finger can also specifically bind to the cognate DNA, often at the one pair finger. Taking all these results into account, it may be concluded that two to three successive C₂H₂ zinc fingers are the most suitable unit to specifically bind to the cognate DNA.

Multiple-adjacent C₂H₂ zinc fingers bind to the DNA, based on the rule that two to three successive fingers are responsible for the specific DNA binding, but the fingers have additional DNA contacts. For example, TFIIIA with nine zinc fingers establishes the DNA binding at fingers 1-3, but briefly touches the DNA at finger 5 and weakly binds to the DNA at fingers 7-9.¹⁶⁻²⁰ Another example is Zac. This has seven zinc fingers and binds to a GC rich DNA duplex. Biochemical and genetic analysis have shown that Zac binds to the DNA at the two consecutive fingers (finger 2-3) and also at the three consecutive fingers (finger 5-7), without involving finger 4.²¹ Accordingly, Zac contacts the DNA at more than three fingers, keeping the two to three finger-DNA binding rule. It is curious that extra fingers are present in multiple-adjacent zinc finger and separated-paired zinc finger proteins, but it is becoming increasingly clear that the extra fingers are engaged in other interactions not only with a secondary locus of DNA but also with distinct molecules such as RNA and proteins.^{13,22-24}

It is interesting to speculate how these tandem zinc fingers have evolved from a single finger. The *Escherichia coli* gene, *arcA*, encodes a helix-turn-helix DNA binding repressor protein for the genes directing aerobic respiration enzymes.²⁵ This gene has an identical eleven-nucleotide sequence flanking a short region, and this genetic organization prompts the region to duplicate spontaneously (Iuchi and Lin, unpublished). Similarly, the identical

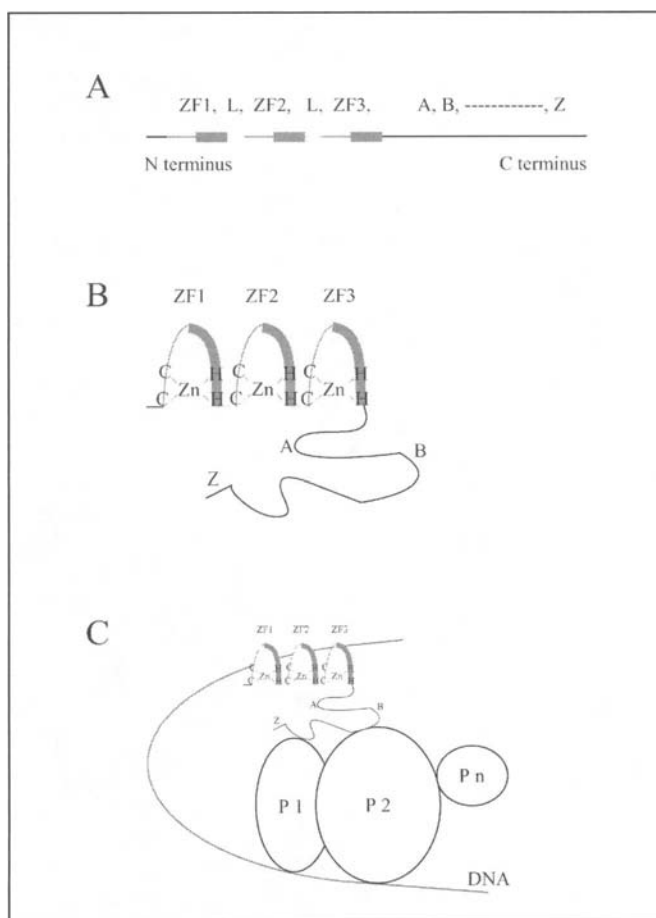


Figure 1. Schematic representation of three states of C₂H₂ zinc finger proteins. A) Primary structure of proteins. Polypeptide has three tandem fingers (ZF1, ZF2 and ZF3), each of which has α helix region (thick line) and is connected by linker (L). Other domains of the polypeptide are presented by A, B, ----, Z. B) Folded zinc finger proteins. Zn²⁺ tetrahedrally binds to the canonical cysteine and histidine residues and stabilizes the zinc finger's tertiary structure. (For details of folded C₂H₂ zinc fingers, see ref. 1 and 3 as well as Fig. 1 of Chapter 8). C) Zinc finger protein-DNA complex. The α helices of three zinc fingers bind to the DNA duplex in the major groove. Various domains of zinc finger proteins, including a part of zinc fingers, interact with other transcription factors.

linker sequences flanking the C₂H₂ zinc finger DNA may allow a zinc finger sequence to duplicate together with the linker itself (Fig. 3). The duplication would in turn stimulate the gene duplication further by using the linkers again or the zinc finger itself. The multiple zinc fingers can gain a better affinity for a DNA sequence and give the zinc finger protein molecules a selective advantage over the original finger proteins to function as transcription factors. This speculation is consistent with the fact that many of the separated-paired C₂H₂ and triple C₂H₂ zinc fingers conserve amino acid residues between tandem zinc fingers.¹³ After the duplication, point mutations would further improve its affinity for the target. Alternative splicing within the gene^{26,27} or recombination with another zinc finger gene²⁸ would allow the finger peptide to acquire tandem hybrid zinc fingers with a very different specificity. It should not be difficult to show that gene duplication indeed causes the evolution in living cells using well-established bacterial and yeast genetic systems.

Table 2. Thermodynamic parameters for zinc finger reactions

Reaction	ZF	K _a (M ⁻¹)	ΔG	ΔH	TΔS	Reference
ZFP + Zn ²⁺ ↔ ZFP•Zn ²⁺						
	1	3.5 × 10 ⁶	-8.8	-9.3	-0.5	11
ZFP•Zn ²⁺ + DNA ↔ ZFP•Zn ²⁺ •DNA						
Zif268	3	2.2 × 10 ¹⁰	NA	NA	NA	33
Zif268	3	2.8 × 10 ⁸	-11	-6.9	4.5	29
TFIIIA	3	1.3 × 10 ⁷	-9	-6.9	2.6	30
SP1	3	2.6 × 10 ⁷	-10	NA	NA	32
WT1	4	8.8 × 10 ⁸	-12	7	19	29
YY1	4	1.8 × 10 ⁶	-8	-11	-2.8	31

ZFP denotes zinc finger peptide. ZF shows number of zinc fingers per finger peptide. The relation between Gibbs free energy, enthalpy and entropy is given by $\Delta G = \Delta H - T\Delta S$, where T is absolute temperature. Negative ΔG value suggests that each above reaction favors the association- over the dissociation- reaction. The unit of energy is kcal/ mol.

Overall Features of the Triple-C₂H₂ Zinc Fingers Binding

The change in Gibbs free energy of zinc finger-DNA binding is similar to or higher than that of the zinc finger folding (Table 2)^{11,29-33} and the binding requires no enzyme action. ΔG is constant over a biologically-meaningful temperature range, 5 to 45 °C.³¹ The reactions are mostly enthalpy- and entropy-driven reactions, but some are only enthalpy- or only entropy-driven. The ΔG values suggest that the C₂H₂ zinc fingers bind to DNA as strongly as some antibodies, whose K_a for their antigens is 10⁵ to 10¹² M⁻¹.³⁴ These zinc finger-DNA complexes are of such high affinity that they routinely display an electrophoretic mobility shift on native polyacrylamide gels.

When an amino acid and guanine are mixed in water, the amino acid starts to associate with the nucleotide base and the reaction soon reaches equilibrium. Tendency of the reaction is described by the affinity constant, $K_a = 1/K_d = [\text{aa-guanine}] / [\text{aa}][\text{guanine}]$. The K_a of amino acids for guanine is in the order arginine > lysine > glutamine > glutamate > glycine.³⁵ Amino acids also have an inherent K_a for the three other bases. Accordingly, it is predictable that the amino acid-base associations are key for zinc finger peptides to specifically bind to the DNA duplexes. Indeed, the arginine-guanine contact, whose association is the greatest of all the combinations, is quite often present in zinc finger-DNA complexes (Fig. 4B and C). However, each amino acid residue of the zinc finger peptides does not have free mobility to access to the favorite bases due to the rigid zinc finger structure. How the zinc finger peptides recognize and bind to the cognate DNA duplex is a big question for the zinc finger-DNA interactions and the answer has come mostly from structure-oriented and genetic oriented analyses of zinc finger-DNA complexes. In particular, the structure of Zif268-DNA and TFIIIA-DNA complexes contributed enormously to the understanding of how the triple fingers align with the DNA duplex and how each finger interacts with the nucleotides of the DNA.^{17,18,36} Three α helices of the Zif268 triple zinc fingers bind in the major groove of the target DNA duplex antiparallel to the primary strand (defined as the strand to which zinc finger contacts most, Fig. 4A and for details of the binding see ref. 2), making hydrogen bonds and forming hydrophobic interactions with nucleotide bases and wrapping around the DNA for almost one turn. In addition to hydrogen bonds and hydrophobic interactions, phosphate contacts also participate in the zinc

finger-DNA complex formation. Phosphate contacts, linking to the DNA backbone, may not be significant in determining the specificity but appear to be important for strengthening the binding.³⁷ Through further analysis of various zinc finger-DNA complexes, it was found that the overall DNA binding mode of Zif268 is shared with other tandem C₂H₂ zinc fingers. These fingers include TFIIIA, SP1, GL1, WT1 and Tramtrack^{17,18,38-41} as well as the single finger GAGA.¹⁴ In addition to these GC-rich DNA binding transcription factors, an AT-rich DNA recognizing transcription factor, CF2II, is also thought to take the same binding mode.²⁷ However, one AT-rich DNA binding transcription factor, Nmp4, is proposed to associate with minor groove of the DNA.⁴²

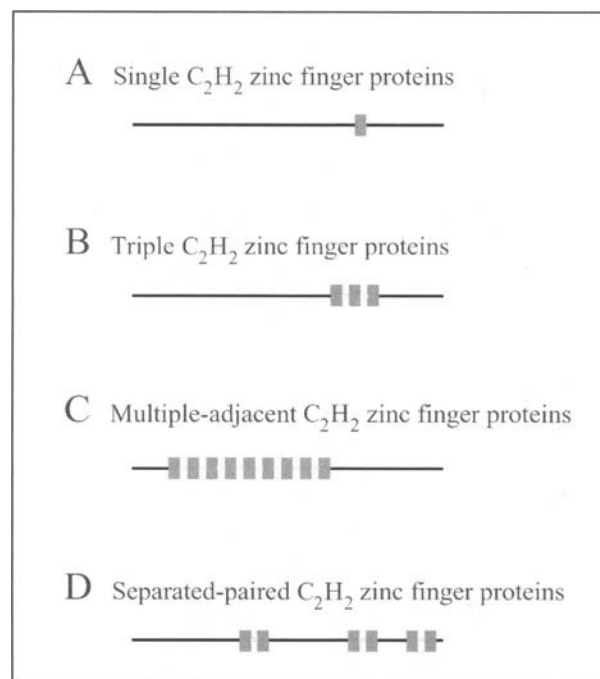


Figure 2. Schematic representation of four classes of C₂H₂ zinc finger proteins. Only one example is shown for each class. Some multiple-adjacent C₂H₂ zinc finger proteins contain more than thirty zinc fingers. The number and the pattern of C₂H₂ zinc fingers indicate how the zinc fingers are involved in the DNA binding (see text).

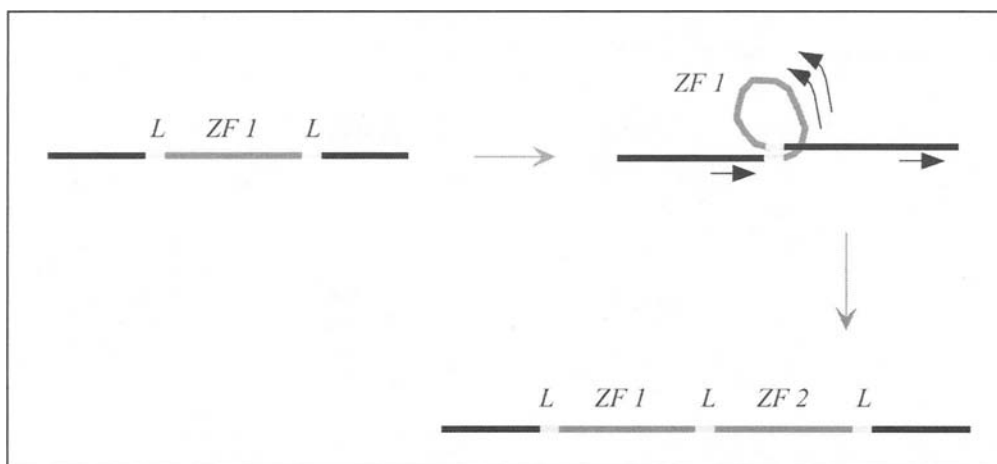


Figure 3. A model of intramolecular duplication of the *zinc finger* region. ZF1 flanked by two identical linker sequences for (L) would form a loop overlapping at the identical sites, or form a hairpin structure complementing within the single strand at the identical sites. When DNA polymerase reads the loop or the hairpin structure twice, then the *zinc finger* region is duplicated together with the linker sequence.

DNA of the Zif268-DNA complex takes a slightly unwound B form so that the major groove is still wide and can be deep.² This form contains 11.3 bp per turn, which is slightly more bases than the B-form itself (10.5 bp per turn).

Side Chain-Base Contacts in C₂H₂ Zinc Finger-DNA Complexes

Amino acid side chains on the surface of α helices in Zif268 are exposed to the cognate DNA duplex in the major groove, and the side chains at position 6, 3, 2 and -1 contact selectively with four successive bases (subsite) (Fig. 4B). Residues at position 6, 3 and -1 bind to three successive bases of the primary strand, the strand contacted most by the side chains, and the residue at position 2 binds to the fourth base present in the complementary strand. In this way, Zif268 finger 3, 2 and 1 recognize the primary strand's subsite, 5'-GCGT-3', 5'-TGGG-3' and 5'-GCGT-3', respectively. Of the four base pairs in the subsites, the base pair(s) at the end is shared by adjacent fingers. Consequently, the Zif268 triple zinc fingers bind to the ten nucleotide base pair, 5'-GCGTGGGCGT-3'. Based on statistics, one can predict with a high accuracy which side chain-base contacts can happen at the key positions of the α helices in the Zif268 finger context.⁴³ Stereochemistry between amino acids and bases can also predict the contacts with a similar accuracy or somewhat less accuracy.⁴⁴

The rule that side chains positioned at 6, 3, 2 and -1 contact bases is well preserved throughout the family of C₂H₂ zinc fingers, but additional contacts can occur in other fingers. Such an example has been observed in the TFIIIA-DNA complex (Fig. 4C).^{17,18} There is an additional contact of the side chain at position 10. Furthermore, some side chains at the regular positions 6, 3, 2 and -1 reach bases out of the subsites. When Zif268 was engineered to bind to AT-rich DNA duplexes, many irregular contacts occurred.^{2,45} In fact, many of these irregularities occurred in the complementary strand but not in the primary strand.⁴⁶ These facts, together with results obtained by mutational analyses, have led to the conclusion that there is no simple code for side chain-base contacts.^{2,47} Superimposition of several finger-DNA complex images showed that the irregular contacts coincide with the slight difference in docking angle of the α helix to DNA. The difference may reveal the influences of all of the

factors involved in specific zinc finger-DNA binding, including amino acid residues within the subsite,⁴⁷ linkers, and adjacent fingers.^{2,37}

Absence of a strict rule in the side chain-base contacts makes it impossible to predict the side chain-base contact with 100%

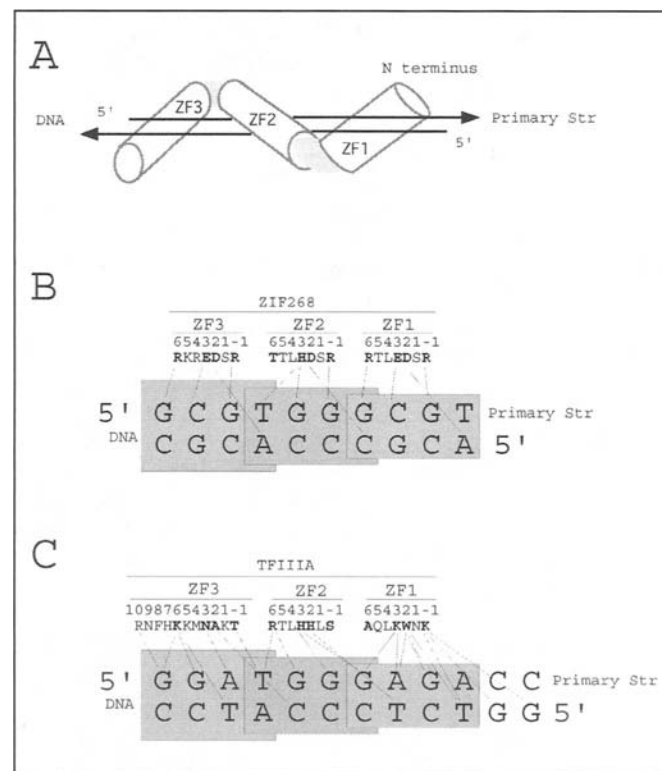


Figure 4. A) Antiparallel binding of zinc fingers to DNA. The orientation of the zinc fingers and the primary strand are antiparallel. B) and C) show amino acid side chain-nucleotide base contacts. Boxes show subsites. The side chain-DNA contacts are drawn, based on the results obtained with the crystallized Zif268-DNA complex B)^{2,36} and with the TFIIIA-DNA complex in solution C).¹⁷ Similar results of the side chain-DNA contact for the TFIIIA-DNA complex has been obtained with the crystallized complex.¹⁸

accuracy, but it rather indicates that C₂H₂ zinc fingers are able to bind to almost any DNA duplexes. The versatility of the zinc finger binding relies on hydrogen bonds that can (i) make long distance contact to bases with and without participation of a water molecule (>2.75 angstroms),^{35,48} (ii) make contacts to more than one base, and (iii) make contacts to bases with flexible angles.³⁵ Zinc fingers' binding also relies on nonspecific hydrophobic interactions. Although it is difficult to predict exactly which zinc finger peptide sequence specifically binds to a DNA sequence, one can obtain desired zinc fingers by the phage display methods and manipulate expression of the target genes with the obtained finger proteins.^{2,49-53}

Linkers

About half of zinc finger proteins have a well-conserved linker, TGEKP, between adjacent fingers.³⁷ The importance of the linker in zinc finger-DNA binding has been revealed by analyzing the effect of mutation on DNA affinity after making substitutions for the conserved residues^{54,55} and by analyzing structures of finger-DNA complexes.^{2,3,56,57} The linker is flexible in solution without DNA. Upon forming the finger-DNA complex, however, the conserved lysine residue of the TGEKP sequence (Fig. 1) interacts with the phosphate backbone. Moreover, linkers contact the C terminus of the preceding α helix by involving threonine and glycine residues so that the zinc finger-DNA complex becomes more stable (C capping).^{17,57,58} Alternative splicing of the *WT1* message disrupts the conserved linker between finger 3 and 4 by inserting the sequence KTS. The change from TGEKP to sequence TGKTSEP is accompanied by a severe decrease in DNA binding. An NMR study of the finger peptide-DNA complex has shown that the insertion increases the flexibility between finger 3 and 4 and abrogates binding of finger 4 to its cognate site.⁴¹ These findings have demonstrated that the conserved linker is not only necessary to promote fingers to fit completely into the DNA major groove, but also necessary to strengthen the DNA binding. A separated-paired zinc finger class, ZAS family, also contains the TGEKP linker,⁵⁹ but fingers belonging to a subtype of the same class, basonuclein-type fingers, have a distinct linker, LR(K)MHK.^{13,60} Tramtrack also has a distinct linker sequence, KRNKVYP.³⁹

Dimerization of Zinc Finger Protein

DNA binding proteins often bind to the target DNA duplex as dimers in order to increase their binding affinity and modulate their regulatory activities.⁶¹ This is true for C₂H₂ zinc finger proteins. Many multiple-adjacent C₂H₂ zinc finger proteins, such as Ikaros, Roaz, GL1, SW15, TRPS-1 and Zac, form homodimers on the target DNA duplex using their zinc fingers.^{23,24,62-64} GL1 forms a homodimer through the hydrophobic surface of zinc finger 1 that is not involved in the DNA binding. Similarly, SW15 dimerizes through finger 1 at the hydrophobic surface of both the β strand and the α helix's C-terminal half. This hydrophobic dimerization was applied to make the homodimer of an engineered two-finger peptide on the DNA.⁶³ However, the hydrophobic binding is not the only dimerization mechanism for zinc fingers. Ikaros contains six zinc fingers. The N-terminal four fingers participate in the specific binding to the DNA, and the C-terminal two fingers, separated from the four and making a pair of fingers, are responsible for the homodimerization.²³ Recently, a mutational analysis of zinc fingers 5 and 6 demonstrated that amino acid residues on the α helix are responsible for the

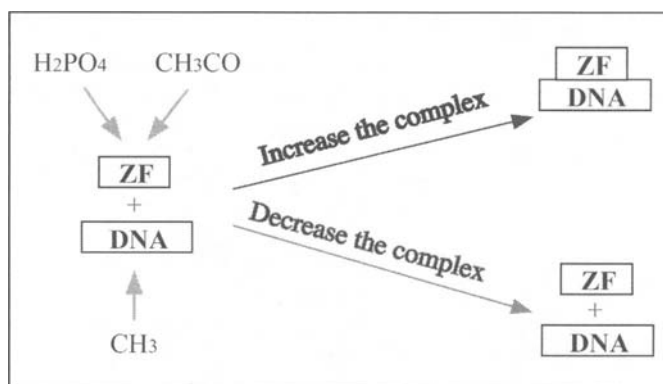


Figure 5. Modification of zinc finger proteins and the target DNA.

specific finger-finger interaction and hence for the Ikaros homodimerization.⁶² These amino acids correspond to those of C₂H₂ zinc fingers with DNA binding activity, and differ from those needed for the hydrophobic dimerization.

Regulation of Zinc Finger Binding to DNA

In vitro zinc finger-DNA binding proceeds as long as both substances are present in a suitable buffer, but the in vivo binding does not happen automatically and proceeds only under certain conditions. The regulation is mediated by modification of either the zinc finger protein itself or the target DNA (Fig. 5). The former can be executed by acetylation and also by phosphorylation, whereas the latter is mediated by methylation. All these modifications increase or decrease the zinc finger protein-DNA complex formation. The direction in which the regulation proceeds, repression or activation, depends on the fingers and the DNA, as well as the regulatory signal involved.

Acetylation of Zinc Fingers

YY1 (Yin Yang 1), with four tandem C₂H₂ zinc fingers, binds to various genes and represses or activates their target gene expression.^{65,66} PCAF (p300/CBP associated factor) acetylates YY1 at the zinc finger domain and inhibits DNA binding. PCAF also acetylates members of the KLF (Krüppel-like factor) family belonging to the triple C₂H₂ zinc finger protein class.^{67,68} It acetylates KLF13 at two lysines of the TGEKK linker between finger 2 and 3 and results in stimulating the fingers to bind to DNA. Another coactivation factor CBP/p300 acetylates a total of six lysines including the linker's two lysines, but it prevents the fingers from binding to DNA. Of the six residues, the lysine of finger 1 directly contacts bases. However, acetylation of this lysine is not enough to interrupt the DNA binding and acetylation of all six lysines is required. Moreover in vitro experiments showed that KLF2 zinc finger cotransfected with PCAF increases the target γ globulin gene expression in vivo while the zinc finger cotransfected with p300 decreases gene expression.⁶⁸ Therefore, it can be concluded that acetylation of KLF2 regulates the target gene expression both positively and negatively. Because lysine residues are abundant in a variety of zinc fingers, acetylation of C₂H₂ zinc fingers is likely to be a common mechanism to regulate gene expression by modulating the zinc fingers' DNA binding activity.

Another domain of YY1, the central glycine-lysine-rich domain, is acetylated by p300 and PCAF. This acetylation has nothing to do with the DNA binding but is required for fully repressing the target gene transcription.⁶⁹

Phosphorylation of Zinc Fingers

During mitosis, Ikaros is phosphorylated on threonine/serine residues of the three linkers and interrupted for its DNA binding.⁷⁰ Phosphorylation of linkers also occurs to Sp1. This suggests that phosphorylation is a global inhibitory mechanism to keep C₂H₂ zinc fingers out of DNA during mitosis. Interestingly, phosphorylation modifies the linkers as the favored sites. A few interpretations can be given. First, modification of the conserved linker is the most effective way to modulate zinc finger binding activity because the linker is essential for the high affinity zinc finger-DNA binding. Second, the modification sites, threonines/serines, are always available in the conserved linkers. Finally, linkers may be accessible even in the finger-DNA complexes so that zinc finger peptides can be separated from the DNA as soon as the regulatory signal is delivered to the fingers. The same arguments can be applied to regulation by acetylation since it also modifies linkers.

A serine residue of the C terminus immediately after a zinc finger can be phosphorylated in the case of Cre1 (catabolite repression) of *Hypocrea jecorina*, perhaps by caseinkinase-II-like enzyme.⁷¹ Surprisingly, this phosphorylation increases the finger's DNA binding activity, suggesting that the unphosphorylated serine residue may mask the finger's DNA binding activity directly by interacting with the finger or indirectly by interacting with other intramolecular domains.

Methylation of DNA

Engineered zinc fingers have been shown to distinguish methylated DNA duplex from unmethylated.⁷² It is becoming increasingly clear that natural zinc fingers differentially recognize methylated and unmethylated CpG (cytidine-guanidine dinucleotide pairs) and play important regulatory roles in the expression of target genes in vivo. CTCF (CCCTC-binding factor) has eleven C₂H₂ zinc fingers and binds to an element present between a promoter and an enhancer in order to block the enhancer activity.⁷³ The CTCF binding sites are widely distributed in vertebrates and several similar CpG-rich sequences are present in ICR (imprinted-control region) of the *ifg2/H19* as well. Electrophoretic mobility shift experiments have demonstrated that CTCF binds to the CpG-rich elements of ICR but only when these are unmethylated.⁷⁴ Although more experiments have to be done to obtain the CTCF function in vivo, the zinc finger proteins are likely to participate in the genomic imprinting process by differentially binding to the methylated- and unmethylated CpG-rich elements.

Kaiso, with triple C₂H₂ zinc fingers, is another example of this type of DNA binding. However, in this case, the zinc fingers recognize the methylated CpG-rich sequence but not the unmethylated sequence.^{75,76} In vivo transfection experiments have shown that kaiso represses expression of genes with methylated binding sites in the promoter.⁷⁵

Acknowledgements

I thank Dr. R. Freeman for information on the Pfam database.

References

- Klug A, Schwabe JW. Protein motifs 5. Zinc fingers. FASEB J 1995; 9(8):597-604.
- Pabo CO, Peisach E, Grant RA. Design and selection of novel Cys2His2 zinc finger proteins. Annu Rev Biochem 2001; 70:313-340.
- Laity JH, Lee BM, Wright PE. Zinc finger proteins: new insights into structural and functional diversity. Curr Opin Struct Biol 2001; 11(1):39-46.
- Miller J, McLachlan AD, Klug A. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. EMBO J 1985; 4(6):1609-1614.
- Brown RS, Sander C, Argos P. The primary structure of transcription factor TFIIIA has 12 consecutive repeats. FEBS Lett 1985; 186(2):271-274.
- Diakun GP, Fairall L, Klug A. EXAFS study of the zinc-binding sites in the protein transcription factor IIIA. Nature 1986; 324(6098):698-699.
- Lander ES, Linton LM, Birren B et al. Initial sequencing and analysis of the human genome. Nature 2001; 409(6822):860-921.
- Hanas JS, Hazuda DJ, Bogenhagen DF et al. *Xenopus* transcription factor A requires zinc for binding to the 5 S RNA gene. J Biol Chem 1983; 258(23):14120-14125.
- Michael SF, Kilfoil VJ, Schmidt MH et al. Metal binding and folding properties of a minimalist Cys2His2 zinc finger peptide. Proc Natl Acad Sci USA 1992; 89(11):4796-4800.
- Cox EH, McLendon GL. Zinc-dependent protein folding. Curr Opin Chem Biol 2000; 4(2):162-165.
- Lachenmann MJ, Ladbury JE, Phillips NB et al. The hidden thermodynamics of a zinc finger. J Mol Biol 2002; 316(4):969-989.
- Miura T, Satoh T, Takeuchi H. Role of metal-ligand coordination in the folding pathway of zinc finger peptides. Biochim Biophys Acta 1998; 1384(1):171-179.
- Iuchi S. Three classes of C₂H₂ zinc finger proteins. Cell Mol Life Sci 2001; 58(4):625-635.
- Omichinski JG, Pedone PV, Felsenfeld G et al. The solution structure of a specific GAGA factor-DNA complex reveals a modular binding mode [see comments]. Nat Struct Biol 1997; 4(2):122-132.
- Nagaoka M, Kaji T, Imanishi M et al. Multiconnection of identical zinc finger: implication for DNA binding affinity and unit modulation of the three zinc finger domain. Biochemistry 2001; 40(9):2932-2941.
- Liao XB, Clemens KR, Tennant L et al. Specific interaction of the first three zinc fingers of TFIIIA with the internal control region of the *Xenopus* 5 S RNA gene. J Mol Biol 1992; 223(4):857-871.
- Wuttke DS, Foster MP, Case DA et al. Solution structure of the first three zinc fingers of TFIIIA bound to the cognate DNA sequence: determinants of affinity and sequence specificity. J Mol Biol 1997; 273(1):183-206.
- Nolte RT, Conlin RM, Harrison SC et al. Differing roles for zinc fingers in DNA recognition: structure of a six-finger transcription factor IIIA complex. Proc Natl Acad Sci USA 1998; 95(6):2938-2943.
- Clemens KR, Zhang P, Liao X et al. Relative contributions of the zinc fingers of transcription factor IIIA to the energetics of DNA binding. J Mol Biol 1994; 244(1):23-35.
- Fairall L, Rhodes D, Klug A. Mapping of the sites of protection on a 5 S RNA gene by the *Xenopus* transcription factor IIIA. A model for the interaction. J Mol Biol 1986; 192(3):577-591.
- Hoffmann A, Ciani E, Boeckardt J et al. Transcriptional activities of the zinc finger protein Zac are differentially controlled by DNA binding. Mol Cell Biol 2003; 23(3):988-1003.
- Fan CM, Maniatis T. A DNA-binding protein containing two widely separated zinc finger motifs that recognize the same DNA sequence. Genes Dev 1990; 4(1):29-42.
- Sun L, Liu A, Georgopoulos K. Zinc finger-mediated protein interactions modulate Ikaros activity, a molecular control of lymphocyte development. EMBO J 1996; 15(19):5358-5369.
- Mackay JB, Crossley M. Zinc fingers are sticking together. Trends Biochem Sci 1998; 23(1):1-4.
- Iuchi S, Lin EC. Adaptation of *Escherichia coli* to respiratory conditions: regulation of gene expression. Cell 1991; 66(1):5-7.
- Hsu T, Gogos JA, Kirsh SA et al. Multiple zinc finger forms resulting from developmentally regulated alternative splicing of a transcription factor gene. Science 1992; 257(5078):1946-1950.
- Gogos JA, Hsu T, Bolton J et al. Sequence discrimination by alternatively spliced isoforms of a DNA binding zinc finger domain. Science 1992; 257(5078):1951-1955.
- Greisman HA, Pabo CO. A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites. Science 1997; 275(5300):657-661.

29. Hamilton TB, Borel F, Romaniuk PJ. Comparison of the DNA binding characteristics of the related zinc finger proteins WT1 and EGR1. *Biochemistry* 1998; 37(7):2051-2058.
30. Liggins JR, Privalov PL. Energetics of the specific binding interaction of the first three zinc fingers of the transcription factor TFIIIA with its cognate DNA sequence. *Proteins* 2000; Suppl 4:50-62.
31. Houbaviy HB, Burley SK. Thermodynamic analysis of the interaction between YY1 and the AAV P5 promoter initiator element. *Chem Biol* 2001; 8(2):179-187.
32. Matsushita K, Sugiura Y. Effect of arginine mutation of alanine-556 on DNA recognition of zinc finger protein Sp1. *Bioorg Med Chem* 2001; 9(9):2259-2267.
33. Miller JC, Pabo CO. Rearrangement of side-chains in a Zif268 mutant highlights the complexities of zinc finger-DNA recognition. *J Mol Biol* 2001; 313(2):309-315.
34. Harlow E, Lane D. *Antibodies*. New York: Cold Spring Harbor Laboratory; 1988.
35. Saenger W. *Principles of nucleic acid structure*. New York: Springer-Verlag; 1983.
36. Pavletich NP, Pabo CO. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* 1991; 252(5007):809-817.
37. Wolfe SA, Nekudova L, Pabo CO. DNA recognition by Cys2His2 zinc finger proteins. *Annu Rev Biophys Biomol Struct* 2000; 29:183-212.
38. Pavletich NP, Pabo CO. Crystal structure of a five-finger GLI-DNA complex: new perspectives on zinc fingers. *Science* 1993; 261(5129):1701-1707.
39. Fairall L, Schwabe JW, Chapman L et al. The crystal structure of a two zinc-finger peptide reveals an extension to the rules for zinc-finger/DNA recognition. *Nature* 1993; 366(6454):483-487.
40. Yokono M, Saegusa N, Matsushita K et al. Unique DNA binding mode of the N-terminal zinc finger of transcription factor Sp1. *Biochemistry* 1998; 37(19):6824-6832.
41. Laity JH, Dyson HJ, Wright PE. Molecular basis for modulation of biological function by alternate splicing of the Wilms' tumor suppressor protein. *Proc Natl Acad Sci USA* 2000; 97(22):11932-11935.
42. Torrungruang K, Alvarez M, Shah R et al. DNA binding and gene activation properties of the Nmp4 nuclear matrix transcription factors. *J Biol Chem* 2002; 277(18):16153-16159.
43. Benos PV, Lapedes AS, Stormo GD. Probabilistic code for DNA recognition by proteins of the EGR family. *J Mol Biol* 2002; 323(4):701-727.
44. Suzuki M, Brenner SE, Gerstein M et al. DNA recognition code of transcription factors. *Protein Eng* 1995; 8(4):319-328.
45. Choo Y, Klug A. Designing DNA-binding proteins on the surface of filamentous phage. *Curr Opin Biotechnol* 1995; 6(4):431-436.
46. Wolfe SA, Grant RA, Elrod-Erickson M et al. Beyond the "recognition code": structures of two Cys2His2 zinc finger/TATA box complexes. *Structure (Camb)* 2001; 9(8):717-723.
47. Gogos JA, Jin J, Wan H et al. Recognition of diverse sequences by class I zinc fingers: asymmetries and indirect effects on specificity in the interaction between CF2II and A+T-rich elements. *Proc Natl Acad Sci USA* 1996; 93(5):2159-2164.
48. Tsui V, Radhakrishnan I, Wright PE et al. NMR and molecular dynamics studies of the hydration of a zinc finger-DNA complex. *J Mol Biol* 2000; 302(5):1101-1117.
49. Choo Y, Isalan M. Advances in zinc finger engineering. *Curr Opin Struct Biol* 2000; 10(4):411-416.
50. Beerli RR, Barbas CF, 3rd. Engineering polydactyl zinc-finger transcription factors. *Nat Biotechnol* 2002; 20(2):135-141.
51. Choo Y, Klug A. Toward a code for the interactions of zinc fingers with DNA: selection of randomized fingers displayed on phage [published erratum appears in *Proc Natl Acad Sci USA* 1995; 92(2):646]. *Proc Natl Acad Sci USA* 1994; 91(23):11163-11167.
52. Rebar EJ, Pabo CO. Zinc finger phage: affinity selection of fingers with new DNA-binding specificities. *Science* 1994; 263(5147):671-673.
53. Isalan M, Choo Y. Engineering nucleic acid-binding proteins by phage display. *Methods Mol Biol* 2001; 148:417-429.
54. Wilson TE, Day ML, Pexton T et al. In vivo mutational analysis of the NGFI-A zinc fingers. *J Biol Chem* 1992; 267(6):3718-3724.
55. Choo Y, Klug A. A role in DNA binding for the linker sequences of the first three zinc fingers of TFIIIA. *Nucleic Acids Res* 1993; 21(15):3341-3346.
56. Elrod-Erickson M, Rould MA, Nekudova L et al. Zif268 protein-DNA complex refined at 1.6 Å: a model system for understanding zinc finger-DNA interactions. *Structure* 1996; 4(10):1171-1180.
57. Laity JH, Dyson HJ, Wright PE. DNA-induced alpha-helix capping in conserved linker sequences is a determinant of binding affinity in Cys(2)-His(2) zinc fingers. *J Mol Biol* 2000; 295(4):719-727.
58. Foster MP, Wuttke DS, Radhakrishnan I et al. Domain packing and dynamics in the DNA complex of the N-terminal zinc fingers of TFIIIA. *Nat Struct Biol* 1997; 4(8):605-608.
59. Wu LC. ZAS: C2H2 zinc finger proteins involved in growth and development. *Gene Expr* 2002; 10(4):137-152.
60. Tseng H, Green H. Basonuclin: A keratinocyte protein with multiple paired zinc fingers. *Proc Natl Acad Sci USA* 1992; 89(21):10311-10315.
61. Ptashne M. *A genetic switch: Gene control and phage λ*. Cambridge: Cell Press & Blackwell Scientific Publications; 1986.
62. McCarty AS, Kleiger G, Eisenberg D et al. Selective dimerization of a C2H2 zinc finger subfamily. *Mol Cell* 2003; 11(2):459-470.
63. Wang BS, Grant RA, Pabo CO. Selected peptide extension contacts hydrophobic patch on neighboring zinc finger and mediates dimerization on DNA. *Nat Struct Biol* 2001; 8(7):589-593.
64. Tsai RY, Reed RR. Identification of DNA recognition sequences and protein interaction domains of the multiple-Zn-finger protein Roaz. *Mol Cell Biol* 1998; 18(11):6447-6456.
65. Shi Y, Lee JS, Galvin KM. Everything you have ever wanted to know about Yin Yang 1. *Biochim Biophys Acta* 1997; 1332(2):F49-66.
66. Thomas MJ, Seto E. Unlocking the mechanisms of transcription factor YY1: are chromatin modifying enzymes the key? *Gene* 1999; 236(2):197-208.
67. Song CZ, Keller K, Chen Y et al. Functional interplay between CBP and PCAF in acetylation and regulation of transcription factor KLF13 activity. *J Mol Biol* 2003; 329(2):207-215.
68. Song CZ, Keller K, Murata K et al. Functional interaction between coactivators CBP/p300, PCAF, and transcription factor FKLF2. *J Biol Chem* 2002; 277(9):7029-7036.
69. Yao YL, Yang WM, Seto E. Regulation of transcription factor YY1 by acetylation and deacetylation. *Mol Cell Biol* 2001; 21(17):5979-5991.
70. Dovat S, Ronni T, Russell D et al. A common mechanism for mitotic inactivation of C2H2 zinc finger DNA-binding domains. *Genes Dev* 2002; 16(23):2985-2990.
71. Cziferszky A, Mach RL, Kubicek CP. Phosphorylation positively regulates DNA binding of the carbon catabolite repressor Cre1 of *Hypocrea jecorina* (*Trichoderma reesei*). *J Biol Chem* 2002; 277(17):14688-14694.
72. Choo Y. Recognition of DNA methylation by zinc fingers. *Nat Struct Biol* 1998; 5(4):264-265.
73. Bell AC, West AG, Felsenfeld G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 1999; 98(3):387-396.
74. Hark AT, Schoenherr CJ, Katz DJ et al. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 2000; 405(6785):486-489.
75. Prokhortchouk A, Hendrich B, Jorgensen H et al. The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. *Genes Dev* 2001; 15(13):1613-1618.
76. Daniel JM, Spring CM, Crawford HC et al. The p120(ctn)-binding partner Kaiso is a bi-modal DNA-binding protein that recognizes both a sequence-specific consensus and methylated CpG dinucleotides. *Nucleic Acids Res* 2002; 30(13):2911-2919.

Zinc Finger Proteins

From Atomic Contact to Cellular Function

Iuchi, S.; Kuldell, N. (Eds.)

2005, XVI, 276 p. 170 illus., 19 illus. in color., Hardcover

ISBN: 978-0-306-48229-8