PROSPECTS
Journal of Cellular Biochemistry 116:2435–2444 (2015)
Journal of Cellular Biochemistry

Artificial Zinc Finger DNA Binding Domains: Versatile Tools for Genome Engineering and Modulation of Gene Expression

Mir A. Hossain, Joeva J. Barrow, Yong Shen, MD Imdadul Haq, and Jörg Bungert*
Department of Biochemistry and Molecular Biology, College of Medicine, Cancer Center, Genetics Institute, University of Florida, Gainesville, Florida 32610

ABSTRACT
Genome editing and alteration of gene expression by synthetic DNA binding activities gained a lot of momentum over the last decade. This is due to the development of new DNA binding molecules with enhanced binding specificity. The most commonly used DNA binding modules are zinc fingers (ZFs), TALE-domains, and the RNA component of the CRISPR/Cas9 system. These binding modules are fused or linked to either nucleases that cut the DNA and induce DNA repair processes, or to protein domains that activate or repress transcription of genes close to the targeted site in the genome. This review focuses on the structure, design, and applications of ZF DNA binding domains (ZFDBDs). ZFDBDs are relatively small and have been shown to penetrate the cell membrane without additional tags suggesting that they could be delivered to cells without a DNA or RNA intermediate. Advanced algorithms that are based on extensive knowledge of the mode of ZF/DNA interactions are used to design the amino acid composition of ZFDBDs so that they bind to unique sites in the genome. Off-target binding has been a concern for all synthetic DNA binding molecules. Thus, increasing the specificity and affinity of ZFDBDs will have a significant impact on their use in analytical or therapeutic settings. J. Cell. Biochem. 116: 2435–2444, 2015. © 2015 Wiley Periodicals, Inc.

KEY WORDS: synthetic DNA binding domain; zinc finger; genome editing; modulation of gene expression

For most DNA-sequence specific binding proteins α-helices are placed into the major groove and specific amino acids engage in base specific contacts [Mechetin and Zharkov, 2014; Slattery et al., 2014]. The ZF typically interacts with three base pairs (bps) of DNA and is composed of an α-helix and two adjacent β-sheets [Pabo et al., 2001; Klug, 2010]. Interactions between the α-helix and one of the β-sheets is mediated by a zinc ion, which is coordinated by two cysteine and two histidine (C2H2, Fig. 1A) or four cysteine (C4) residues. The C2H2-ZF is the most common DNA binding motif found in eukaryotic transcription factors and its mode of DNA binding is very well understood [Pabo et al., 2001; Klug, 2010]. The RNA polymerase III transcription factor TFIIIA was among the first eukaryotic transcription factors that has been purified to homogeneity and for which the corresponding coding DNA was identified and sequenced [Pelham and Brown, 1980; Ginsberg et al., 1984]. Early biochemical work in the Klug and Wu laboratories demonstrated that TFIIIA contained Zn and consisted of repeating blocks of about 3 kDa [Hanas et al., 1983; Miller et al., 1985]. TFIIIA contains 9 C2H2 ZFs that make extensive contacts with the internal control region (ICR) of the 5S rRNA gene, creating a 50 bp DNAse I footprint [Seifart et al., 1989]. However, not all of the nine ZFs of TFIIIA participate in canonical DNA interactions. Most ZF-transcription factors contain fewer ZFs and bind to a large number of regulatory DNA elements [Pabo et al., 2001; Klug, 2010]. For example, transcription factor Sp1 contains 3 C2H2 ZFs and interacts with CpG-rich DNA in regulatory DNA elements that direct or modulate transcription by RNA polymerase II (Pol II) [Suske, 1999].

The first co-crystal structure of a ZF protein with a DNA fragment was that of the mouse transcription factor Zif268 [Pavletich and Pabo, 1991]. The structure revealed that 3 ZFs follow the helical path and that residues at the N-terminus of the α-helix interact with bases in the major groove. The binding of Zif268 is antiparallel in that the N-terminal residues of the ZF make contacts with the 3′end of the DNA binding site and the C-terminal residues contact the 5′end. Work on Zif268 has shown that residues −1, 3, and 6, relative to the start site of the α-helix, make contacts with 3 specific bases in one of the DNA strands [Pavletich and Pabo, 1991; Pabo et al., 2001; Klug, 2010]. The interaction of one ZF with a triple bp sequence is a common feature of DNA binding ZF proteins. However, often residue 2 of the α-helix contacts the next base at the 3′end of the opposite

Grant sponsor: National Institutes of Health; Grant numbers: R01 DK083389, R01 DK 52356.
*Correspondence to: Jörg Bungert, Ph.D., Department of Biochemistry and Molecular Biology, College of Medicine, University of Florida, 1600 SW Archer Road, PO Box: 100245, Gainesville, FL 32610.
E-mail: jbungert@ufl.edu
Manuscript Received: 7 May 2015; Manuscript Accepted: 11 May 2015
Accepted manuscript online in Wiley Online Library (wileyonlinelibrary.com): 18 May 2015
DOI 10.1002/jcb.25226 • © 2015 Wiley Periodicals, Inc.

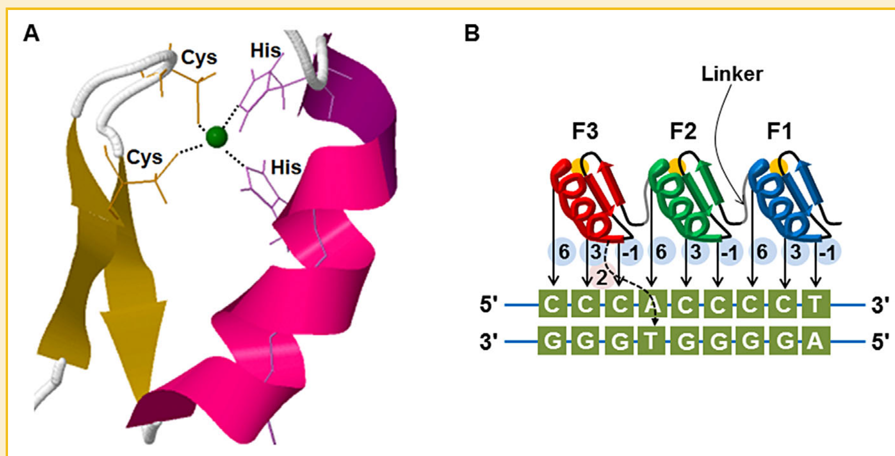

Fig. 1. Structure and DNA binding patterns of C2H2 zinc fingers. (A) Structure of the zinc finger outlining two β-sheets (yellow), an α-helix (pink), and a zinc atom (green circle) coordinated by two cysteine (Cys) and two histidine (His) residues. (B) Interaction of a 3 ZF protein with a DNA sequence of 9 bps. As outlined, residues −1, 3, and 6 make specific contacts with the nucleotides in the major groove within a triple bp, and residue 2 also contacts a nucleotide of a neighboring bp.

strand highlighting the important fact that ZFs not only recognize a triple bp sequence but also the sequence context with respect to neighboring bps (Fig. 1B) [Pabo et al., 2001; Klug, 2010].

The individual ZFs within DNA-binding transcription factors are most often connected by 5 AA linker regions with the conserved sequence TGEKP, which is found in all krüppel-like C2H2 ZF proteins [Pabo et al., 2001; Klug, 2010]. The linker regions appear to play an active role in DNA binding and data suggest that they may stabilize protein/DNA interactions at the C-terminal region of the α-helix [Wuttke et al., 1997]. Interestingly, linker regions have been shown to adopt an ordered configuration upon binding of the ZF protein to DNA [Laity et al., 2000a]. Variations of the linker length are associated with altered modes of DNA binding and can either enhance or inhibit the binding of specific ZFs in the context of a ZFDBD [Pabo et al., 2001; Klug, 2010].

ZFs not only interact with DNA but also with RNA and other proteins [Pabo et al., 2001; Klug, 2010]. This should be kept in mind when designing artificial ZF proteins. For example, the erythroid-specific transcription factor GATA-1 interacts with the co-activator cAMP-responsive element binding protein (CEBP)-binding protein (CBP) via one of the two ZFs [Blobel et al., 1998]. In addition, the GATA-1 ZF also interacts with Sp1 via ZF–ZF interactions [Imanishi et al., 2010]. The linker regions of the ZF proteins prominently participate in the ZF–ZF contacts which are mostly electrostatic [Imanishi et al., 2010].

DESIGN AND GENERATION OF ARTIFICIAL ZF-DNA BINDING PROTEINS

The various modes by which ZFs contact specific bases either within or outside the target triple bp binding site make it difficult to generate straightforward rules and codes for the design of proteins intended to bind unique DNA sequences in the genome [Mandell and Barbas, 2006; Gersbach et al., 2014]. Nevertheless, as discussed below, there are several widely used tools that can predict the optimal target DNA sequence for recognition by tailored ZFDBDs.

Early on, the Klug and Berg laboratories pioneered the design of ZF proteins with altered DNA binding specificities [Desjarlais and Berg, 1992, 1993; Choo and Klug, 1994a; Isalan et al., 1997]. Specific mutations in key residues of the Zif268 DNA binding interface changed the specificity of ZFs toward different DNA sequences. The development of selection procedures like phage display that allowed identification of specific DNA binding proteins from a large library of ZF proteins with randomized alterations of specific amino acid residues provided a breakthrough in the design of artificial DNA binding proteins [Rebar and Pabo, 1994; Choo and Klug, 1994b; Wu et al., 1995]. Furthermore, the Pabo group developed an elegant three step random mutagenesis and selection procedure based on Zif268 to generate a 3-ZF protein that interacts with a specific 9 bp sequence [Greisman and Pabo, 1997]. This selection scheme also considered neighboring effects by adjacent triple bps.

More recent selection procedures involved bacterial one-hybrid (B1H) or bacterial two hybrid (B2H) systems [Maeder et al., 2008; Persikov et al., 2015]. The most extensive screen for DNA binding ZF proteins using the B1H system has recently been published by Persikov et al. [2015]. The authors generated a large number of mutant three ZF proteins in which AAs at position −1, 1, 2, 3, 5, and 6 were altered in two of the ZFs. The computational analysis confirmed the importance of the core AA residues at positions −1, 2, 3, and 6 for the specificity and affinity of protein DNA interactions but also revealed novel aspects of neighboring effects as well as interactions within the context of individual ZFs.

The selection of specific ZFs capable of interacting with specific triple bp sequences led to the generation of ZF libraries [Bae et al., 2003]. ZFs from these libraries can be linked together to generate ZF proteins with determined binding specificities. This modular assembly approach was successful in generating highly specific ZFDBDs. However, the general applicability of the modular assembly was questioned by experiments showing that most of the assembled

ZF proteins failed to activate a reporter gene in B2H assays [Ramirez et al., 2008]. Subsequent studies demonstrated that most of the ZF proteins that failed in the B2H assay revealed correct binding selectivity in a protein binding microarray [Lam et al., 2011]. Nevertheless, both studies revealed that modular assembly is most successful for ZFs that bind to GNN nucleotides. An explanation for the GNN preference is that most artificial ZFDBDs are based on either Zif268 or Sp1 ZF proteins that bind to GNN. Furthermore in GNN-binding ZFs an arginine at position 6 interacts with the 5′guanine via two hydrogen bonds which contributes to stronger protein/DNA interactions [Pavletich and Pabo, 1991; Elrod-Erickson et al., 1998]. Triplets such as ANN or CNN are typically recognized with lower affinity. The TNN triplets are particularly limited as ZFs can only bind to TGA, TGG, or TAG.

The variability in DNA binding modes both with respect to individual ZFs and neighboring effects renders the prediction of DNA binding specificity challenging. Gupta et al. [2014] recently used the known binding specificity of 678 two-ZF modules to construct a random forest-based predictive model for ZF-specificity [Gupta et al., 2014]. Likewise, Persikov et al. [2014] used structural models and prediction algorithm to generate position weight matrices (PWMs) for ZFs at four nucleotide positions reflecting the interactions of the four core AAs (−1, 2, 3, and 6) [Persikov and Singh, 2014]. Even though there are limitations to this approach it is quite encouraging that for up to 80% of the ZF proteins binding preferences could be predicted.

To target a unique site in the human genome ZFDBDs need to contain at least 6 ZFs [Gersbach et al., 2014]. The reason for this is that a DNA sequence larger than 18 bps is predicted to be present only once in the human genome. As mentioned before, the AA sequence linking two adjacent ZFs is relatively well conserved and usually consists of TGEKP [Pabo et al., 2001; Klug, 2010]. Several studies have shown that increasing the number of ZFs beyond 3 does not necessarily increase the specificity and affinity of the protein. It was shown that if 6 ZFs are linked via canonical spacers, the protein is strained and may not adopt an ideal binding configuration in which the subsequent ZFs follow the major grove of the DNA [Pabo et al., 2001; Klug, 2010]. Pabo and coworkers have shown that extending the linker between the 3rd and 4th ZF leads to order of magnitudes increased affinities and specificities [Kim and Pabo, 1998]. Likewise, Moore et al. [2001] have shown that strings of two ZF units connected via extended linkers increased affinity and specificity by orders of magnitudes [Moore et al., 2001]. Both groups used target sites that either do or do not contain nucleotide gaps between the ZFs connected by extended linkers. The two-unit approach led to the generation of ZFDBDs with very little tolerance of variations in the targeted site [Moore et al., 2001]. However, it should be mentioned that other experiments failed to reproduce the beneficial effect of increasing the linker length on the binding affinity [Neuteboom et al., 2006]. The differences between these two studies could be due to the fact that different backbones for the ZF-design (Zif268 vs. Sp1) were used. The Pabo group has solved the crystal structure of 2 copies of three ZFDBDs bound to adjacent sites mimicking a six ZFDBD with sufficient flexibility between the three finger modules. The authors demonstrated that the protein is still strained when bound to DNA, which suggests that extended linkers may provide advantages to ZFDBDs other than conferring flexibility [Peisach and Pabo, 2003].

Other studies investigated the effect of linker length and sequence on the binding of Sp1 ZFs [Imanishi et al., 2000; Imanishi and Sugiura, 2002]. The authors introduced polyglycine, -glutamine, or -arginine into the linker and demonstrated that the altered linker in some cases caused DNA bending and that the different AA sequences had different effects on the affinity of protein/DNA interactions. Yan et al. [2007] introduced α-helix creating peptide sequences into the linker between two 3ZF-binding domains and showed that the resulting ZFDBD binds to sequences with up to 10 bp gaps between the two 3ZF units [Yan et al., 2007]. Whether increased specificity with extended linker regions could be a guiding principle or whether this depends on the sequence or conformation of the target DNA sequence remains to be established. For example, one study has shown that an extended linker between ZFs can lead to impaired ZF binding to the cognate site [Laity et al., 2000b]. In this case it was shown that alternative splicing leading to an additional three residues (KTS) within the canonical linker (TGEKP) between ZF three and ZF four of tumor suppressor protein WT1 increased its flexibility and abrogated DNA binding of the fourth finger. Therefore, it is clear that the distance between ZFs is an important determinant for optimal ZF-DNA interactions. The effect of the linker length on binding specificity is likely mediated by the sequence context of the target site and/or the conformation of the DNA.

Several platforms are available that assist in the design and generation of ZFDBDs. OPEN (Oligomerized Pool Engineering) is an open-source, publicly available resource that is based on a library of ZF-pools constructed by the Zinc Finger Consortium (http://www.zinfingers.org) [Maeder et al., 2008]. The pool consists of randomized libraries of ZFs in the context of a 3ZF protein. A different method was developed by the Barbas laboratory [Mandell and Barbas, 2006; Gersbach et al., 2014]. This method uses algorithms that are based on the knowledge that exists on the mode of ZF-DNA interactions to determine the AA sequence of ZFs that would bind specific triple bps in the DNA (http://www.zincfingertools.org). The rules not only take into account the interaction of ZFs with triple bps but also the effect of neighboring DNA sequences in the context of a DNA binding domain containing 3 ZFs. One aspect not taken into account when designing synthetic ZFDBDs is the DNA conformation. It has been shown that many protein/DNA interactions are governed by both sequence and shape readout [Rohs et al., 2010]. Nevertheless, the available platforms are extremely helpful for the initial design and generation of ZFDBDs. However, the specificity and affinity of the artificial ZFDBDs need to be examined using in vitro and in vivo approaches as outlined below.

AFFINITY AND SPECIFICITY

Assessing the specificity and affinity of artificial ZFDBDs in vitro and in vivo is important for their use as analytical or therapeutic tools. Electrophoretic mobility shift assays (EMSAs) and isothermal titration calorimetry are frequently used to examine specificity and dissociation constant (Kd) of DNA binding proteins in vitro [Churion et al., 2014; Wang et al., 2014b]. Kds for ZF proteins are often in the

nanomolar to femtomolar range [Pabo et al., 2001; Klug, 2010]. As mentioned before, the Pabo and Klug laboratories developed ZFDBDs with extended linker regions that bind in the femtomolar to picomolar range [Kim and Pabo, 1998; Moore et al., 2001]. The in vivo occupancy of ZFs at specific and at off-target sites can be determined by chromatin-immunoprecipitation (ChIP), allowing determination of genome wide occupancy when followed by high through-put sequencing (ChIP-seq) [Furey, 2012]. ChIP-Exo-Seq, in which the ChIP selected DNA is subjected to exonuclease digestion before high-throughput sequencing, provides binding data with higher resolution [Rhee and Pugh, 2011].

Occupancy of ZFDBDs in vivo is determined by many parameters including specificity/affinity, concentration, chromatin accessibility, and potential interactions with other proteins. Several recent genome wide studies on the in vivo occupancy of artificial ZFDBDs revealed wide-spread binding to off-target sites. Grimmer et al. [2014] recently generated 6 ZF-containing, HA-tagged ZFDBDs targeting the Sox2 promoter in MCF7 breast cancer cells. The authors subjected the cells to ChIP-seq. using HA-specific antibodies and to RNA seq. The data showed that the ZFDBDs interacted with thousands of sites in the genome [Grimmer et al., 2014]. However, despite extensive off-target site binding, only a few genes revealed expression changes in cells expressing the ZFDBD. Furthermore, the inclusion of a KRAB repression domain increased the number of ZFDBD binding sites in the genome. KRAB domains are often found in C2H2 ZF proteins and repress transcription by modifying histones [Lupo et al., 2013]. Burdach et al. [2014] analyzed the genome wide occupancy of the naturally occurring 3ZF-containing Krüppel-like factor 3 (KLF3) and found that mutations in a domain that interacts with the co-repressor C-terminal binding protein (CtBP), altered the binding specificity of the ZFDBD [Burdach et al., 2014]. These data clearly show that in vivo occupancy is guided not just by the DNA binding domain but also by protein:protein interactions and chromatin accessibility. Theoretically, an increase in the number of ZFs should increase the binding affinity for a target site and should also increase the specificity of interactions if the concentration of the ZFDBD is restricted to low levels. In this respect, Grimmer et al. [2014] found that not all of the 6 ZFs in the ZFDBD are involved in sequence-specific interactions. In fact, the genome-wide binding site analysis suggests that only the middle 4 ZFs were engaged in protein/DNA interactions [Grimmer et al., 2014]. This suggests, as discussed by the authors, that increasing the number of ZFs may indeed increase the specificity of the ZFDBDs. According to the findings in the Pabo and Klug laboratories, increasing the length of the linker region every 2 or 3 ZFs to reduce potential strain may allow the design of highly specific ZF proteins [Kim and Pabo, 1998; Moore et al., 2001]. The use of dimeric ZFDBDs that only function when they form dimers at the target site could increase the specificity of DNA interactions. Dimeric ZFs are most often used in genome editing procedures but they have also been applied to reconstituting enzymatic activity of epigenetic modifiers at specific sites in the genome [Nomura and Barbas, 2007; Gersbach et al., 2014].

Another challenge in targeting ZFDBDs to specific sites in the genome is the control of ZFDBD expression levels in the cell. Increasing the expression levels of the ZFDBDs will generally lead to increased unspecific binding events. It is important to keep the concentration of the proteins sufficiently low so that they can efficiently interact with a specific site while minimizing off-target binding. Thus, the combination of generating ZFDBDs with high specificity/affinity and reducing expression levels is expected to significantly decrease off-site binding events. Expression levels can in principle be controlled by placing the ZFDBD encoding DNA under control of strong or weak eukaryotic promoters or by using inducible systems to transiently express the ZFDBDs at will [Pabo et al., 2001; Klug, 2010; Gersbach et al., 2014]. Beerli et al. [2000] generated ZF transcription factors that were fused to steroid hormone receptor binding sites and were thus subject to regulation by hormone ligands [Beerli et al., 2000]. A negative feed-back loop could also be incorporated into the ZFDBD expression plasmid to limit its expression (Fig. 2). For example, a partial binding site for the ZFDBD could be inserted downstream of the transcription start site.

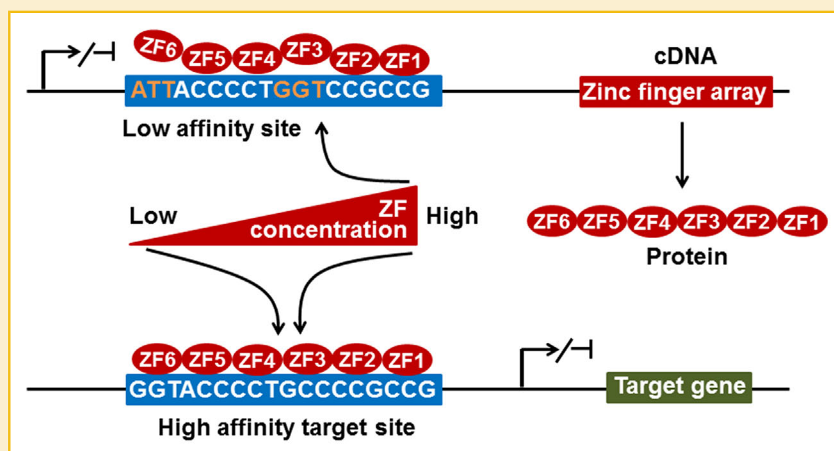

Fig. 2. Proposed mechanism of controlling ZFDBD expression. Shown below is the perfect target sequence for a designed ZFDBD. An imperfect sequence is placed downstream of the promoter controlling expression of the ZFDBD (top). At low concentrations, the ZFDBD will bind the perfect target site but not the imperfect site, allowing increased ZFDBD expression. At high concentrations the ZFDBD will bind to its own promoter and reduce expression.

At low expression levels the ZFDBD will bind to the perfect target site in the genome but not to the imperfect site in the downstream promoter of its own gene. At high expression levels the ZFDBD will also bind its own gene and inhibit transcription. Previous studies have shown that targeting a ZFDBD downstream of the transcription start site impairs transcription elongation [Choo et al., 1994; Barrow et al., 2014]. This feedback approach would monitor expression levels that lead to unspecific binding.

If ZFDBDs are targeted to heterochromatic regions, accessibility could become an issue. To allow binding of ZFDBDs to inaccessible sites in the genome, cells could be treated with histone deacetylase (HDAC) or DNA methyltransferase (DMTs) inhibitors [Kubicek et al., 2012]. However, this may not be necessary as recent studies demonstrated that ZFDBDs can be targeted to repressed chromatin and are able to bind methylated DNA [Sasai et al., 2010; Buck-Koehntop et al., 2012]. For example, Deng et al. [2014] have shown that a ZFDBD interacts with a normally inaccessible target site in the silenced fetal γ-globin gene [Deng et al., 2014]. Thus, ZFDBDs may also be employed as pioneer transcription factors to initiate the access over regulatory DNA elements.

As mentioned previously, specificity of ZFDBDs in vivo is primarily assessed by ChIP based assays. ChIP assays normally include a prolonged incubation of cells in 1–2% formaldehyde [Furey, 2012]. It is possible that under these conditions many transient and unproductive protein/DNA interactions are cross-linked. Most DNA binding proteins slide along accessible DNA or randomly sample sites in the genome before they engage in DNA sequence specific interactions [von Hippel and Berg, 1989]. If chromosomal sites are accessible and if there are no competing proteins, ZFDBDs, particularly if they are expressed at high levels, may frequently occupy these sites and will be efficiently crosslinked. This interpretation is in line with the observation by Grimmer et al. [2014] that most off-site target events remain without functional consequence on nearby genes. The Auble laboratory recently developed a modified ChIP protocol with reduced crosslinking time [Poorey et al., 2013]. This crosslinking kinetics (CLK) method allows determining the on- and off-rate of the interactions between proteins and specific sites in chromatin in vivo and thus represents an ideal tool to determine in vivo occupancy of ZFDBDs.

APPLICATION OF ZFDBDS

There are many applications for ZFDBDs as either analytical or therapeutic tools. ZFDBDs can be fused to transcriptional activation or repression domains to alter expression of specific genes (Fig. 3A and B), to epigenetic modifiers to alter chromatin accessibility at gene loci (Fig. 3C), or to protein/protein interaction domains to force interactions between long-distance gene regulatory elements and gene promoters (Fig. 3E). Furthermore, ZFDBDs are also used for gene editing by fusing them with nucleases that initiate DNA repair or DNA recombination processes (Fig. 4).

SYNTHETIC ZF TRANSCRIPTION FACTORS

Transcription factors are modular and usually contain multiple separable domains that confer DNA binding, transcription activation, or repression, protein/protein interactions, or regulatory domains that are post-transcriptionally modified (e.g., phosphorylation or

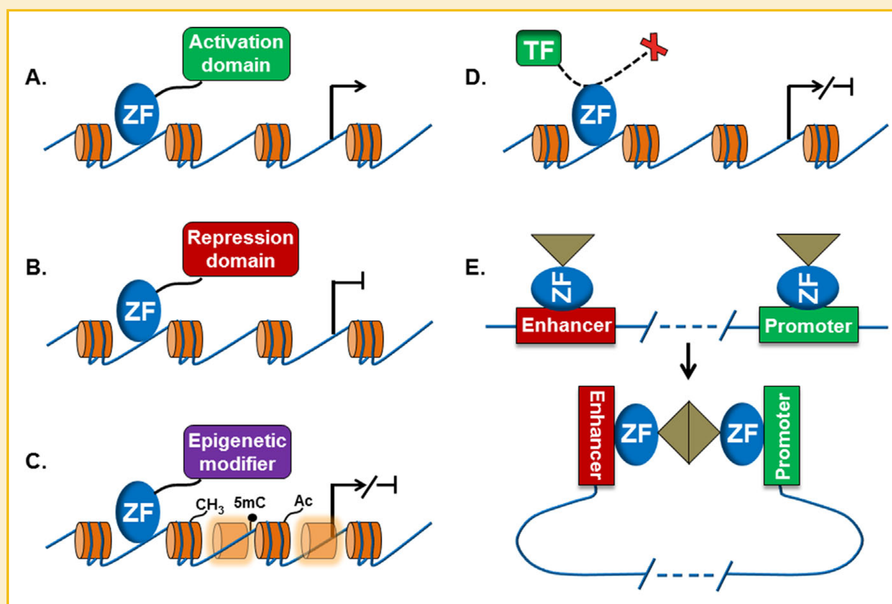

Fig. 3. Applications for ZFDBDs. (A) ZFDBDs can be expressed together with an activation domain and targeted to a promoter to activate expression of a specific gene in the genome. (B) ZFDBDs can be expressed together with a repression domain and targeted to a promoter to repress expression of a specific gene. (C) ZFDBDs can be expressed together with an enzymatic activity that introduces epigenetic changes (e.g., DNA methylation, histone modifications) to a genomic region targeted by the ZFDBD. (D) ZFDBDs can be expressed without effector domains and targeted to specific transcription factor binding sites to neutralize the function of transcription factors at specific gene loci. (E) ZFDBDs can be expressed together with protein dimerization domains (triangles) to mediate interactions between distant regulatory sites in the genome (e.g., enhancer and promoter).

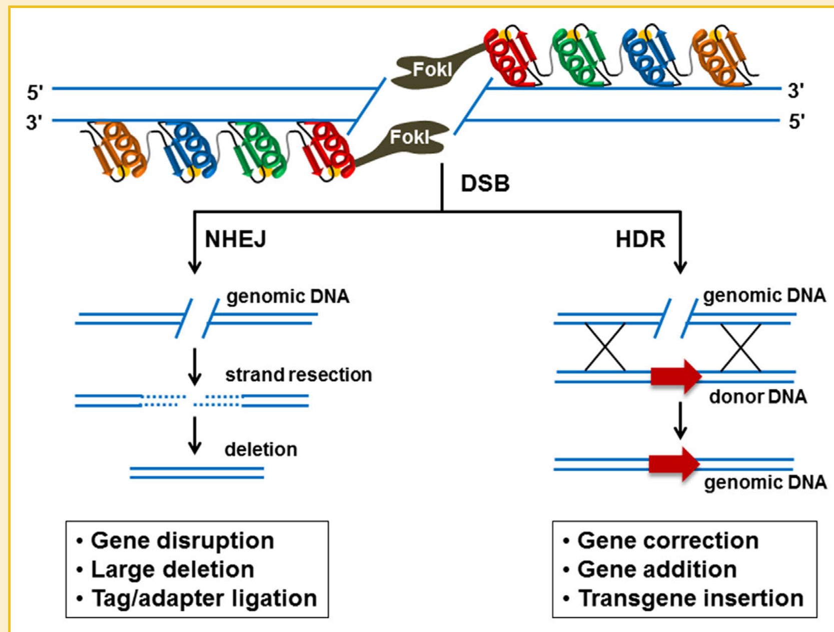

Fig. 4. Genome editing by ZF-Nucleases. ZFDBDs can be expressed together with DNA nucleases to induce DNA repair processes in the cell. Shown on top are two 4-ZF proteins fused to two domains of the DNA endonuclease FokI. The nuclease activity of FokI is reconstituted only when both proteins are targeted to neighboring recognition sequences. The double strand break can be repaired either by nonhomologous end-joining (NHEJ, on the left) or by homologous recombination. NHEJ is the preferred way of repair in the absence of homologous sequences and often leads to deletions of sequences flanking the break. Homologous recombination is used if a homologous target sequence is provided and is used to correct mutations or to introduce new genes or gene-tags.

methylation) [Ptashne and Gann, 1990]. Thus ZFDBDs can be fused to different protein domains to target activation or repression domains to specific gene promoters. Commonly used activation domains are derived from the strong Herpes Simplex Virus (HSV) transcription activator VP16 [Gräslund et al., 2005]. VP64 consists of a tetrameric repeat of the minimal activation domain of VP16 [Gräslund et al., 2005]. To repress transcription ZFDBDs are frequently fused to the KRAB (Krüppel-associated box) domain [Lupo et al., 2013]. In one study authors fused a ZFDBD to VP64 or to the KRAB repression domain and showed enhanced or diminished erbB-gene transcription, respectively [Beerli et al., 1998]. Likewise, two groups reported on the successful targeting of a ZFDBD-VP64 activator to the repressed γ-globin gene and demonstrated activation of the normally silenced fetal globin gene in adult erythroid cells [Wilber et al., 2010; Costa et al., 2012]. This finding is significant because it (a) highlights the ability of ZFDBDs to target inaccessible chromatin and (b) demonstrates that ZF transcription factors are promising tools for developing new therapies for sickle cell disease and thalassemias. A number of studies used ZFDBDs to target activation or repression domains to viral genes or to tumor-suppressor as well as oncogenes [Pabo et al., 2001; Klug, 2010; Gersbach et al., 2014]. For example, Wang et al. [2014] showed that ZFDBD-VP64 transactivators targeted to the 5′Long Terminal Repeat (LTR) reactivate latent HIV-1 [Wang et al., 2014a]. Furthermore, ZF transcription factors targeted to the VEGF-A gene were shown to alter expression in cell culture and in animal models [Liu et al., 2001].

ZFDBDs have been used to target epigenetic modifiers to specific sites in the genome. The KRAB repression domain interacts with chromatin modifying enzymes and has been shown to cause increase in H3K9me3, a mark associated with repressed chromatin [Lupo et al., 2013]. Rivenbark et al. [2012] targeted the DNA methyltransferase (DNMT) 3A via a ZFDBD to the Sox2 and Maspin promoters. The targeted DNA methylation led to inhibition of expression of these genes [Rivenbark et al., 2012]. In another study ZFDBDs were targeted to the murine *Fosb* gene [Heller et al., 2014]. The ZFDBDs were fused to either the p65 activation domain derived from transcription factor NFκB, which increased histone acetylation, or to the minimal catalytic domain of G9A, a histone H3K9 methyltransferase. It was shown that the targeted histone acetylation or methylation at the Fosb promoter controlled the drug response in regions of the brain harboring the reward system.

CHROMATIN STRUCTURAL ALTERATIONS USING ZFDBDS

In an elegant series of experiments the Blobel laboratory used ZFDBDs to alter chromatin interaction patterns in the β-globin gene locus. In the first study the investigators targeted a ZFDBD fused to the protein/protein interaction domain of Ldb1 to the adult β-globin gene promoter in erythroid cells that express no GATA1 [Deng et al., 2012]. Previous studies have shown that transcription factor GATA1 is critical for long-range interactions between the adult β-globin gene promoter and the locus control region (LCR), a powerful regulatory DNA element located 50 Kb upstream of the β-globin gene [Dean, 2011]. Ldb1 is a component of a large protein complex

that contains the DNA binding proteins GATA1 and Tal 1, which bind to composite GATA-E-box motifs [Dean, 2011]. The self-dimerization of Ldb1 is thought to mediate the long-distance interactions in the globin gene locus. The authors demonstrated that expression of the ZFDBD/Ldb1 fusion protein in GATA1deficient erythroid cells restored interactions between the LCR and the adult β-globin gene promoter and activated its expression [Deng et al., 2012]. This interaction was mediated by Ldb1 targeted to the β-globin gene via the ZFDBD and by additional Ldb1 recruited to the LCR by endogenous proteins. In a subsequent study the authors demonstrated that a ZFDBD/Ldb1 fusion protein targeted to the γ-globin gene promoter, which is normally silenced in adult erythroid cells, established interactions between the LCR and the γ-globin gene and reactivated fetal globin gene expression [Deng et al., 2014]. These studies have broad implications because they demonstrate that ZFDBDs or other DNA binding domains can be used to alter chromatin architecture and change expression patterns in complex gene loci.

DISSECTING THE FUNCTION OF REGULATORY ELEMENTS AND CAUSATIVE SNPS

ZFDBDs have also been used to block the binding of endogenous transcription factors at specific sites (Fig. 3D) [Choo et al., 1994; Barrow et al., 2012, 2014]. This has first been demonstrated by Choo et al. [1994] who showed that a 3 ZF containing ZFDBD was able to repress transcription of a reporter gene when targeted to a sequence downstream of the transcription start site [Choo et al., 1994]. Similarly, a 6-ZF containing ZFDBD targeting a KLF1 binding site in the murine adult β-globin gene promoter was shown to interfere with the binding of KLF1 and to repress adult β-globin gene expression [Barrow et al., 2012]. Since ZFDBDs have been shown to interact with repressed chromatin domains it may be feasible to target these proteins to repressor binding sites and to reactivate expression of silenced genes, for example, the γ-globin gene in adult erythroid cells.

In the competition studies ZFDBDs were used without fusion partners. The use of ZFDBDs without additional protein domains may decrease binding to off-target sites due to minimizing protein: protein interactions. Furthermore, these proteins may be directly delivered to target cells due to the cell penetrating activity of ZFDBDs [Gaj et al., 2014]. The small size of ZFDBDs that lack effector domains may render this process particularly efficient. It may be important to note that the cell-penetrating activity of ZFDBDs may allow transient exposure of cells which may reduce off-site occupancy.

Functionally linking noncoding genetic variants with altered gene expression patterns is a challenging mission. Roughly 4,000 genes have been reported to be associated with human diseases and most of the significant SNP variants are thought to be located in gene proximal or distal DNA regulatory elements [Kellis et al., 2014]. In the post-GWAS era, one important issue is to confirm whether a statistically associated SNP is located in a biologically functional element and to analyze how it contributes to disease progression. ZFDBDs can be a useful tool to dissect the functionality of a causative SNP. Furthermore, theoretically ZFDBDs could mimic differential allelic expression by causing the dissociation of transcription factors with specific SNP associated DNA regulatory elements.

GENOME EDITING BY ZF-NUCLEASES

ZFDBDs have been used to target endonucleases (e.g., Fok1) to specific genomic loci to induce a double strand break that will be repaired by the DNA repair machinery using non-homologous end joining (NHEJ) or homology directed repair (HDR, Fig. 4) [Pabo et al., 2001; Klug, 2010; Gersbach et al., 2014]. If no DNA template with homology to the region targeted by the ZFDBD and cut by the nuclease is provided, the cell will repair the DNA break preferentially using NHEJ which is often accompanied by deletions. If the break is targeted to the coding region of the gene NHEJ can alter the reading frame which often results in the occurrence of a premature stop codon leading to either nonsense mediated decay or production of a non-functional protein. NHEJ can also be used to delete transcription factor binding sites. If a DNA template with homology to the DNA flanking the cut site is provided, the cell will preferentially repair the break using HDR (Fig. 4). The targeted homologous recombination allows replacement of a DNA segment with another segment, for example, replacement of a mutant gene or exon with a wild-type copy. In addition, HDR can also be used to add a DNA sequence to an endogenous gene that encodes for a specific protein tag. Orlando et al. [2010] describe a new technique by which short homology sequences will direct ligation into endogenous loci thus allow tagging genes by NHEJ [Orlando et al., 2010].

Reducing off-target binding is particularly important for genome editing. Investigators developed a procedure that requires the targeting of two ZFDBDs to reconstitute nuclease activity at a specific site [Pabo et al., 2001; Klug, 2010; Gersbach et al., 2014]. Each ZFDBD contains 3–4 ZFs and is fused with only a portion of the nuclease. Nuclease activity is only reconstituted when the two ZFDBDs bind to adjacent sites in the genome (Fig. 4). Thus, off-site binding of single ZFDBDs will not result in a DNA break, although it may still interfere with other processes in the cell. Sequence specific nuclease activity can be monitored using in vitro cleavage site selection or in vivo integration assays [Gabriel et al., 2011]. The combination of in vitro cleavage site selection and in silico abstraction provides a promising tool for identifying off-site targets of ZF nucleases [Sander et al., 2013].

Zinc finger nucleases (ZFNs) have been successfully used to correct genetic diseases in mouse and human cell lines as well as in mouse disease models. The University of Pennsylvania together with Sangamo recently published the results of an ongoing clinical trial that produced promising results [Tebas et al., 2014]. In this study a ZFN was targeted to the HIV co-receptor CCR5 in T-cells. Previous studies have shown that HIV exposed individuals with mutations in the CCR5 gene show protection from HIV infection [Liu et al., 1996]. Furthermore, allogenic transplants of cells from individuals carrying the mutations into HIV infected patients demonstrated protection and long term benefits [June and Levine, 2012]. Preceding the current phase 1 trial with CCR5 targeted ZFN the investigators

performed toxicity tests in mice and demonstrated that the ZFN did not cause negative side effects [Tebas et al., 2014]. These are promising findings but future studies have to address the global occupancy of the ZFNs and a careful analysis of off-site double-strand cuts in the genome.

Recently, investigators were successful in correcting the sickle cell mutation in human hematopoietic stem/progenitor cells using ZFNs [Hoban et al., 2015]. Importantly, the investigators did not detect any off-site nuclease activity in these cells. However, as mentioned before, ZFDBDs bind to many other sites in the genome and may interfere with other processes by competing with other binding activities. This may not be so much of an issue in gene editing as most approaches attempt to express the ZFNs only transiently.

ZFDBD DELIVERY

In most cases ZFDBDs, ZF-transcription factors, or ZFNs are delivered to cells via virus mediated transfection [Pabo et al., 2001; Klug, 2010; Gersbach et al., 2014]. If the virus is allowed to integrate its DNA into the host genome it may cause alterations of gene expression at or near the site of integration. For genome editing purposes it is desirable to introduce the ZFNs and putative therapeutic DNA used for recombination only transiently, so that cells derived from the procedure no longer express the components. Several inducible systems are available to transiently express ZFNs [Pabo et al., 2001; Klug, 2010; Gersbach et al., 2014]. If ZFDBDs are used as competitors or artificial transcription factors, they need to be expressed for a long period of time, unless epigenetic changes introduced by artificial ZF transcription factors are stably maintained during subsequent cell divisions.

Several studies have shown that ZFDBDs and fusion proteins thereof can be delivered to cells by protein transduction. Tachikawa et al. [2004] fused ZF transcription factors that were targeted to the VEGF-A gene to cell penetrating peptides and demonstrated direct protein delivery [Tachikawa et al., 2004]. More recently it was demonstrated that ZFDBDs have intrinsic cell penetrating capabilities [Gaj et al., 2014; Gaj and Liu, 2015]. This is a feature of the positively charged ZFs themselves. Gaj et al. [2014] fused a non-DNA binding 2-ZF containing protein with emerald GFP (EMGFP), which does not have intrinsic cell penetrating activity. The incubation of the protein with HeLa cells led to a very high fraction of cells that revealed fluorescence. More recently, it was demonstrated that inclusion of two instead of one nuclear localization domains further increased the cell penetrating efficiency [Liu et al., 2015]. ZFDBDs without large effector domains may be even more efficient in penetrating the cell membrane compared to larger proteins containing activation, repression, or dimerization domains. However, the benefit of directly delivering the ZFDBDs, ZF-transcription factors, or ZFNs to cells could be hampered by instability and short half-life of these proteins. It was shown that two different ZFNs exhibited a half-life of about 2hrs in 293 T-cells and that they were subjected to ubiquitination and proteasome mediated degradation [Ramakrishna et al., 2013]. It is not clear if instability was due to the nuclease domains. Other studies have shown that the half-life of ZF proteins depends on the specific amino acid sequence and/or structure of the protein [Kang and Kim, 2000].

CONCLUSIONS AND FUTURE DIRECTIONS

The last two decades brought enormous progress in the development of synthetic DNA binding activities that can alter the genome or modulate chromatin structure or gene expression. The CRISPR/Cas9, TALE, and ZF systems have each advantages and disadvantages. For all three systems off-target binding is a concern and future work will have to focus on increasing DNA binding specificity and affinity. ZFDBDs have the advantage of being relatively small and of being able to penetrate cell membranes. Thus, in the future ZFDBDs may be used without a DNA intermediate, which not only reduces concerns about DNA integration mediated mutagenesis, but also may allow precise dosing of these proteins for experimental or therapeutic purposes.

ACKNOWLEDGMENTS

We thank our colleagues in the Bungert laboratory for stimulating discussions. The research in the author's laboratory was funded through grants from the National Institutes of Health (NIH, R01 DK083389 and R01 DK 52356 to J.B)

REFERENCES

Bae KH, Kwon YD, Shin HC, Hwang MS, Ryu EH, Park KS, Yang HY, Lee DK, Lee Y, Park J, Kwon HS, Kim HW, Yeh BI, Lee HW, Sohn SH, Yoon J, Seol W, Kim JS. 2003. Human zinc fingers as building blocks in the construction of artificial transcription factors. Nat Biotechnol 21:275–280.

Barrow JJ, Li Y, Hossain M, Huang S, Bungert J. 2014. Dissecting the function of the adult β-globin downstream promoter region using an artificial zinc finger DNA-binding domain. Nucleic Acids Res 42:4363–4374.

Barrow JJ, Masannat J, Bungert J. 2012. Neutralizing the function of a β-globin-associated cis-regulatory DNA element using an artificial zinc finger DNA-binding domain. Proc Natl Acad Sci USA 109:17948–17953.

Beerli RR, Schopfer U, Dreier B, Barbas CF. 2000. Chemically regulated zinc finger transcription factors. J Biol Chem 275:32617–32627.

Beerli RR, Segal DJ, Dreier B, Barbas CF. 1998. Toward controlling gene expression at will: Specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. Proc Natl Acad Sci USA 95:14628–14633.

Blobel GA, Nakajima T, Eckner R, Montminy M, Orkin SH. 1998. CREB-binding protein cooperates with transcription factor GATA-1 and is required for erythroid differentiation. Proc Natl Acad Sci USA 95:2061–2066.

Buck-Koehntop BA, Stanfield RL, Ekiert DC, Martinez-Yamout MA, Dyson HJ, Wilson IA, Wright PE. 2012. Molecular basis for recognition of methylated and specific DNA sequences by the zinc finger protein Kaiso. Proc Natl Acad Sci USA 109:15229–15234.

Burdach J, Funnell AP, Mak KS, Artuz CM, Wienert B, Lim WF, Tan LY, Pearson RC, Crossley M. 2014. Regions outside the DNA-binding domain are critical for proper in vivo specificity of an archetypal zinc finger transcription factor. Nucleic Acids Res 42:276–289.

Choo Y, Klug A. 1994a. Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions. Proc Natl Acad Sci USA 91:11168–11172.

Choo Y, Klug A. 1994b. Toward a code for the interactions of zinc fingers with DNA: Selection of randomized fingers displayed on phage. Proc Natl Acad Sci USA 91:11163–11167.

Choo Y, Sánchez-García I, Klug A. 1994. In vivo repression by a site-specific DNA-binding protein designed against an oncogenic sequence. Nature 372:642–645.

Churion K, Liu Y, Hsiao HC, Matthews KS, Bondos SE. 2014. Measuring Hox-DNA binding by electrophoretic mobility shift analysis. Methods Mol Biol 1196:211–230.

Costa FC, Fedosyuk H, Neades R, de Los Rios JB, Barbas CF, Peterson KR. 2012. Induction of fetal hemoglobin in vivo mediated by a synthetic γ-Globin zinc finger activator. Anemia 2012:507894.

Dean A 2011. In the loop: Long range chromatin interactions and gene regulation. Brief Funct Genomics 10:3–10.

Deng W, Lee J, Wang H, Miller J, Reik A, Gregory PD, Dean A, Blobel GA. 2012. Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. Cell 149:1233–1244.

Deng W, Rupon JW, Krivega I, Breda L, Motta I, Jahn KS, Reik A, Gregory PD, Rivella S, Dean A, Blobel GA. 2014. Reactivation of developmentally silenced globin genes by forced chromatin looping. Cell 158:849–860.

Desjarlais JR, Berg JM. 1992. Toward rules relating zinc finger protein sequences and DNA binding site preferences. Proc Natl Acad Sci USA 89:7345–7349.

Desjarlais JR, Berg JM. 1993. Use of a zinc-finger consensus sequence framework and specificity rules to design specific DNA binding proteins. Proc Natl Acad Sci USA 90:2256–2260.

Elrod-Erickson M, Benson TE, Pabo CO. 1998. High-resolution structures of variant Zif268-DNA complexes: Implications for understanding zinc finger-DNA recognition. Structure 6:451–464.

Furey TS 2012. ChIP-seq and beyond: New and improved methodologies to detect and characterize protein-DNA interactions. Nat Rev Genet 13:840–852.

Gabriel R, Lombardo A, Arens A, Miller JC, Genovese P, Kaeppel C, Nowrouzi A, Bartholomae CC, Wang J, Friedman G, Holmes MC, Gregory PD, Glimm H, Schmidt M, Naldini L, von Kalle C. 2011. An unbiased genome-wide analysis of zinc-finger nuclease specificity. Nat Biotechnol 29:816–823.

Gaj T, Liu J. 2015. Direct protein delivery to mammalian cells using cell-permeable Cys 2-His2 zinc-finger domains. J Vis Exp.

Gaj T, Liu J, Anderson KE, Sirk SJ, Barbas CF. 2014. Protein delivery using Cys2-His2 zinc-finger domains. ACS Chem Biol 9:1662–1667.

Gersbach CA, Gaj T, Barbas CF. 2014. Synthetic zinc finger proteins: The advent of targeted gene regulation and genome modification technologies. Acc Chem Res 47:2309–2318.

Ginsberg AM, King BO, Roeder RG. 1984. Xenopus 5S gene transcription factor, TFIIIA: Characterization of a cDNA clone and measurement of RNA levels throughout development. Cell 39:479–489.

Greisman HA, Pabo CO. 1997. A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites. Science 275:657–661.

Grimmer MR, Stolzenburg S, Ford E, Lister R, Blancafort P, Farnham PJ. 2014. Analysis of an artificial zinc finger epigenetic modulator: Widespread binding but limited regulation. Nucleic Acids Res 42:10856–10868.

Gräslund T, Li X, Magnenat L, Popkov M, Barbas CF. 2005. Exploring strategies for the design of artificial transcription factors: Targeting sites proximal to known regulatory regions for the induction of gamma-globin expression and the treatment of sickle cell disease. J Biol Chem 280: 3707–3714.

Gupta A, Christensen RG, Bell HA, Goodwin M, Patel RY, Pandey M, Enuameh MS, Rayla AL, Zhu C, Thibodeau-Beganny S, Brodsky MH, Joung JK, Wolfe SA, Stormo GD. 2014. An improved predictive recognition model for Cys(2)-His(2) zinc finger proteins. Nucleic Acids Res 42:4800–4812.

Hanas JS, Hazuda DJ, Bogenhagen DF, Wu FY, Wu CW. 1983. Xenopus transcription factor A requires zinc for binding to the 5 S RNA gene. J Biol Chem 258:14120–14125.

Heller EA, Cates HM, Peña CJ, Sun H, Shao N, Feng J, Golden SA, Herman JP, Walsh JJ, Mazei-Robison M, Ferguson D, Knight S, Gerber MA, Nievera C, Han MH, Russo SJ, Tamminga CS, Neve RL, Shen L, Zhang HS, Zhang F, Nestler EJ. 2014. Locus-specific epigenetic remodeling controls addiction- and depression-related behaviors. Nat Neurosci 17:1720–1727.

Hoban MD, Cost GJ, Mendel MC, Romero Z, Kaufman ML, Joglekar AV, Ho M, Lumaquin D, Gray D, Lill GR, Cooper AR, Urbinati F, Senadheera S, Zhu A, Liu PQ, Paschon DE, Zhang L, Rebar EJ, Wilber A, Wang X, Gregory PD, Holmes MC, Reik A, Hollis RP, Kohn DB. 2015. Correction of the sickle cell disease mutation in human hematopoietic stem/progenitor cells. Blood 125: 2597–2604.

Imanishi M, Hori Y, Nagaoka M, Sugiura Y. 2000. DNA-bending finger: Artificial design of 6-zinc finger peptides with polyglycine linker and induction of DNA bending. Biochemistry 39:4383–4390.

Imanishi M, Imamura C, Higashi C, Yan W, Negi S, Futaki S, Sugiura Y. 2010. Zinc finger-zinc finger interaction between the transcription factors, GATA-1 and Sp1. Biochem Biophys Res Commun 400:625–630.

Imanishi M, Sugiura Y. 2002. Artificial DNA-bending six-zinc finger peptides with different charged linkers: Distinct kinetic properties of DNA bindings. Biochemistry 41:1328–1334.

Isalan M, Choo Y, Klug A. 1997. Synergy between adjacent zinc fingers in sequence-specific DNA recognition. Proc Natl Acad Sci USA 94:5617–5621.

June C, Levine B. 2012. Blocking HIV′s attack. Sci Am 306:54–59.

Kang JS, Kim JS. 2000. Zinc finger proteins as designer transcription factors. J Biol Chem 275:8742–8748.

Kellis M, Wold B, Snyder MP, Bernstein BE, Kundaje A, Marinov GK, Ward LD, Birney E, Crawford GE, Dekker J, Dunham I, Elnitski LL, Farnham PJ, Feingold EA, Gerstein M, Giddings MC, Gilbert DM, Gingeras TR, Green ED, Guigo R, Hubbard T, Kent J, Lieb JD, Myers RM, Pazin MJ, Ren B, Stamatoyannopoulos JA, Weng Z, White KP, Hardison RC. 2014. Defining functional DNA elements in the human genome. Proc Natl Acad Sci USA 111:6131–6138.

Kim JS, Pabo CO. 1998. Getting a handhold on DNA: Design of poly-zinc finger proteins with femtomolar dissociation constants. Proc Natl Acad Sci USA 95:2812–2817.

Klug A 2010. The discovery of zinc fingers and their applications in gene regulation and genome manipulation. Annu Rev Biochem 79:213–231.

Kubicek S, Gilbert JC, Fomina-Yadlin D, Gitlin AD, Yuan Y, Wagner FF, Holson EB, Luo T, Lewis TA, Taylor B, Gupta S, Shamji AF, Wagner BK, Clemons PA, Schreiber SL. 2012. Chromatin-targeting small molecules cause class-specific transcriptional changes in pancreatic endocrine cells. Proc Natl Acad Sci USA 109:5364–5369.

Laity JH, Dyson HJ, Wright PE. 2000a. 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 295:719–727.

Laity JH, Dyson HJ, Wright PE. 2000b. Molecular basis for modulation of biological function by alternate splicing of the Wilms′ tumor suppressor protein. Proc Natl Acad Sci USA 97:11932–11935.

Lam KN, van Bakel H, Cote AG, van der Ven A, Hughes TR. 2011. Sequence specificity is obtained from the majority of modular C2H2 zinc-finger arrays. Nucleic Acids Res 39:4680–4690.

Liu J, Gaj T, Wallen MC, Barbas CF. 2015. Improved cell-penetrating zinc-finger nuclease proteins for precision genome engineering. Mol Ther Nucleic Acids 4:e232.

Liu PQ, Rebar EJ, Zhang L, Liu Q, Jamieson AC, Liang Y, Qi H, Li PX, Chen B, Mendel MC, Zhong X, Lee YL, Eisenberg SP, Spratt SK, Case CC, Wolffe AP. 2001. Regulation of an endogenous locus using a panel of designed zinc

finger proteins targeted to accessible chromatin regions. Activation of vascular endothelial growth factor A. J Biol Chem 276:11323–11334.

Liu R, Paxton WA, Choe S, Ceradini D, Martin SR, Horuk R, MacDonald ME, Stuhlmann H, Koup RA, Landau NR. 1996. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. Cell 86:367–377.

Lupo A, Cesaro E, Montano G, Zurlo D, Izzo P, Costanzo P. 2013. KRAB-Zinc finger proteins: A repressor family displaying multiple biological functions. Curr Genomics 14:268–278.

Maeder ML, Thibodeau-Beganny S, Osiak A, Wright DA, Anthony RM, Eichtinger M, Jiang T, Foley JE, Winfrey RJ, Townsend JA, Unger-Wallace E, Sander JD, Müller-Lerch F, Fu F, Pearlberg J, Göbel C, Dassie JP, Pruett-Miller SM, Porteus MH, Sgroi DC, Iafrate AJ, Dobbs D, McCray PB, Cathomen T, Voytas DF, Joung JK. 2008. Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Mol Cell 31:294–301.

Mandell JG, Barbas CF. 2006. Zinc Finger Tools: Custom DNA-binding domains for transcription factors and nucleases. Nucleic Acids Res 34:W5 16–W523.

Mechetin GV, Zharkov DO. 2014. Mechanisms of diffusional search for specific targets by DNA-dependent proteins. Biochemistry (Mosc) 79:496–505.

Miller J, McLachlan AD, Klug A. 1985. Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. EMBO J 4: 1609–1614.

Moore M, Klug A, Choo Y. 2001. Improved DNA binding specificity from polyzinc finger peptides by using strings of two-finger units. Proc Natl Acad Sci USA 98:1437–1441.

Neuteboom LW, Lindhout BI, Saman IL, Hooykaas PJ, van der Zaal BJ. 2006. Effects of different zinc finger transcription factors on genomic targets. Biochem Biophys Res Commun 339:263–270.

Nomura W, Barbas CF. 2007. In vivo site-specific DNA methylation with a designed sequence-enabled DNA methylase. J Am Chem Soc 129:8676–8677.

Orlando SJ, Santiago Y, DeKelver RC, Freyvert Y, Boydston EA, Moehle EA, Choi VM, Gopalan SM, Lou JF, Li J, Miller JC, Holmes MC, Gregory PD, Urnov FD, Cost GJ. 2010. Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. Nucleic Acids Res 38:e152.

Pabo CO, Peisach E, Grant RA. 2001. Design and selection of novel Cys2His2 zinc finger proteins. Annu Rev Biochem 70:313–340.

Pavletich NP, Pabo CO. 1991. Zinc finger-DNA recognition: Crystal structure of a Zif268-DNA complex at 2.1. A. Science 252:809–817.

Peisach E, Pabo CO. 2003. Constraints for zinc finger linker design as inferred from X-ray crystal structure of tandem Zif268-DNA complexes. J Mol Biol 330:1–7.

Pelham HR, Brown DD. 1980. A specific transcription factor that can bind either the 5S RNA gene or 5S RNA. Proc Natl Acad Sci USA 77:4170–4174.

Persikov AV, Singh M. 2014. De novo prediction of DNA-binding specificities for Cys2His2 zinc finger proteins. Nucleic Acids Res 42:97–108.

Persikov AV, Wetzel JL, Rowland EF, Oakes BL, Xu DJ, Singh M, Noyes MB. 2015. A systematic survey of the Cys2His2 zinc finger DNA-binding landscape. Nucleic Acids Res 43:1965–1984.

Poorey K, Viswanathan R, Carver MN, Karpova TS, Cirimotich SM, McNally JG, Bekiranov S, Auble DT. 2013. Measuring chromatin interaction dynamics on the second time scale at single-copy genes. Science 342:369–372.

Ptashne M, Gann AA. 1990. Activators and targets. Nature 346:329–331.

Ramakrishna S, Kim YH, Kim H. 2013. Stability of zinc finger nuclease protein is enhanced by the proteasome inhibitor MG132. PLoS One 8:e54282.

Ramirez CL, Foley JE, Wright DA, Müller-Lerch F, Rahman SH, Cornu TI, Winfrey RJ, Sander JD, Fu F, Townsend JA, Cathomen T, Voytas DF, Joung JK. 2008. Unexpected failure rates for modular assembly of engineered zinc fingers. Nat Methods 5:374–375.

Rebar EJ, Pabo CO. 1994. Zinc finger phage: Affinity selection of fingers with new DNA-binding specificities. Science 263:671–673.

Rhee HS, Pugh BF. 2011. Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. Cell 147:1408–1419.

Rivenbark AG, Stolzenburg S, Beltran AS, Yuan X, Rots MG, Strahl BD, Blancafort P. 2012. Epigenetic reprogramming of cancer cells via targeted DNA methylation. Epigenetics 7:350–360.

Rohs R, Jin X, West SM, Joshi R, Honig B, Mann RS. 2010. Origins of specificity in protein-DNA recognition. Annu Rev Biochem 79:233–269.

Sander JD, Ramirez CL, Linder SJ, Pattanayak V, Shoresh N, Ku M, Foden JA, Reyon D, Bernstein BE, Liu DR, Joung JK. 2013. In silico abstraction of zinc finger nuclease cleavage profiles reveals an expanded landscape of off-target sites. Nucleic Acids Res 41:e181.

Sasai N, Nakao M, Defossez PA. 2010. Sequence-specific recognition of methylated DNA by human zinc-finger proteins. Nucleic Acids Res 38: 5015–5022.

Seifart KH, Wang L, Waldschmidt R, Jahn D, Wingender E. 1989. Purification of human transcription factor IIIA and its interaction with a chemically synthesized gene encoding human 5 S rRNA. J Biol Chem 264:1702–1709.

Slattery M, Zhou T, Yang L, Dantas Machado AC, Gordân R, Rohs R. 2014. Absence of a simple code: How transcription factors read the genome. Trends Biochem Sci 39:381–399.

Suske G 1999. The Sp-family of transcription factors. Gene 238:291–300.

Tachikawa K, Schröder O, Frey G, Briggs SP, Sera T. 2004. Regulation of the endogenous VEGF-A gene by exogenous designed regulatory proteins. Proc Natl Acad Sci USA 101:15225–15230.

Tebas P, Stein D, Tang WW, Frank I, Wang SQ, Lee G, Spratt SK, Surosky RT, Giedlin MA, Nichol G, Holmes MC, Gregory PD, Ando DG, Kalos M, Collman RG, Binder-Scholl G, Plesa G, Hwang WT, Levine BL, June CH. 2014. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. N Engl J Med 370:901–910.

von Hippel PH, Berg OG. 1989. Facilitated target location in biological systems. J Biol Chem 264:675–678.

Wang P, Qu X, Wang X, Zhu X, Zeng H, Chen H, Zhu H. 2014a. Specific reactivation of latent HIV-1 with designer zinc-finger transcription factors targeting the HIV-1 5′-LTR promoter. Gene Ther 21:490–495.

Wang S, Linde MH, Munde M, Carvalho VD, Wilson WD, Poon GM. 2014b. Mechanistic heterogeneity in site recognition by the structurally homologous DNA-binding domains of the ETS family transcription factors Ets-1 and PU.1. J Biol Chem 289:21605–21616.

Wilber A, Tschulena U, Hargrove PW, Kim YS, Persons DA, Barbas CF, Nienhuis AW. 2010. A zinc-finger transcriptional activator designed to interact with the gamma-globin gene promoters enhances fetal hemoglobin production in primary human adult erythroblasts. Blood 115:3033–3041.

Wu H, Yang WP, Barbas CF. 1995. Building zinc fingers by selection: Toward a therapeutic application. Proc Natl Acad Sci USA 92:344–348.

Wuttke DS, Foster MP, Case DA, Gottesfeld JM, Wright PE. 1997. 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 273:183–206.

Yan W, Imanishi M, Futaki S, Sugiura Y. 2007. Alpha-helical linker of an artificial 6-zinc finger peptide contributes to selective DNA binding to a discontinuous recognition sequence. Biochemistry 46:8517–24.