Human zinc fingers as building blocks in the construction of artificial transcription factors

Kwang-Hee Bae[1,†], Young Do Kwon[1,2,†], Hyun-Chul Shin[1], Moon-Sun Hwang[1], Eun-Hyun Ryu[1], Kyung-Soon Park[1], Hyo-Young Yang[1], Dong-ki Lee[1], Yangsoon Lee[1], Jinwoo Park[1], Heung Sun Kwon[1], Hyun-Won Kim[3], Byung-Il Yeh[3], Hyean-Woo Lee[3], Soon Hyung Sohn[3], Joonho Yoon[3], Wongi Seol[1], and Jin-Soo Kim[1,*]

Published online 18 February 2003; doi:10.1038/nbt796

We describe methods for generating artificial transcription factors capable of up- or downregulating the expression of genes whose promoter regions contain the target DNA sequences. To accomplish this, we screened zinc fingers derived from sequences in the human genome and isolated 56 zinc fingers with diverse DNA-binding specificities. We used these zinc fingers as modular building blocks in the construction of novel, sequence-specific DNA-binding proteins. Fusion of these zinc-finger proteins with either a transcriptional activation or repression domain yielded potent transcriptional activators or repressors, respectively. These results show that the human genome encodes zinc fingers with diverse DNA-binding specificities and that these domains can be used to design sequence-specific DNA-binding proteins and artificial transcription factors.

Genes are regulated at the transcriptional level by transcription factors that interact with specific DNA regulatory elements to activate or inhibit transcription initiation. Many transcriptional factors have modular structures that consist of a DNA-binding domain and an effector (activation or repression) domain. The effector domains of many transcription factors remain active when transferred to heterologous DNA-binding domains[1]. This suggests that it may be possible to construct novel transcription factors by preparing DNA-binding domains with the desired DNA-binding specificities and then linking them to the appropriate effector domains.

A wide range of organisms has transcription factors containing zinc-finger motifs, which bind to DNA in a sequence-specific manner. There are several types of zinc fingers, but members of the Cys_2-His_2 class are ideal for generating artificial transcription factors owing to their diversity and modular structure. For example, novel DNA-binding proteins have been created by using phage display to alter the DNA-binding specificity of a zinc-finger protein (ZFP)[2–7]. Although phage display is an effective method, it is labor-intensive and time-consuming. Moreover, because phage display uses *in vitro* selection, the selected ZFPs might not function efficiently *in vivo*. With phage display, ZFPs are selected by binding to 'naked' DNA rather than to chromatin, whereas in eukaryotic cells DNA is packaged into nucleosomes to form chromatin.

To circumvent these problems, we developed an approach for constructing DNA-binding proteins that recognize the desired DNA sites in eukaryotic cells. We isolated zinc-finger domains with diverse DNA-binding specificities by screening, in yeast, plasmid libraries that encode zinc fingers derived from the human genome. These zinc fingers were then used as modular building blocks to construct novel DNA-binding proteins and, subsequently, designer transcription factors.

Results

Zinc fingers derived from the human genome. We modified the yeast one-hybrid system[8–10] to select zinc fingers that exhibit diverse DNA-binding specificities. This system includes two essential components: a reporter plasmid and a plasmid that encodes a hybrid transcription factor (Fig. 1A). The hybrid transcription factor consists of two domains: a DNA-binding domain and a transcriptional activation domain. We used the Gal4 transcriptional activation domain fused to the Zif268 protein as our hybrid transcription factor. Zif268 is a zinc-finger protein that contains three zinc-finger domains; together, the three-finger protein binds specifically to a 9 base pair (bp) DNA sequence (5′-GCG TGG GCG-3′). Each zinc finger of Zif268 recognizes a 3 bp subsite within the 9 bp compound binding site.

To isolate zinc fingers with diverse DNA-binding specificities, we cloned human DNA segments that encode zinc fingers. The DNA segments were amplified using PCR. The PCR products were then cloned into a yeast expression plasmid that contained the Gal4-Zif268 hybrid transcription factor. In this library, zinc finger 3 (the C-terminal finger) of Zif268 was replaced with zinc-finger sequences amplified from the human genome (Fig. 1A).

We used both yeast *HIS3* and *Escherichia coli LacZ* as reporter genes. The reporter plasmids were constructed by inserting three copies of the target DNA sequences adjacent to the coding regions of the reporter genes. The target sequences have the format 5′-XXX TGG GCG-3′, where XXX corresponds to a 3 bp target subsite and TGG GCG corresponds to the DNA sequences recognized by finger 1 (the N-terminal finger) and finger 2 (the middle finger) of Zif268 (Fig. 1A). In total, two sets of 64 different reporter plasmids were prepared (one set for each of the two reporter genes).

We identified zinc fingers with diverse DNA-binding specificities by isolating yeast cells in which transcription of the reporter genes

[1]*ToolGen Inc., 461-6 Jeonmin-Dong, Yusung-Gu, Daejeon, 305-390, South Korea.* [2]*School of Biological Sciences, Seoul National University, Seoul, 151-742, South Korea.* [3]*Department of Biochemistry and Institute of Basic Medical Sciences and IFBB, Yonsei University, Wonju College of Medicine, Wonju, 220-701, South Korea.* **Corresponding author (e-mail: jsk@toolgen.com).* *[†]These authors contributed equally to this work.*

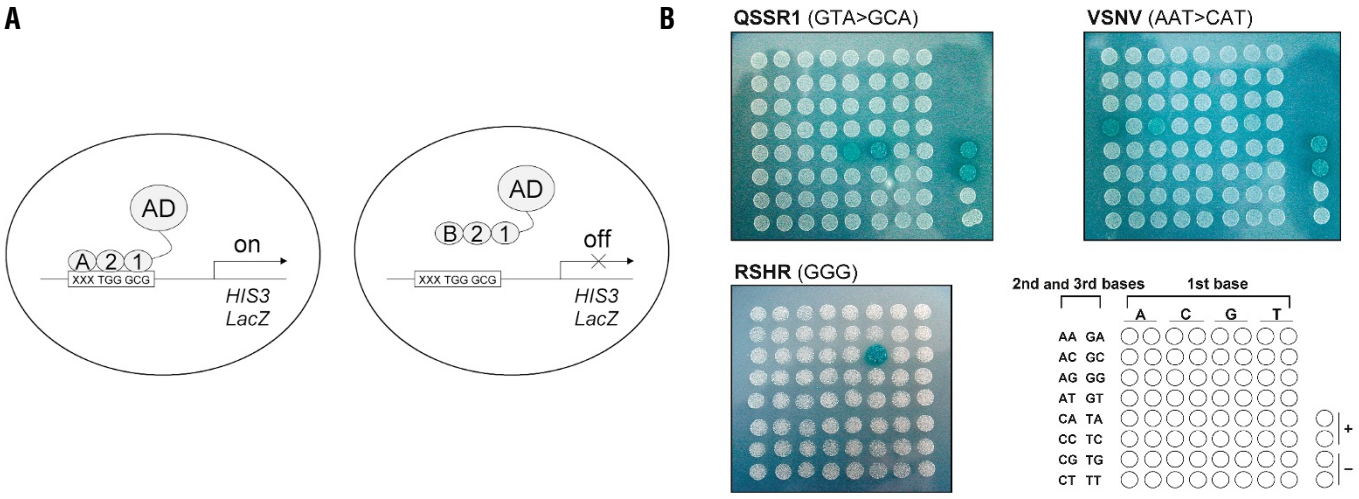

Figure 1. *In vivo* screening system and zinc 'fingerprints'. (A) Zinc-finger screening system in yeast. Zinc-finger domain 'A' recognizes the 3-bp target DNA subsite (designated XXX) upstream of the reporter gene, and transcription of the *HIS3* or *LacZ* reporter gene is activated. As a result, the yeast grows on a medium lacking histidine and forms a blue colony when grown on a medium containing X-gal (see Fig. 1B). In contrast, zinc-finger domain 'B' does not recognize the target subsite. In this case, the reporter gene is not activated; thus the yeast does not grow on a medium lacking histidine and forms a white colony on a medium containing X-gal (see Fig. 1B). AD, transcriptional activation domain; 1, zinc finger 1 of Zif268; 2, zinc finger 2 of Zif268. (B) Zinc fingerprints. The DNA-binding specificities of selected zinc fingers (QSSR1, VSNV, and RSHR) that had been fused to fingers 1 and 2 of Zif268 were determined in yeast cells using 64 LacZ reporter sets, each of which contained one of the 64 triplet subsites. '+' indicates a positive control, which contains the third finger of Zif268 and a reporter plasmid containing its subsite sequence, GCG. '–' indicates a negative control, which contains only the first and second fingers of Zif268 and a control reporter plasmid with no binding sequence. The sequences of the functional 3 bp subsites for each of the ZFPs are shown in parentheses after the protein name, but can also be deduced from the positions of the blue colonies using the key in the bottom right. 'Fingerprints' for other zinc fingers are shown in Supplementary Figure 4 online.

was activated. The amino acid sequence of each selected zinc finger was deduced from the DNA sequence, and each zinc finger was named using the single-abbreviation of the four amino acid residues at positions –1, 2, 3, and 6 in the α-helix of the zinc finger. These amino acid residues are expected to make contact with bases in the target DNA subsite[11–15].

The DNA-binding specificities of the selected zinc fingers in yeast cells were determined using the 64 *LacZ* reporter plasmids, each of which contains one of the 64 distinct triplet subsites within the 9 bp target DNA sequences (Fig. 1B). Many of the selected zinc fingers showed a unique DNA-binding signature. For example, cells harboring the plasmid that encodes the QSSR1 zinc finger fused to fingers 1 and 2 of the Zif268 protein and the Gal4 activation domain formed blue colonies only when mated with cells that contained the GTA and GCA reporter plasmids. This means that, under the conditions of the assay, the QSSR1 finger recognizes the GTA and GCA triplets but not the other 62 triplet subsites. After screening ~2,000 zinc fingers present in our library, we obtained a total of 56 zinc fingers that showed distinct DNA-binding specificities in our yeast system (Table 1). As a group, these zinc fingers recognize 25 subsites.

For selected fingers, we compared the amino acid residues at the critical base-contacting positions with those expected from the zinc

Table 1. Lists, target subsites, and K_d values of selected human zinc-finger domains[a]

Zinc fingers	Target subsite	K_d value[b] (nM)	Zinc fingers	Target subsite	K_d value (nM)	Zinc fingers	Target subsite	K_d value (nM)
CSNR1[c]	GAV	0.13 (GAC)	**QSNI**	MAA	2.9 (CAA)	**RDER1[c]**	GHG	0.078 (GCG)
CSNR2	GAV	ND	**QSNR1[c]**	GAA	0.059	**RDER2**	GCG	ND
DSAR	RTC	0.021 (GTC)	**QSNR2**	GAA	ND	**RDER3**	GCG	ND
DSCR	GCC	0.22	**QSNR3v**	GAA	ND	**RDER4**	GCG	ND
HSNK[c]	RAC	10 (GAC)	**QSNR4**	GAA	ND	**RDER5**	GCG	ND
HSSR	GTT	0.043	**QSNK**	DAA	0.15 (GAA)	**RDER6**	GCG	ND
ISNR[c]	GAW	0.0062 (GAT)	**QSNT**	AAA	ND	**RDHR1**	GGG	0.049
ISNV	AAT	ND	**QSNV1**	MAA	ND	**RDHR2**	GGG	ND
KSNR	GAG	0.094	**QSNV2[c]**	NAA	0.23 (CAA)	**RDHT[c]**	NGG	0.38 (AGG)
QAHR	GGA	0.47	**QSNV3**	MAA	ND	**RDKI**	GGG	ND
QFNR	GAG	3.7	**QSNV4**	MAA	ND	**RDKR**	RGG	0.25 (GGG)
QGNR	GAA	0.069	**QSSR1[c]**	GYA	0.046 (GTA)	**RSHR[c]**	GGG	0.010
QSHR1	RGA	ND	**QSSR2**	GYA	ND	**RSNR[c]**	GAG	0.040
QSHR2[c]	GGA	ND	**QSSR3**	GYA	ND	**SSNR**	GAG	0.075
QSHR3[c]	RGA	ND	**QSTR**	GYA	0.051 (GTA)	**VSNV[c]**	AAT	0.14
QSHR4	GGA	ND	**QSTV**	GAV	ND	**VSSR[c]**	GTD, GCD	0.12 (GTG)
QSHR5	GRA, MGA	0.11 (GGA)	**QTHQ[c]**	HGA	0.034 (CGA)	**VSTR[c]**	GCW	0.81 (GCT)
QSHT[c]	HGA	0.96 (AGA)	**QTHR1**	VGA, GAA	0.089 (GGA)	**WSNR**	GGW	0.073 (GGT)
QSHV	HGA	3.6 (CGA)	**QTHR2**	GGA	ND	**Zif268**	GCG	0.056

[a]The sites in parentheses were used to determine the K_d value of the corresponding ZFPs.
[b]The K_d values were determined by EMSA using purified ZFPs, each of which was composed of finger 1 and 2 of Zif268 and a zinc-finger of interest.
[c]Selected zinc-finger domains used in construction of three-finger proteins.
ND, not determined; W, A or T; R, A or G; H, A or C or T; M, A or C; V, A or C or G; D, A or G or T; Y, C or T; N, A or C or G or T.

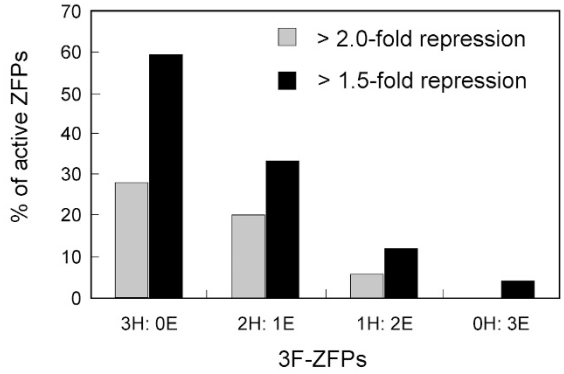

Figure 2. Analyses of chimeric zinc-finger proteins assembled by domain shuffling. The 153 ZFPs were constructed as described in the text. The plasmids encoding each of these ZFPs were cotransfected into HEK293 cells with corresponding reporter plasmids. The reporter plasmids carried the firefly luciferase gene under the control of synthetic promoters in which appropriate binding sites are incorporated at positions near the initiator element (see Supplementary Fig. 3 online). Repression levels (fold repression) were obtained by dividing the luciferase activities from the cells transfected with the empty plasmid by those from the cells transfected with plasmids encoding ZFPs. 'Active ZFPs' were defined as those that showed greater than 1.5-fold repression (low stringency) or 2.0-fold repression (high stringency). The chimeric ZFPs are divided into four groups based on their domain composition (H, human; E, engineered): '3H: 0E', ZFPs composed of three human domains and no engineered domain; '2H: 1E', two human domains and one engineered domain; '1H: 2E', one human domain and two engineered domains; and '0H: 3E', no human domain and three engineered domains. For each group, the percentages of ZFPs that exhibited repression activity of more than 2.0- or 1.5-fold are shown. 3F-ZFPs, three-finger ZFPs.

finger–DNA recognition code described previously[11–13]. Amino acid residues at positions –1, 3, and 6 in the α-helix of zinc fingers play important roles in base-specific interactions[14–15]. Although this code is of limited utility in designing new site-specific DNA-binding proteins[16–17], we found it quite useful for analyzing the selected zinc fingers. Overall, the amino acid residues at the critical base-contacting positions of the selected zinc fingers matched those predicted by the code in at least two of the three critical positions.

We also performed electrophoretic mobility shift assays (EMSAs) to characterize binding of the selected zinc fingers to DNA *in vitro* (Table 1). The recombinant ZFPs expressed in and purified from *E. coli* contained fingers 1 and 2 of Zif268 and a selected zinc finger at the C terminus. ZFPs that allowed yeast cells to grow on histidine-deficient plates and produced blue colonies in our yeast assays bound tightly to the corresponding probe DNAs *in vitro*. Most of the dissociation constants were within the sub-nanomolar range. A few ZFPs, such as the HSNK, QFNR, QSHV, and QSNI proteins, bound with lower affinity to their target sequences, with dissociation constants in the nanomolar range. However, these proteins still showed at least a 50-fold preference for specific target DNA sequences when compared with binding to random DNA sequences (data not shown).

Zinc-finger domain shuffling to construct new DNA-binding proteins. One of the features that make zinc-fingers ideal for protein engineering is their modular relationship with each other. Our next goal was to test whether the human zinc fingers isolated using our *in vivo* screening could function in a modular fashion to assemble new DNA-binding proteins. We also sought to ascertain how these naturally occurring zinc fingers would fare relative to engineered zinc fingers created by replacing critical base-contacting residues. We constructed three types of three-finger chimeric ZFPs. (i) ZFPs composed exclusively of the human zinc fingers described above. We selected eighteen zinc-finger domains after considering each domain's DNA-binding specificity and affinity (see Supplementary Table 1 online). (ii) ZFPs composed exclusively of engineered zinc fingers. We chose 18 engineered domains (see Supplementary Table 1 online) that had been selected by phage display and thoroughly characterized[6,18–21]. (iii) Hybrid ZFPs composed of both the 18 human domains and the 18 engineered domains. (The engineered zinc fingers described in refs. 22 and 23 were not included in our experiments because they were not available at the time. Therefore, the engineered zinc fingers we used may not represent the best of those currently available.)

To generate the expression plasmids for these ZFPs, we prepared DNA that encodes the chimeric three-finger proteins described above by two different methods. In the first method, nucleic acids encoding the 36 selected zinc fingers were mixed, and three-finger constructs were randomly picked for further analysis. (A total of 33 constructs were prepared using this method.) Each construct was sequenced to determine which zinc-finger domains it encoded, and then a target DNA sequence was synthesized for each randomly assorted three-finger protein. In the second method, segments of DNA encoding various zinc-finger domains were assembled to form chimeric three-finger proteins; the three fingers selected were those most likely to bind to target DNA sequences in the promoter regions of known genes. (A total of 120 constructs were prepared using this method.) A computer algorithm was used to match the preferred recognition sites for the individual zinc fingers with naturally occurring target DNA sequences. Zinc fingers that included an aspartate residue at position 2 were analyzed with special consideration because this residue can form a hydrogen bond with a base outside the 3 bp subsite recognized by the zinc fingers[14,15]. The computer algorithm accounted for this additional specificity. If a randomly assembled three-finger protein with an aspartate at position 2 violated the rule for interaction with a base outside the 3-bp subsite, we excluded the protein from our analyses.

We assessed the function of these proteins in transient cotransfection assays[24–26]. In these assays, ZFPs efficiently repressed VP16-activated transcription of a reporter gene when the ZFPs bound to DNA sequences near the transcription start site. We used a similar *in vivo* repression assay to determine whether the newly constructed

Table 2. Binding sites and identity of ZFPs used in *VEGF* activation

ZFP ID	Binding position[a]	F1	F2	F3	Target sequence (from 5′ to 3′)[b]	
				Intended	Experimentally determined[d]	
F480-P	–633R[c]	RSHR	RDHT	RSHR	GGG TGG GGG	GGG TGG GGG
F475-V	–455	QSHR2	RDHT	RSNR	GAG CGG GGA	n.d.
F435-P	–90R[c], –391R[c]	QSHR2	RDHT	RSHR	GGG TGG GGA	GGG TGG GGA
F121-P	+434	QSHT	RSHR	RDHT	TGG GGG TGA	TGG GGG CGA
F83-P	+359	RDHT	QSHT	RSNR	GAG AGA CGG	GAG AGA CGG

[a]Numbers in the binding position indicate a 3′ end of the binding site relative to transcriptional initiation sequence. [b]'Intended' sequences are those present in the VEGF promoter and 'experimentally determined' sequences are the consensus sequences obtained from SELEX experiments. [c]R, reverse strand. [d]Underlined letters indicate that the corresponding ZFPs recognize more than one base as shown in Table 1. The SELEX data also showed more heterogeneity in these positions.

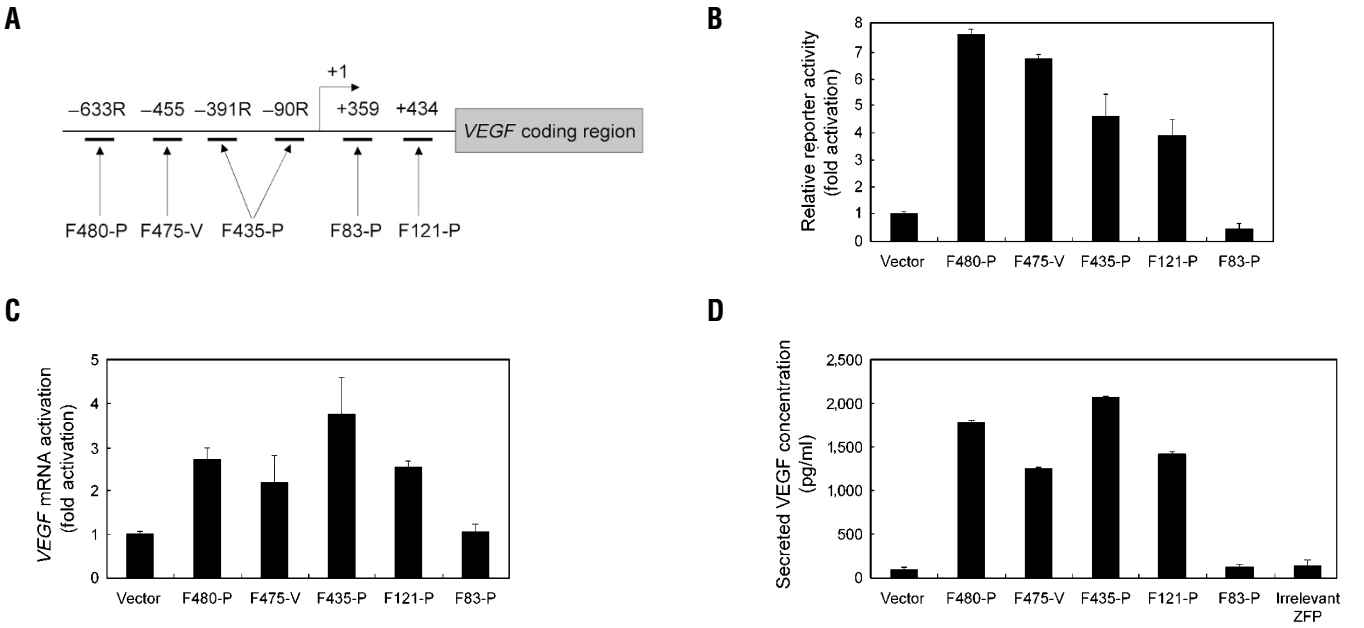

Figure 3. Regulation of expression of the endogenous *VEGF* gene by ZFPs assembled by domain shuffling. (A) Binding sites in the *VEGF* promoter for the ZFPs used in this study. '+1' indicates the transcription initiation site; 'P' and 'V' represent the p65 and VP16 activation domains, respectively, which had been fused individually to the sequence of the C terminus of the ZFPs. (B) Activation of the luciferase reporter gene under the control of the *VEGF* promoter (nucleotides –950 to +450) by ZFPs. Plasmids encoding ZFP activators were co-transfected into HEK 293 cells along with a reporter construct encoding the firefly luciferase gene under the control of the *VEGF* promoter. Luciferase activity was measured 2 d after transfection. (C) Activation of the endogenous *VEGF* gene by the indicated ZFP activators. We measured, by quantitative RT-PCR, the relative levels of *VEGF* mRNA in HEK 293 cells that had been transiently transfected with the corresponding ZFPs. (D) Amounts of VEGF protein secreted from HEK 293 cells that had been transiently transfected with the corresponding ZFPs; measured with an ELISA assay kit (Chemicon, Temecula, CA). 'Vector' indicates a control vector that encode no ZFPs. An 'irrelevant ZFP' was also used as a negative control.

three-finger proteins were functional in human cells. The assay used luciferase reporter constructs that had the appropriate target DNA sequences inserted near the transcription start site.

Our results showed that the human (H) zinc fingers were much better building blocks for constructing new DNA-binding proteins than were the engineered (E) domains (Fig. 2 and Supplementary Table 2 online). Out of 96 ZFPs that were composed exclusively of human domains ('3H: 0E'), 28% of the proteins were highly active (that is, they yielded greater than 2.0-fold repression), and 59% showed greater than 1.5-fold repression. In contrast, only 4% (1 out of 25) of the ZFPs composed exclusively of engineered domains ('0H: 3E') showed greater than 1.5-fold repression, and none of these ZFPs showed greater than 2.0-fold repression. Results with the hybrid proteins, which were composed of both human and engineered domains ('2H: 1E' and '1H: 2E'), fell between those with the all-human and all-engineered ZFPs; only 20% ('2H: 1E') and 6% ('1H: 2E') of these hybrid ZFPs showed greater than 2.0-fold repression, and 33% ('2H: 1E') and 12% ('1H: 2E') showed greater than 1.5-fold repression. These results clearly indicate that human zinc fingers selected *in vivo* are more useful than mutated domains when the domain shuffling method is used to assemble new ZFPs. (We ruled out the possibility that ZFPs composed of human zinc fingers were simply expressed at higher levels in human cells than were those composed of engineered domains by performing western blot analysis with representative ZFPs (see Supplementary Fig. 1 online).)

Why are the human domains better building blocks than the engineered domains? Possible answers include the following. (i) The *in vivo* selection conditions under which we screened the human zinc fingers might have been more stringent than the phage display conditions under which the engineered fingers were selected. In this regard, we note that several of the engineered domains, when tested in our *in vivo* yeast assay, did not activate the reporter genes (data not shown). (ii) Most of the human zinc fingers contain a serine residue at position 2, and this may account for their enhanced ability to serve as functional modules in multifinger ZFPs. A serine residue at position 2 might be able to interact with any of the four bases positioned adjacent to the 3-bp subsite, because serine can function as a hydrogen-bond donor or acceptor. Unlike most of the human zinc fingers, engineered domains contain diverse amino acid residues at position 2. This may limit their utility as modular building blocks.

Designer transcription factors composed of human zinc fingers. To determine whether ZFPs composed of human zinc fingers can regulate transcription of an endogenous gene, we assembled ZFPs that bind to elements in the human vascular endothelial growth-factor (*VEGF*) gene. This gene was chosen because it has been successfully upregulated using engineered ZFPs constructed by site-specific mutagenesis[27]. The *VEGF* promoter sequence that corresponded to the region from nucleotides –950 to +450 was scanned to identify sites nine nucleotides long that are potential target sites for recognition by three-finger ZFPs. Five DNA constructs encoding these ZFPs were assembled with the domain shuffling method described above. The proteins were expressed in *E. coli*, purified, and used in systematic evolution of ligands by exponential enrichment (SELEX) analyses to determine their DNA-binding specificities. All four ZFPs designed to target the *VEGF* promoter showed the expected DNA-binding specificities (Table 2).

We then generated artificial transcription factors by fusing the three-finger domains to either the p65 or VP16 activation domains and inserting this construct into an expression plasmid. Fig. 3A shows the locations of binding sites for these ZFPs in the *VEGF* promoter. The human *VEGF* promoter contains two DNase I–hypersensitive regions that are accessible to transcriptional factors; binding of engineered ZFP transcription factors to these sites activates *VEGF* gene expression[27]. We found that regardless of the location of the ZFP

binding sites, the four ZFPs we tested activated not only the luciferase reporter gene under the control of the *VEGF* promoter (Fig. 3B), but also the endogenous *VEGF* gene itself (Fig. 3C). Also, an enzyme-linked immunosorbent assay (ELISA) performed with medium from the transiently transfected cells indicated that these ZFPs upregulated production of the VEGF protein 13- to 21-fold (Fig. 3D). Control cells that had been transfected with the control plasmid (which contained no ZFP-coding sequences) did not show any change in *VEGF* mRNA or protein levels. One ZFP, termed F83, did not affect the levels of *VEGF* mRNA or protein in these assays (Figs. 3C,D). This may be due to the binding of some other protein at the ZFP target site or to the local chromatin structure, which might have rendered the target site inaccessible to the ZFP. There was no strong correlation between the levels of *VEFG* expression by these ZFPs and their DNA-binding affinities or their expression levels in cells (data not shown).

To investigate the specificity of ZFP transcription factors on a genome-wide scale, we performed DNA microarray experiments with 293 cell lines that had been stably transfected with DNA constructs encoding each of the following three zinc-finger transcription factors: F121-p65, F435-p65, and F475-VP16. Fifty-one out of 7,458 genes were co-regulated (49 were upregulated and 2 were downregulated more than twofold) by all three ZFP transcriptional activators (Supplementary Table 3 online). Most of these co-regulated genes seem to be closely associated with VEGF function. Many are regulated by VEGF, involved in angiogenesis or hypoxia, or expressed in vascular endothelial cells. Therefore, it is likely that these genes are downstream targets of VEGF. We note, however, that dozens of other genes were regulated by one or two of our ZFP activators but not by all three (data not shown). Considering that these ZFPs that recognize nine bp sites, it is possible that they regulate genes other than VEGF. The use of four-, five-, or six-finger proteins should help to improve the specificity. Taken together, these data indicate that our ZFPs, assembled by shuffling naturally occurring zinc-finger domains, function in cells as transcriptional regulators of specific genes.

Discussion

Two approaches exist for making DNA-binding proteins that bind to predetermined DNA sequences. In one, site-directed mutagenesis is used to replace amino acid residues at a few key positions with residues that interact with the desired bases in the DNA target sequence[27–29]. The zinc finger–DNA 'recognition code', compiled from extensive biochemical and structural studies by many groups[11–13], has made this approach practical. The other approach is based on zinc-finger phage display[2–7]. ZFPs with altered DNA-binding specificities can be selected from a library of zinc-finger variants in which the amino acid residues at key positions are randomly mutated.

Our approach, termed GeneGrip, differs from these approaches in several ways. (i) Our selection and screening procedures[8–10] are done *in vivo*. Since genomic DNA is packaged as chromatin in eukaryotic cells, we reasoned that ZFPs selected *in vivo* may be more useful for regulating gene expression in higher eukaryotic cells than those selected *in vitro* (using phage display) or in prokaryotic cells[30]. Indeed, using our system we were able to construct ZFPs that functioned efficiently in human cells. (ii) Unlike previous approaches, in which mutations were incorporated at key positions in a given zinc-finger framework[27,31–33], our approach uses intact, wild-type zinc fingers derived from DNA sequences in the human genome. A naturally occurring linker sequence was used to connect individual zinc fingers. ZFPs composed of human zinc fingers may be preferable in therapeutic applications because they would be less likely to induce a host immune response than would proteins composed of mutated zinc fingers (although an immune response involving the linkers between the ZFP moieties and the effector domains is still possible.) (iii) Unlike previous protocols such as phage display selection, our shuffling approach is easily scalable. Thousands of highly active ZFPs can be constructed simultaneously. For example, shuffling 20 domains to make three-finger proteins would yield 8,000 ($= 20 \times 20 \times 20$) ZFPs in a single step.

Designer zinc-finger transcription factors generated using the GeneGrip technique could be used to regulate the expression of as yet uncharacterized genes *in vivo*, an approach that could help in identifying new genes and determining their functions. In addition, a ZFP library could be screened to isolate improved phenotypes induced by specific ZFPs. Our technology should thus find wide application in basic research, medicine, and biotechnology.

Experimental protocol

Construction of a yeast expression plasmid for a zinc-finger library. We constructed an expression plasmid (termed pPCFM-Zif) encoding a zinc-finger transcription factor by modification of pPC86 (ref. 34). To construct plasmids encoding human zinc-finger libraries, DNA segments from human genomic DNA encoding zinc fingers were amplified by PCR and the 100 bp PCR products were inserted into pPCFM-Zif. The plasmid library was prepared from a total of 1.2×10^6 *E. coli* transformants. Gap-repair cloning[35] was also used to construct plasmids that encode individual zinc-finger domains. The construction of these plasmids are described in detail in the Supplementary Experimental Protocol online.

Reporter plasmids were prepared by inserting one of 64 pairs of complementary oligonucleotides (primer 4 in Supplementary Table 4 online) that contained three copies of a 9-bp target sequence into pRS315His (a gift from R. Reed at Johns Hopkins University's School of Medicine, Baltimore, MD) and pLacZi (Clontech, Palo Alto, CA).

***In vivo* selection of zinc-finger domains.** Yeast mating was used to facilitate identification of zinc fingers that bind to each 3-bp target site. The binding affinity and specificity of each zinc finger fused to fingers 1 and 2 of Zif268 were determined both in yeast and by EMSA. These methods are described in detail in Supplementary experimental protocol online.

Construction of three-finger proteins using selected zinc fingers as modular building blocks. A mammalian expression plasmid, pcDNA3 (Invitrogen, Carlsbad, CA), was used as a parental vector for expressing ZFPs in mammalian cells. DNA segments that encode individual zinc-finger domains were subcloned into the plasmid, and the resulting plasmids were used as starting material for ZFP construction. The scheme for constructing plasmids that encode new three-finger proteins is outlined in Supplementary Figure 2 online. The constructed ZFPs were tested for their DNA binding ability and affinity in mammalian cells as described previously[24–26]. The reporter plasmid for the assay was constructed using pGL3-TATA/Inr[24–26] which harbors the firefly luciferase gene as the reporter. The sequence of and method for making the reporter are described in Supplementary Figure 3 online. In addition, SELEX was performed to test whether these proteins recognize the appropriate target DNA sequences. These methods are described in detail in Supplementary Experimental Protocol online.

Note: Supplementary information is available on the Nature Biotechnology website.

Acknowledgments

We thank Randall Reed for providing yeast strain yWAM2 and plasmids pRS315HIS and pPC86. We also thank S.I. Lee, S.I. Kim, J.W. Kim, and H.R. Lee for technical assistance and K. LaMarco for carefully reading our manuscript. This study was supported by a grant of the National Research Laboratory Program from the Ministry of Science & Technology, Republic of Korea (M1-0104-00-0048).

Competing interests statement

The authors declare competing financial interests: see the Nature Biotechnology website (http://www.nature.com/naturebiotechnology) for details.

Received 6 August 2002; accepted 3 January 2003

1. Brent, R. & Ptashne, M.A. Eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. Cell **43**, 729–736 (1985).
2. Rebar, E.J. & Pabo, C.O. Zinc-finger phage: affinity selection of fingers with new DNA-binding specificities. *Science* **263**, 671–673 (1994).
3. Jamieson, A.C., Kim, S.-H. & Wells, J.A. *In vitro* selection of zinc fingers with altered DNA-binding specificity. *Biochemistry* **33**, 5689–5695 (1994).
4. Choo, Y. & Klug, A. Toward a code for the interactions of zinc fingers with DNA: selection of randomized fingers displayed on phage. *Proc. Natl. Acad. Sci.* **91**, 11163–11167 (1994).
5. Wu, H., Yang, W.P. & Barbas, C.F. III. Building zinc fingers by selection: toward a therapeutic application. *Proc. Natl. Acad. Sci. USA* **92**, 344–348 (1995).
6. Greisman, H.A. & Pabo, C.O. A general strategy for selecting high-affinity zinc-finger proteins for diverse DNA target sites. *Science* **275**, 657–661 (1997).
7. Isalan, M., Klug, A. & Choo, Y. A rapid, generally applicable method to engineer zinc fingers illustrated by targeting the HIV-1 promoter. *Nat. Biotechnol.* **19**, 656–660, (2001).
8. Wang, M.M. & Reed, R.R. Molecular cloning of the olfactory neuronal transcription factor Olf-1 by genetic selection in yeast. *Nature* **364**, 121–126 (1993).
9. Cheng, X., Boyer, J.L. & Juliano, R.L. Selection of peptides that functionally replace a zinc finger in the Sp1 transcription factor by using a yeast combinatorial library. *Proc. Natl. Acad. Sci. USA* **94**, 14120–14125 (1997).
10. Bartsevich, V.V. & Juliano, R.L. Regulation of the MDR1 gene by transcriptional repressors selected using peptide combinatorial libraries. *Mol. Pharmacol.* **58**, 1–10 (2000).
11. Desjarlais, J.R. & Berg, J.M. Redesigning the DNA-binding specificity of a zinc-finger protein: a data base–guided approach. *Proteins Struct. Funct. Genet.* **12**, 101–104 (1992).
12. Choo, Y. & Klug, A. Physical basis of a protein-DNA recognition code. *Curr. Opin. Struct. Biol.* **7**, 117–125 (1997).
13. Wolfe, S.A., Greisman, H.A., Ramm, E.I. & Pabo, C.O. Analysis of zinc fingers optimized via phage display: evaluating the utility of a recognition code. *J. Mol. Biol.* **285**, 1917–1934 (1998).
14. Pavletich, N.P. & Pabo, C.O. Zinc finger–DNA recognition: crystal structure of a Zif268-DNA complex at 2.1Å. *Science* **252**, 809–817 (1991).
15. Elrod-Erickson, M., Benson, T.E. & Pabo, C.O. High-resolution structures of variant Zif268-DNA complexes: implications for understanding zinc finger–DNA recognition. *Structure.* **6**, 451–464 (1998).
16. Wolfe, S.A., Grant, R.A., Elrod-Erickson, M. & Pabo, C.O. Beyond the 'recognition code': structures of two Cys2His2 zinc finger/TATA box complexes. *Structure* **8**, 717–723 (2001).
17. Pabo, C.O., Peisach, E. & Grant, R.A. Design and selection of novel Cys2His2 zinc-finger proteins. *Annu. Rev. Biochem.* **70**, 313–340 (2001).
18. Segal, D.J., Dreier, B., Beerli, R.R. & Barbas, C.F. III. Toward controlling gene expression at will: selection and design of zinc-finger domains recognizing each of the 5′-GNN-3′ DNA target sequences. *Proc. Natl. Acad. Sci. USA* **96**, 2758–2763 (1999).
19. Dreier, B., Beerli, R.R., Segal, D.J., Flippin, J.D. & Barbas, C.F. III. Development of zinc-finger domains for recognition of the 5′-ANN-3′ family of DNA sequences and their use in the construction of artificial transcription factors. *J. Biol. Chem.* **276**, 29466–29478 (2001).
20. Choo, Y. & Klug, A. Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions. *Proc. Natl. Acad. Sci. USA* **91**, 11168–11172 (1994).
21. Desjarlais, J.R. & Berg, J.M. Length-encoded multiplex binding site determination: application to zinc-finger proteins. *Proc. Natl. Acad. Sci. USA* **91**, 11099–11103 (1994).
22. Liu, Q., Xia, Z. & Case, C.C. Validated zinc-finger protein designs for all 16 GNN DNA triplet targets. *J. Biol. Chem.* **277**, 3850–3856 (2002).
23. Sera, T. & Uranga, C. Rational design of artificial zinc-finger proteins using a nondegenerate recognition code table. *Biochemistry* **41**, 7074–7081 (2002).
24. Kim, J.-S. & Pabo, C.O. Transcriptional repression by zinc-finger peptides: exploring the potential for applications in gene therapy. *J. Biol. Chem.* **272**, 29795–29800 (1997).
25. Kim, J.-S. & Pabo, C.O. Getting a handhold on DNA: design of poly-zinc-finger proteins with femtomolar dissociation constants. *Proc. Natl. Acad. Sci. USA* **95**, 2812–2817 (1998).
26. Kang, J.S. & Kim, J.-S. Zinc-finger proteins as designer transcription factors. *J. Biol. Chem.* **275**, 8742–8748 (2000).
27. Liu, P.Q. *et al.* Regulation of an endogenous locus using a panel of designed zinc-finger proteins targeted to accessible chromatin regions. *J. Biol. Chem.* **276**, 11323–11334 (2001).
28. Nardelli, J., Gibson, T. & Charnay, P. Zinc finger–DNA recognition: analysis of base specificity by site-directed mutagenesis. *Nucleic Acids Res.* **20**, 4137–4144 (1992).
29. Taylor, W.E. *et al.* Designing zinc-finger ADR1 mutants with altered specificity of DNA binding to T in UAS1 sequences. *Biochemistry* **34**, 3222–3230 (1995).
30. Joung, J.K., Ramm, E.I. & Pabo, C.O. A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. *Proc. Natl. Acad. Sci. USA* **97**, 7382–7387 (2000).
31. Beerli, R.R., Dreier, B. & Barbas, C.F. III. Positive and negative regulation of endogenous genes by designed transcription factors. *Proc. Natl. Acad. Sci. USA* **97**, 1495–1500 (2000).
32. Zhang, L. *et al.* Synthetic zinc-finger transcription factor action at an endogenous chromosomal site. Activation of the human erythropoietin gene. *J. Biol. Chem.* **275**, 33850–33860 (2000).
33. Ren, D., Collingwood, T.N., Rebar, E.J., Wolffe, A.P. & Camp, H.S. PPARγ knockdown by engineered transcription factors: exogenous PPARγ2 but not PPARγ1 reactivates adipogenesis. *Genes Dev.* **16**, 27–32 (2002).
34. Chevray, P.M. & Nathans, D. Protein interaction cloning in yeast: identification of mammalian proteins that interact with the leucine zipper of Jun. *Proc. Natl. Acad. Sci. USA* **89**, 5789–5793 (1992).
35. Hudson, J.R. *et al.* The complete set of predicted genes from *Saccharomyces cerevisiae* in a readily usable form. *Genome Res.* **7**, 1169–1173, (1997).