Nucleic Acids Research Advance Access published January 8, 2013

Nucleic Acids Research, 2013, 1–11
doi:10.1093/nar/gks1357

Using defined finger–finger interfaces as units of assembly for constructing zinc-finger nucleases

Cong Zhu[1], Ankit Gupta[1,2], Victoria L. Hall[1], Amy L. Rayla[1], Ryan G. Christensen[3], Benjamin Dake[4], Abirami Lakshmanan[1], Charlotte Kuperwasser[4], Gary D. Stormo[3] and Scot A. Wolfe[1,2,*]

[1]Program in Gene Function and Expression, University of Massachusetts Medical School, Worcester, MA, USA 01605, [2]Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA, USA 01605, [3]Department of Genetics, Washington University School of Medicine, St Louis, MO, USA 63108 and [4]Molecular Oncology Research Institute (MORI), Tufts University School of Medicine, Boston, MA, USA 02111

Received November 7, 2012; Revised December 4, 2012; Accepted December 5, 2012

ABSTRACT

Zinc-finger nucleases (ZFNs) have been used for genome engineering in a wide variety of organisms; however, it remains challenging to design effective ZFNs for many genomic sequences using publicly available zinc-finger modules. This limitation is in part because of potential finger–finger incompatibility generated on assembly of modules into zinc-finger arrays (ZFAs). Herein, we describe the validation of a new set of two-finger modules that can be used for building ZFAs via conventional assembly methods or a new strategy—finger stitching—that increases the diversity of genomic sequences targetable by ZFNs. Instead of assembling ZFAs based on units of the zinc-finger structural domain, our finger stitching method uses units that span the finger–finger interface to ensure compatibility of neighbouring recognition helices. We tested this approach by generating and characterizing eight ZFAs, and we found their DNA-binding specificities reflected the specificities of the component modules used in their construction. Four pairs of ZFNs incorporating these ZFAs generated targeted lesions *in vivo*, demonstrating that stitching yields ZFAs with robust recognition properties.

INTRODUCTION

Zinc-finger nucleases (ZFNs) are chimeric fusions between a programmable zinc-finger array (ZFA) and the nuclease domain of FokI (1). These artificial nucleases are powerful tools for genome modification, as they can generate a site-specific double-strand break (DSB) within the genome to promote a number of different types of genome editing (2,3). ZFNs can disrupt the function of a protein-coding gene when an imprecisely repaired DSB creates a frameshift in the coding sequence (2,3). These DSBs can also be used for the introduction of tailor-made changes to the genome by dramatically stimulating the rate of homologous recombination at a locus with an exogenously supplied donor DNA (2,3). ZFNs have been used in a variety of model and non-model organisms to facilitate reverse genetic approaches to study gene function or construct disease models for analysis (3–8). Engineered nucleases also have potential application as gene therapy-based therapeutics (9–15), where the first of these reagents are now in advanced clinical trials for treatment of AIDS (16).

The use of ZFNs has primarily been limited by the ease with which ZFAs can be created to selectively target a desired genomic region. Excluding purchase from commercial sources, selection-based approaches provide the most reliable method for creating ZFAs with novel DNA-binding specificity (17–23). Bacterial-based selection systems have somewhat simplified the process of creating ZFAs with novel specificity (24–26), but these systems still require effort on the part of end-users to generate functional constructs. Many zinc-finger proteins bind to DNA in an apparently modular fashion (27–32). Based on this supposition, a comprehensive archive of single-finger modules should enable the ready assembly of any multi-finger ZFA, where the resultant recognition site is a composite of the specificities of the incorporated finger modules (31,33–35). Using this approach, many

*To whom correspondence should be addressed. Tel: +1 508 856 3953; Fax: +1 508 856 5460; Email: scot.wolfe@umassmed.edu

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

© The Author(s) 2013. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from http://nar.oxfordjournals.org/ at University of Massachusetts Medical School on February 14, 2013

laboratories have generated ZFAs for incorporation into ZFNs (36–43). However, success rates with such modularly assembled ZFNs have typically been modest (<30%) (39–41). The inconsistency in these systems could reflect insufficient specificity or affinity of the modules within the published archives (44), or incompatibility between the assembled fingers, which has been dubbed 'context-dependent effects' (17,28,31,45). The primary source of this incompatibility likely resides with mismatched residues at the finger–finger interface that degrade or alter specificity (41) in some cases because of recognition overlap between neighbouring fingers (30,46–48).

The impact of interface incompatibility can be reduced by limiting the number of unproven finger–finger interfaces that are generated during assembly (2). The use of two-finger modules for ZFA assembly reduces, but does not eliminate, the number of unproven interfaces (49–51). Unproven interfaces can be eliminated entirely through the context-dependent assembly (CoDA) of two-finger modules, where three-finger ZFAs are constructed from two-finger units with a common overlapping central finger (52). CoDA-generated ZFNs have favourable success rates (~50%) *in vivo*; however, the vast majority of the modules that have been tested as ZFNs recognize NG-type (gn**NG**nn) dinucleotide junctions at the finger–finger interface, which are well understood and, therefore, are not the limiting interface for the expansion of ZFN targeting density. Moreover, some of the CoDA modules assigned to target non-NG junctions actually prefer NG junction sequences (50).

Recently, we described the selection of two-finger modules recognizing GR**NN**YG dinucleotide junctions using our bacterial one-hybrid (B1H) system (50). This approach focused on randomizing the recognition residues at the finger–finger interface in a library of two-finger modules to select optimal modules for each target sequence. The specificity of the selected modules was, subsequently, validated to identify modules with the most favourable recognition properties. Additional mutagenesis expanded the breadth of this archive to contain modules spanning 162 six base pair target sites that can be used together or in conjunction with other single-finger archives (41) for the assembly of ZFAs. This archive expands the current collection of publicly available two-finger modules, in particular those that recognize non-NG interfaces. However, their assembly into larger ZFAs remains predicated on the concatenation of modules at dinucleotide junctions that have well-defined sequence preferences (e.g. GK or AN) (50). Appropriate interface residues for many of the other dinucleotide junctions remain poorly defined. Moreover, based on our analysis of zinc-finger specificity, it seems that the appropriate choice of these recognition residues can be impacted by the identity of the residues at position +3 in each finger (Gupta and Wolfe, unpublished results), confirming the complexity in zinc-finger–DNA recognition that has been observed in other mutagenesis studies (45,53,54).

Most ZFA-assembly methods use finger archives composed of single- or multi-finger modules, where these units are delimited by the structural motif of the zinc finger. However, if finger–finger interfaces represent critical grammar for the successful assembly of functional ZFAs, then construction units wherein this most dynamic feature of recognition is fixed may yield higher success rates. This type of assembly approach has been used in the construction of three-finger ZFAs from 1.5 finger modules by Choo and colleagues (19), where an intervening phage display selection step could be used to optimize the recognition properties of the assembled proteins. We extended this approach by choosing units of assembly with fixed elements of overlap, such that an additional selection is not required, and ZFAs containing any number of fingers could be assembled if complementary monomeric units were present in the archive. Moreover, this method allows the construction of hybrid fingers with novel specificity that is not present within the archive. To serve as an initial archive for this assembly approach, we selected a set of GANNAG two-finger modules that can be assembled into ZFAs either through the standard modular-assembly or via finger 'stitching'. We demonstrated that ZFAs assembled through either method function robustly as ZFNs when assembled into nucleases. These results highlight the advantage of using defined finger–finger interfaces for the construction of artificial ZFAs.

MATERIALS AND METHODS

Animal husbandry

Zebrafish were handled according to established protocols (55) and in accordance with Institutional Animal Care and Use Committee (IACUC) guidelines of the University of Massachusetts Medical School.

2F-library construction

The two-finger (2F) library was designed with scheme RSDNLXX XXXNLTR, using codons VNS(+5)VNS(+6) NNW(−1)NNW(+1)NNW(+2) (V: A/C/G, N:A/C/G/T, S:G/C and W:A/T) for the five randomized residues. 2F-libraries were constructed as previously described (50). Briefly, individual F1 and F2 libraries were independently constructed via cassette mutagenesis of annealed randomized oligonucleotides into pBluescript vector containing the appropriate zinc finger backbone derived from Zif268. The 2F-library was constructed from these single-finger libraries by overlapping polymerase chain reaction (PCR) assembly. This 2F-library was then ligated into the B1H expression vector 1352-omega-UV2 between unique BssHII and Acc65I restriction enzyme sites, such that the ω-subunit of the RNA polymerase is fused at the N-terminus of the two zinc fingers, and the Engrailed homeodomain follows the fingers at the C-terminus (Figure 1). After electroporation into bacterial cells, 1×10^8 cells (five times the theoretical size of the library) were plated on ten 2xYT–carbenicillin plates (150 × 15 mm) and grown at 37°C for 14 h. 1352-omega-UV2 plasmids containing the 2F-library were isolated from pooled surviving colonies and used for selections.

Downloaded from http://nar.oxfordjournals.org/ at University of Massachusetts Medical School on February 14, 2013

Zinc-finger binding site cloning and 2F-module B1H selections

The 16 GANNAG zinc-finger binding sites (ggccTAATTACCTGANNAGGacg) were cloned between the EcoRI and NotI sites in the pH3U3-mcs reporter vector (57). The homeodomain (Engrailed) binding site TAATTA (underlined) is present 3 bp away and on the strand opposite to the zinc-finger binding site to minimize any interference between the homeodomain and the zinc fingers. Selections for 2F-modules were performed as described previously (50). The zinc-finger library (20 ng) and the reporter vector (1 μg) containing the zinc-finger target site were co-transformed via electroporation in the selection strain that lacks endogenous expression of the ω-subunit of RNA polymerase (US0*ΔhisBΔpyrFΔrpoZ*). The 2×10^7 co-transformed cells were plated on selective NM minimal medium plates containing different concentrations of competitive inhibitor (3-aminotriazole; 3-AT) and inducer (isopropyl-β-D-thiogalactoside; IPTG) and grown at 37°C until a moderate number of colonies were visible. Selections were performed under different stringencies for each target site by varying 3-AT or IPTG concentration to achieve about surviving 1000 colonies per plate. In some instances, it was necessary to further increase stringency by adjusting the strength of the binding site. This was accomplished by altering either the engrailed binding site from canonical sequence TAATTA to mutant sequence TAATGC or the 6 bp 2-finger module target site from GANNAG to TANNAG. Post-selection, 2F-modules from eight surviving colonies of various sizes (typically four large, two medium and two small) were sequenced to identify functional amino acid sequences for further evaluation. The success of the selection was judged by the diversity of sequences obtained from these selections, with the expectation that successful selections will converge on a small number of functional residues at the critical recognition positions.

Cloning B1H-selected 2F modules into 3F F1-GCG constructs

To determine the binding specificities of 2F-modules, a 'GCG' binding anchor zinc finger (recognition helix: RSDTLAR) was fused at the N-terminus of the 2F-module via overlapping PCR. After overlapping PCR, the 3F-ZFA was cloned into 1352-omega-UV2 vector between the Acc65I and BamHI sites for expression as an omega fusion.

Constrained variation-B1H method

To determine binding site specificities of 2F-modules, the constrained variation-B1H assay was performed as described previously (56). After transformation into the selection strain, 1×10^6 cells containing the zinc-finger plasmid (1352-omega-UV2-ZFP) and the 6 bp randomized binding site library plasmid (pH3U3) were plated on selective NM minimal medium plates (100 × 15 mm) containing 50 μM of IPTG and 1 or 2 mM of 3-AT and grown at 37°C for 22–30 h. The surviving colonies were pooled, and the binding site plasmid was isolated for identification of the functional DNA sequences. The binding site region was PCR amplified, barcoded and sequenced via Illumina sequencing, and then binding specificities were determined from these data using the log-odds method (50,56,58).

Creating multi-finger ZFPs

All stitching finger ZFPs and six-finger traditional-modular-assembly ZFPs were created by gene synthesis through Genscript USA (Piscatacwa, NJ, USA) or Invitrogen (Calsbad, CA, USA). In some cases, the specificity of the two-finger module was adjusted from GDNNMG to GDNNMA by altering the residues at positions −1, 1 and 2 in the N-terminal cap from RSD to QRG (50). The traditional-modular-assembly four- or five-finger ZFPs were created based on existing six-finger ZFPs by PCR amplification of the desired finger subsets. These ZFPs were flanked by Acc65I/BamHI sites to facilitate the cloning of these ZFPs into 1352-omega-UV2 vector for B1H-binding site selection or pCS2-DD or RR vectors for creating ZFNs for activity assay.

B1H-binding site selections using the 28 bp library

The selections for 3F and 4F ZFAs were performed as previously described (41,58). The $1–5 \times 10^7$ US0 selection strain cells co-transformed with the 1352-omega-UV2 ZFA expression plasmid and the 28 bp pH3U3 library plasmid were plated on NM minimal medium selective plates lacking uracil and containing 3-AT (2.5, 5 or 10 mM) as the competitor and grown at 37°C for 36–72 h. The number of surviving bacterial colonies on each plate was estimated, and then these colonies were pooled, and the population of recovered DNA sequences was determined via Illumina sequencing. Unique sequences were ranked based on the number of recovered reads. From this list, an overrepresented sequence motif was determined with MEME (59,60) using as input the number of unique sequences from the top of the list that correspond to the estimated number of colonies on the selection plate (typically >1000).

ZFN injections and lesion analysis

For gene targeting in zebrafish, ZFAs were cloned in pCS2 vectors containing the DD/RR obligate heterodimer version of the *FokI* nuclease domain (61,62). pCS2–ZFN constructs were linearized with NotI, and mRNA was transcribed using the Message Machine SP6 kit from Ambion. ZFN mRNAs were injected into the blastomere of one-cell-stage zebrafish embryos as previously described (24). ZFN-injected embryos (8–30) with normal appearance and uninjected embryos were collected 24 h post-fertilization (h.p.f.) and incubated in 50 mM of NaOH (15 μl/embryo) for 15 min at 95°C to isolate genomic DNA and then neutralized with 0.5 M of Tris–HCl (4 μl/embryo). The DNA solution was centrifuged for 1 min at 13 000 r.p.m., and supernatant was taken for lesion analysis. For initial validation of ZFN activity, the region flanking the ZFN target site was amplified using the Phire Hot Start DNA Polymerase (New England Biolabs), and restriction fragment length polymorphism (RFLP) analysis or *T7 Endonuclease I*

Downloaded from http://nar.oxfordjournals.org/ at University of Massachusetts Medical School on February 14, 2013

(or T7E1, New England Biolabs) assay was performed (39,63). For T7E1 assay, The PCR products were denatured and re-annealed using the following program: 95°C for 180 s, 85°C for 20 s, 75°C for 20 s, 65°C for 20 s, 55°C for 20 s, 45°C for 20 s, 35°C for 20 s and 25°C for 20 s with a 0.1°C/s decrease rate in between steps. This allows the hybridization of mutant and wild-type DNA strands to form duplex DNA, which then can be detected by T7E1 nuclease assay. After re-annealing, 10 µl of the products were then treated with 2U of T7E1 for 1 h in the presence of NEB buffer 2. The T7E1 treated and untreated DNA was then subjected to electrophoresis on 3% agarose gel (Ultra pure 1000, Invitrogen). The gel images were analysed using ImageJ, and lesion rate was calculated as described previously (64).

Transfection of ZFNs in HEK 293 T cells

Low passage HEK 293T cells were plated in 12-well plates overnight, such that they were 75% confluent the next day. Transfection was performed using TransIt-LT1 (Mirus Bio) according to the manufacturer's protocol with 1 µg of total DNA made up of 40% for each ZFN and 20% of a GFP transfection-efficiency reporter. Cells were passaged once to a larger plate and harvested ~3 days post-transfection by lysis in QuickExtract (EpiCentre) solution using 150 µl per million cells before PCR amplification of the ZFN cut sites.

LacZα blue–white assay

To further determine the types of the mutations induced by the ZFNs in the zebrafish or human cell genome, we also cloned the target fragments, such that they generate a sequence lacking stop codons in-frame with LacZα gene on pBluescript-KS(-) vector between XbaI and KpnI sites. The lengths of the target fragments cloned were between 60 and 90 bp, such that it would have minimal impact on the function of the translated LacZ peptide in the α-complementation assay using XL1-Blue *Escherichia coli* cell (the amplicons were chosen, such that there are no stop codons in the reading frame). The small indels induced by the ZFN through non-homologous end joining (NHEJ) pathway would disrupt the reading frame of the LacZα on the pBluescript-KS (-) vector, so that there is no functional LacZα peptide produced and consequently no active β-galactosidase. *E. coli* colonies harbouring such fragments seem to be white on plates containing X-gal and IPTG. On the other hand, colonies containing wild-type sequence without indels would produce active LacZα peptide and appear to be blue colonies on these plates. The identities of the sequence (or the types of the lesions) are then identified through sequencing the inserts from white colonies.

RESULTS

Selection of two-finger modules recognizing GANNAG target sites

Leveraging our previous success in selecting functional two-finger modules recognizing GRNNYG sites using the B1H system (50), we used a similar two-stage approach for the identification of modules targeting GANNAG sites (Figure 1). In the first stage, modules complementary to each of the 16 GANNAG interfaces were selected from a two-finger library that consists of $\sim 2 \times 10^7$ variants, where positions +5 and +6 of the N-terminal finger and positions −1, +1 and +2 of the C-terminal finger were randomized. The other recognition positions within the fingers were fixed, where Asn is present at position +3 of each finger to mediate Adenine recognition. The stringency of the selection conditions for each target site was optimized to obtain a few hundred surviving colonies on each selection plate to restrict survival to the most favourable fingers for target recognition.

The resulting pools of selected modules for the sixteen gANNAg target sites trended towards a consensus sequence (Figure 2, Supplementary Table S1). Comparison of these recovered sequences with clones recovered from the corresponding gANNCg selections (50) reveals the complexity inherent in recognition at the finger–finger interface. Fundamentally, the two-finger libraries used in these selections differed at only a single position (either Thr or Asn present at position +3 of Finger 1), yet in some cases, this led to a dramatic difference in the recovered residues at the interface positions. Some of the dinucleotide junctions, in particular those for the AN-type junctions, result in the recovery of similar finger interface residues in both the ANNA and ANNC selections. However, others, in particular those for the CN-type junctions, result in the recovery of different residue sets at the finger–finger interface from each library. This disparity alludes to the presence of context-dependent effects at the finger-finger interface.

Identification of two-finger modules preferring each dinucleotide junction

To identify two-finger modules with the most favourable recognition properties for each GANNAG dinucleotide junction, we analysed their DNA-binding specificity using the B1H system (41,57). We objectively selected a small number of clones from each recovered pool for analysis. The DNA-binding specificity of each module was characterized as a three-finger ZFA by appending a single finger recognizing a 'GCG' triplet to the N-terminus. In this context, each two-finger module was characterized using a reporter system containing a six base pair randomized binding site library flanking the finger 1 recognition sequence (50,56) (Figure 1B). Recovered binding sites from the randomized library for each two-finger module were pooled and then characterized by Illumina sequencing to determine the preferred recognition motif. Using this approach, we characterized 77 candidate two-finger modules to identify those with the strongest preference for each of the 16 GANNAG dinucleotide junctions (Supplementary Figure S1). Based on this analysis, we have identified modules that are compatible with each of the 16 dinucleotide junctions (Figure 3; Supplementary Table S2). For 14 of these modules, the desired dinucleotide junction is the most

Downloaded from http://nar.oxfordjournals.org/ at University of Massachusetts Medical School on February 14, 2013

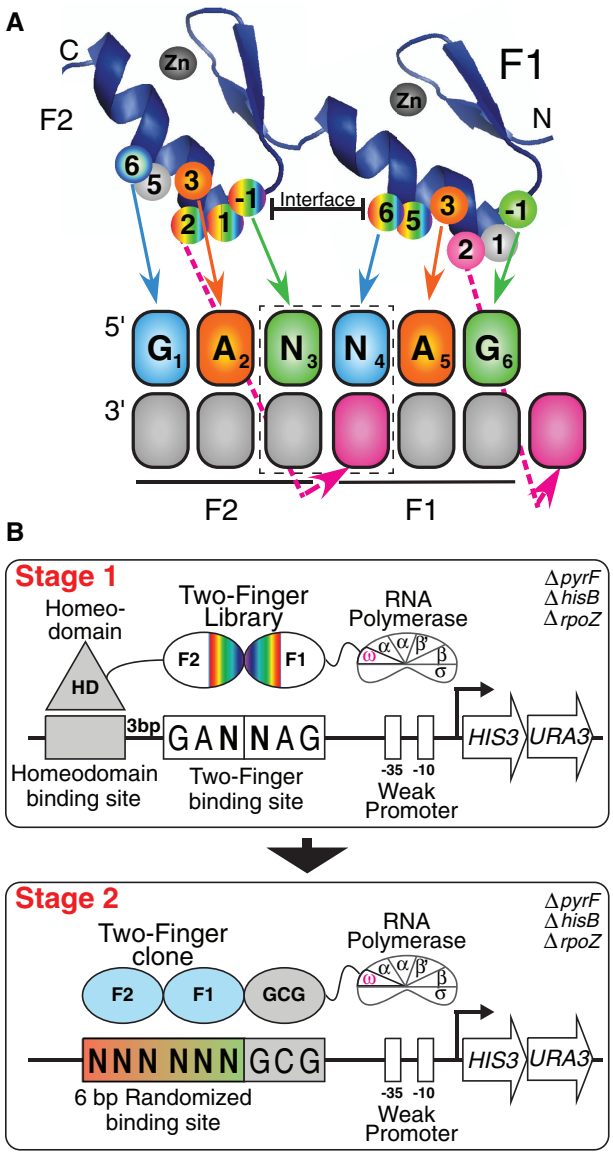


Figure 1. Selection of GANNAG finger sets. (**A**) Schematic of two-finger ZFP library used in these selections with the specificity determinants mapped to their recognition positions in their binding site. The dashed box indicates the position of the dinucleotide junction. This library contains randomized amino acids at the finger–finger interface at positions +5 and +6 of finger 1 (randomized with VNS codons) and positions −1, +1 and +2 of finger 2 (randomized with NNW codons), where the numbering scheme refers to the position of the residue relative to the start of the recognition helix. The finger 1 residues at positions −1, 1 and 2 (R, S and D) represent the N-terminal cap, and the finger 2 residues at positions 5 and 6 (T and R) represent the C-terminal cap. (**B**) Schematic representation of the two-stage process used to identify two-finger modules with the desired sequence preference. In Stage 1, the B1H system is used to select two-finger modules complementary to each target site. The randomized two-finger module library is fused between the DNA-binding domain of the Engrailed homeodomain and the ω-subunit of the RNA polymerase. The fixed 6-bp GANNAG target site is present on the His3/Ura3 reporter plasmid between the homeodomain binding site and the −35 box. In Stage 2, the DNA binding specificity of candidate two-finger modules obtained from the first stage of the selection are interrogated. Each two-finger module is fused to an N-terminal finger (RSDTLAR) that binds to the 'GCG' triplet adjacent to the 6 bp randomized zinc-finger binding region on the reporter plasmid. The recovered binding sites are determined by Illumina sequencing, and then a binding site motif is calculated from these sequences (56).

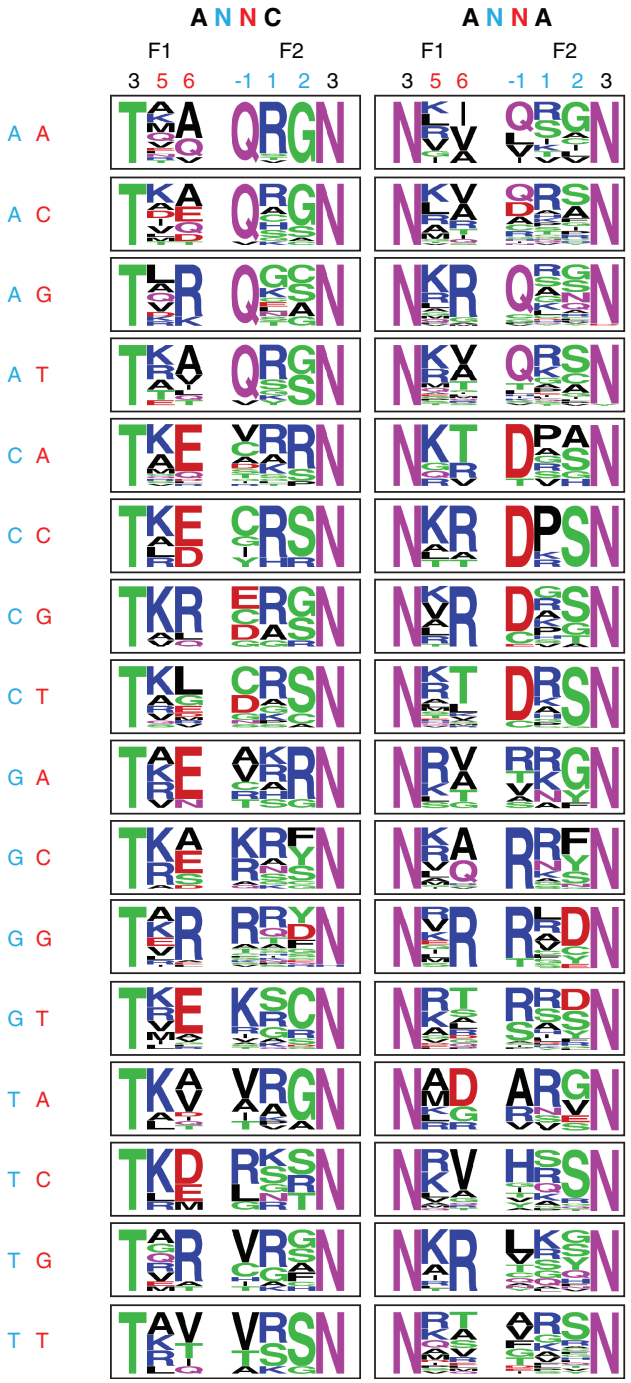


Figure 2. Comparison of the preferred residues recovered for the dinucleotide junctions in gANNCg (50) and gANNAg selections. In both selections, residues were selected at positions +5 and +6 of finger 1 and positions −1, +1 and +2 of finger 2. Amino acids at positions +3 of the two fingers were fixed in the library. The recovered sequences are displayed as frequency logos. The only fixed difference between the two-finger libraries used in these selections occurs at position +3 of the finger 1, which is threonine (T) in the gANNCg selections and asparagine (N) in gANNAg selections.

prevalent sequence recovered in the binding site selections, although in some cases, the preference for this sequence is only modest. For the two remaining junctions (gaACag & gaCCag), we identified modules that recognize

Downloaded from http://nar.oxfordjournals.org/ at University of Massachusetts Medical School on February 14, 2013

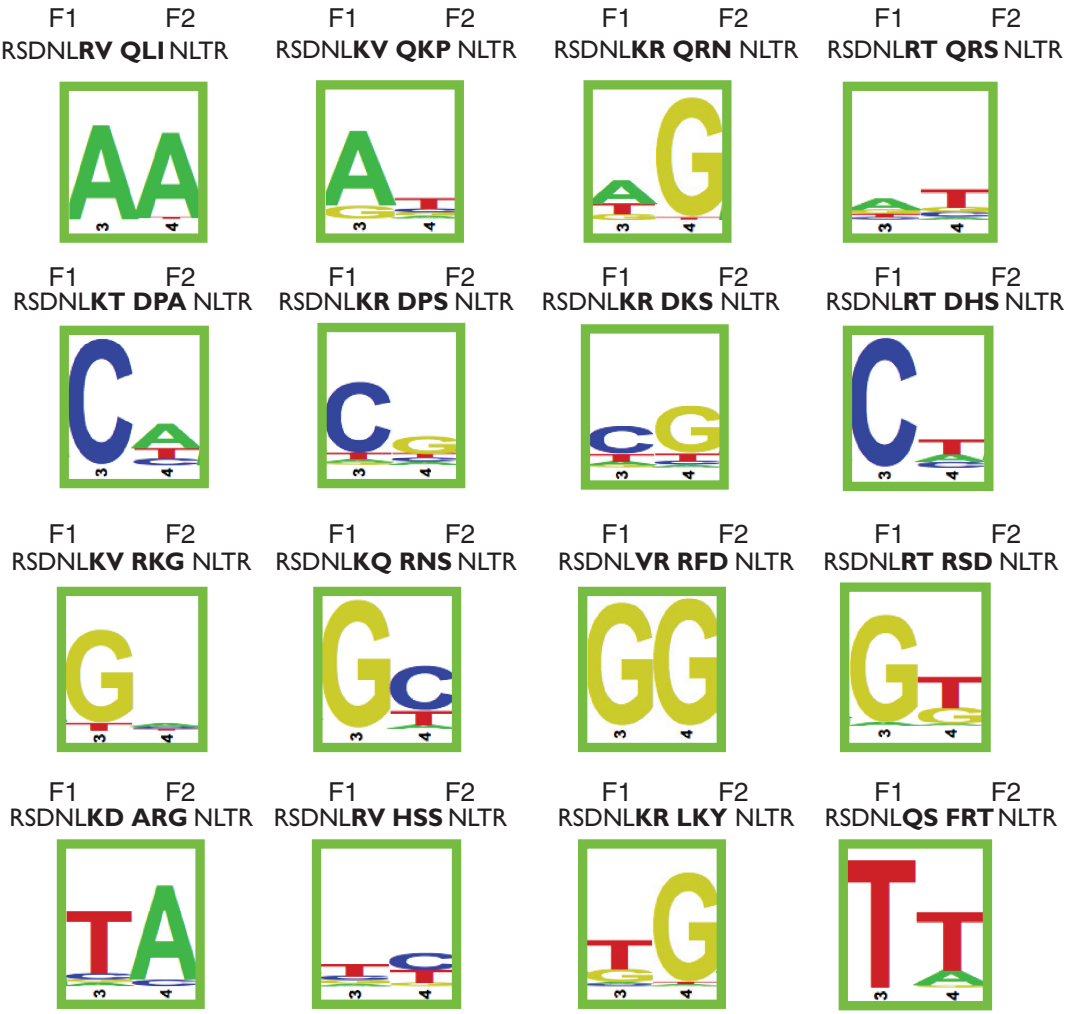


Figure 3. DNA binding specificities of the two-finger GANNAG modules with the most favourable specificity for each of the 16 different dinucleotide junctions. The DNA binding specificities were determined using B1H system (Supplementary Figure S1), and the sequence logo at the 2-bp interface is shown. Amino acid residues at positions −1 to +6 of the recognition helix of each finger are shown. The five amino acid residues recovered in the ANNA interface selection are indicated in bold font.

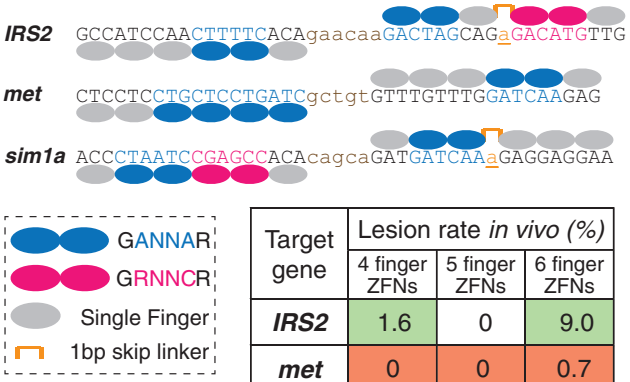
the intended target site as the second and third most preferred site, respectively. These modules constitute our validated GANNAG archive for ZFA assembly.

Demonstrating the functionality of the GANNAG two-finger modules

The validated GANNAG two-finger modules expand our existing archives (41,50) for generating ZFAs via modular assembly. To demonstrate the functionality of these new modules, we generated ZFNs via a traditional modular assembly approach for three targets (*IRS2, met* and *sim1a*) in the zebrafish genome, where each ZFA incorporates at least one new GANNAG module (Figure 4). In some cases, the specificity of the two-finger module was adjusted from GANNAG to GANNAA by altering the residues (from RSD to QRG) in the N-terminal cap (50). These targets were chosen, such that six-finger ZFAs can be constructed for each half site providing the opportunity to compare the activity of four-, five- and six-finger proteins (Supplementary Table S3). In two of these ZFAs, we used a recently described THPRAPIPKP linker between fingers from Sangamo BioSciences that allows a single base pair to be skipped between intervening modules (65).

ZFNs containing these ZFAs were tested as pairs of four-, five- or six-finger proteins for each target site in zebrafish embryos. Dose response curves were used to identify an optimal concentration of ZFN mRNA for each target. This initial survey revealed that all of the *met* ZFNs were highly toxic to embryos. Based on this analysis, each ZFN dose was calibrated to a level where ~50% of the embryos developed normally at 24 h.p.f.. ZFN-injected normal embryos were subsequently analysed for lesions by T7 Endonuclease I (T7EI) analysis of PCR products spanning the target site (39,63) (Supplementary Figure S2). This revealed that the *IRS2* and *sim1a* ZFNs were active, whereas the *met1* ZFNs showed minimal activity (Figure 4). Interestingly the *IRS2* and *sim1a* ZFNs displayed opposite trends

Downloaded from http://nar.oxfordjournals.org/ at University of Massachusetts Medical School on February 14, 2013



Target gene	Lesion rate *in vivo* (%)		
	4 finger ZFNs	5 finger ZFNs	6 finger ZFNs
IRS2	1.6	0	9.0
met	0	0	0.7
sim1a	11.9	3.2	1.2

Figure 4. Schematic representation of the three pairs of ZFNs targeting *IRS2*, *met* and *sim1a*. Only the six-finger constructs are shown. Finger number was reduced by progressively removing fingers from the N-terminus of these constructs. The positions of two-finger modules (GANNAG) described herein or fingers from other archives [GRNNCG (50) and single fingers (41)] within these ZFAs are indicated, as is the position of the THPRAPIPKP linker that allows a single base pair (underlined lowercase base) to be skipped between intervening modules (65). The efficiency of lesion generation in zebrafish embryos by these ZFAs as a function of the number of fingers is indicated in the chart, where red highlights indicate the high toxicity of the *met* ZFNs.

with regards to activity and finger number. The *IRS2* ZFNs had the greatest activity with the longer ZFAs, whereas the *sim1a* ZFNs displayed the opposite trend.

Stitching together ZFAs with novel DNA binding specificity

These newly characterized two-finger modules expand the archive of modules for generating ZFNs via traditional modular assembly (50). In addition these newly identified two-finger modules can serve as building blocks for a novel ZFA assembly method: finger stitching (Figure 5A). This new strategy takes advantage of a common feature of the two-finger modules targeting GANNAG interfaces. Each finger contains an asparagine (Asn) at position +3 of the recognition helix that recognizes A_2 and A_5 of the '$G_1A_2N_3N_4A_5G_6$' 6-bp target site (Figure 1A). We envisioned that the two Asn residues might serve as bookends for the selected interface residues recognizing the dinucleotide junction N_3N_4 and thereby preserve their DNA binding specificity on incorporation into multi-finger proteins. Thus, in this approach, ZFAs are assembled by joining interfaces from compatible two-finger modules that share identical residues at the +3 position to create large arrays. For example, a three-finger protein recognizing a sequence $GAN_3N_4\mathbf{A}N_6N_7AG$ would be constructed from two two-finger modules recognizing GAN_3N_4AG and GAN_6N_7AG, where the bold **A** indicates the position of recognition overlap. Because our previously selected GANNCG modules also contain Asn at position +3 within the C-terminal finger of the two-finger module, these can be incorporated to generate ZFAs recognizing $GAN_3N_4AN_6N_7CG$ target sites. In addition,

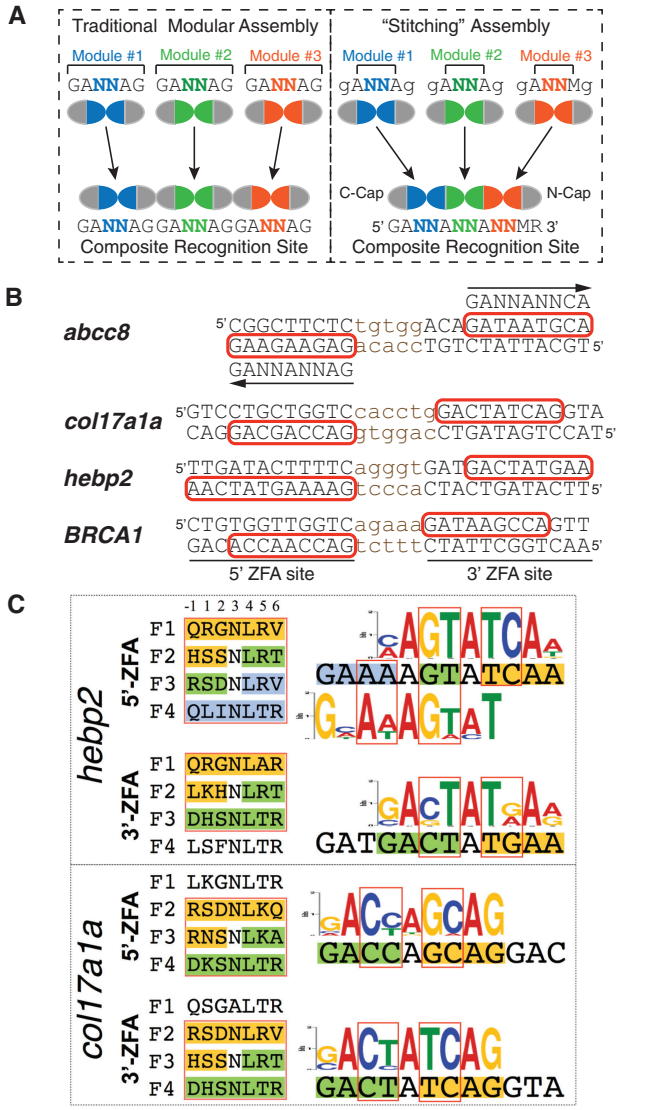


Figure 5. 'Finger stitching' for ZFA assembly using GANNAG modules. (**A**) Schematic comparison of the traditional modular assembly approach (left) with our stitching approach (right). Instead of assembling whole finger units, stitching assembles segments between the +3 positions of the neighbouring recognition helices and then caps these at the N- and C- terminus (N-cap and C-cap). Two additional components extend the targetable sequences: the ability to incorporate GANNCG modules as the last unit in a stitched array, and the use of an alternate N-cap specific for a 3′ adenine (50), which allows the final specified base to be either G or A depending on the choice of the N-cap. (**B**) Target sites for the four pairs of 'stitched' ZFNs, where the binding site for each monomer is indicated in capital letters and the recognition element of the three or four finger stitched ZFA is boxed in red on the primary recognition strand. In some cases, a ZFA contains three stitched fingers and one additional single-finger module. For the *abcc8* target site, the composite recognition site for the stitched portion of the array is indicated above or below the primary recognition sequence, where the arrow denotes the 5′–3′ orientation. (**C**) To assess the quality of ZFAs generated through this approach, we assembled three-finger stitched ZFAs spanning portions of the target site and determined their DNA binding specificities using the B1H system. The recognition helices for these fingers are indicated to the left of the target sites, where the positions of the stitched fingers are boxed in red. The segments of the stitched fingers that arise from a common 2F-module share a common colour. Likewise, the positions of the dinucleotide junctions between fingers in the recognition motifs for these fingers are boxed in red, and the subsites recognized by the stitched finger segments are differentially coloured.

Downloaded from http://nar.oxfordjournals.org/ at University of Massachusetts Medical School on February 14, 2013

we have developed an alternate N-terminal cap that is specific for adenine (50); consequently, the 3′-most base recognized by a stitched array can be G or A depending on the N-cap that is used. Similarly, the finger stitching can be used to create ZFPs of any desired length by extending the array through the overlap of additional modules off either terminus.

To demonstrate that this novel assembly approach can create functional ZFAs and ZFNs, we chose three target genes (*abcc8, col17a1a* and *hebp2*) in the zebrafish genome containing ZFN sites that could be targeted using 'stitched' ZFAs, and where these target sites contain non-NG interfaces within the stitched fingers (Figure 5B). We also designed a pair of ZFNs for a human target gene, *BRCA1*, applying the same criteria. Genes encoding these ZFAs were generated by gene synthesis, where canonical TG(E/Q)KP linkers were used to connect all fingers in these arrays. As a first step in validating this approach, we determined the DNA binding specificities of these ZFAs in the B1H system using a randomized 28 bp library (41,50). Because our 28 bp library ($\sim 10^8$ unique members) can more effectively sample all possible recognition sequences for a three-finger than four-finger ZFA, we determined the DNA binding specificities of three-finger subsets for many of these ZFAs to provide a clearer assessment of their specificity. The determined DNA-binding specificities demonstrate in many cases that ZFAs assembled using the stitching method recognize their intended target sites with reasonable fidelity both at the dinucleotide junction sequences and the neighbouring adenines (Figure 5C and Supplementary Figure S3). However, there are instances where the dinucleotide preference is more degenerate in the stitched fingers than observed in the parent modules (e.g. abcc8-3p, Supplementary Figure S3), suggesting that in some cases, these assemblies are influenced by context-dependent effects.

Stitched ZFAs yield functional ZFNs

Given the favourable specificity of the stitched ZFAs, we evaluated their functionality as ZFNs. For each of the ZFNs targeting the zebrafish genomic sites, the optimal mRNA concentration was determined via a dose response curve as previously described, and ZFN activity was assessed at the optimal dose in healthy embryos at 24 h.p.f. The activity of the *BRCA1* ZFNs was assessed by transfection of expression plasmids encoding these ZFNs into HEK 293T cells. Genomic DNA was harvested 64 h after transfection for lesion analysis. For all samples, lesion rates were determined by enzymatic digestion of PCR products spanning the target region (either T7EI or site-specific restriction enzyme) relative to untreated control samples. All four ZFN pairs induced lesions at frequencies between 1 and 11.4% (Table 1 and Supplementary Figures S4–S7), where the lesion sequences were consistent with the types of mutations expected for ZFN activity (Supplementary Figure S8).

Table 1. Lesion rates for 'stitched' ZFNs

Target gene	Lesion rate *in vivo* (%)
abcc8	1.1
col17a1a	11.4
hebp2	2.8
*brca1**	2.7

*In 293T cells.

DISCUSSION

Although ZFN technology has been successfully used in a multitude of systems for genomic modification (2,3), one of the major barriers in adoption is the need for a simple approach to generate functional ZFNs for nearly any target site. Traditional-modular-assembly of ZFAs, although becoming more facile as the quality of the finger archives improves, still suffers from either moderate success rates (39–41,51) or moderate targeting density (50). The functional assembly of these units is complicated by the influence of context-dependent effects at the finger–finger interface (17,28,31,41,45). The CoDA method described by the Zinc Finger Consortium bypasses this problem by using two separate archives of two-finger modules that share common N-terminal or C-terminal fingers, which permits the assembly of three-finger proteins through overlap at these common units (52). Although straightforward, this system is inherently limited to the creation of three-finger ZFAs, which can restrict the precision of these ZFNs in complex metazoan genomes, and assessment of its modules has focused on ZFAs recognizing NG-type junctions (Supplementary Table S4). We have sought to bypass this limitation through the development of a new assembly method wherein finger–finger interface units provide the grammar for assembly, which ensures that the finger–finger interface is always compatible. We assembled four ZFN pairs using this stitching approach focused on ZFAs that contain non-NG junctions at the finger recognition sequences. Remarkably, the specificity of the modules comprising these ZFAs on assembly was generally preserved when compared with the determined specificities of the primary two-finger modules that compose the archive. Moreover, all of these ZFNs were functional when tested in zebrafish or in human cells, demonstrating that this approach can produce ZFAs that can function in the context of a complex genome.

We believe that the success of this stitching approach stems primarily from using the +3 positions to demarcate the units of assembly. The recognition preference of residues is probably best understood for the +3 position in canonically binding fingers (54). Obviously, our stitching approach ignores potential context-dependent effects along each recognition helix that is splinted together from two different modules. For example, if one considers position +3 as a pivot point for the docking of the zinc finger within the major groove, the length and bulk of the residues at the flanking recognition positions (−1 and +6) may influence the geometry of

Downloaded from http://nar.oxfordjournals.org/ at University of Massachusetts Medical School on February 14, 2013

finger binding, which could lead to sub-optimal recognition in some instances for stitched fingers. Thus, we do not expect that this new methodology will be completely free of complications, but we anticipate that it will perform favourably when compared with the traditional modular assembly approach by minimizing the incompatibility of fingers that are joined together.

Our strategy is adapted from the 1.5-finger assembly method described by Isalan *et al.* (19); however, their strategy focused solely on the assembly of three-finger ZFAs and used selection in many cases to generate a functional three-finger protein. Our approach does not require selection and can be adapted for the creation of ZFAs of any desired length. One limitation of our approach is the requirement for a suitable archive of two-finger modules that can be used for targeting the desired DNA sequence. Currently, our archive is limited to the GANNAG and GANNCG two-finger modules, but this set can be readily expanded through the selection of additional archives of modules or by co-opting two-finger modules with good specificity from archives that have been generated for other systems (39,52). Although the available archive for this assembly is currently somewhat sparse, subsets of stitched modules can be combined through standard modular assembly with one- and two-finger units from existing archives to broaden the sequences that can be readily targeted.

To facilitate the discovery of ZFN target sites that are accessible using this approach, we have modified our existing website for the identification of ZFN target sites within a user-input sequence element (http://pgfe.umassmed.edu/ZFPmodularsearchV2.html) to include the incorporation of stitched finger sets or subsets within a ZFA. The Web interface ranks a set of target sites based on the quality of the ZFA that can be assembled and outputs the amino acid and DNA sequence for the ZFAs to facilitate their creation through gene synthesis. This new assembly method coupled with the standard modular assembly approach increases the density of ZFN target sites in the zebrafish genome to approximately one every 110 bp, where 98% of the protein coding genes have a ZFN target site (Supplementary Table S5). The number of target sites that are accessible could be greatly expanded through the creation of additional two-finger module archives, where it should be readily feasible to generate a validated set of all 256 possible GNNNNG units allowing virtually any site to be targeted by varying the number of fingers and the spacer between ZFN binding sites.

Although our new archive of modules and our new assembly method increases the density of ZFN target sites, our zinc finger-based systems do not have the flexibility in targeting that has recently been demonstrated with the Transcription Activator-Like Effector Nuclease (TALEN)-based platform (63,66–69). However, ZFNs remain an important platform for targeted genomic editing that may have advantages over TALENs for certain applications, in particular therapeutics. Because each zinc finger recognizes three base pairs as opposed to one base pair for each TALE module (70–73), ZFNs are inherently more compact than TALENs. Thus, for nuclease-based gene therapy applications using viral delivery systems (74), ZFNs constitute a more compact cargo than TALENs, and as such, they may prove to be more amenable to use in certain settings.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–5 and Supplementary Figures 1–8.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the other members of the Wolfe laboratory for insightful comments and discussions. They thank N. Lawson and his laboratory for their insightful advice and zebrafish husbandry training. They thank M. Noyes for his advice for improving the stringency of the B1H selections. They thank J. Zhu for her assistance with the website construction.

FUNDING

U.S. National Institutes of Health (NIH) [GM068110 to S.A.W., HL093766 to N. Lawson and S.A.W.], [HG000249 to G.D.S.]; Breast Cancer Research Foundation and Silvian Foundation (to C.K.). Funding for open access charge: NIH [GM068110].

Conflict of interest statement. None declared.

REFERENCES

1. Kim,Y.G., Cha,J. and Chandrasegaran,S. (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. USA*, **93**, 1156–1160.
2. Urnov,F.D., Rebar,E.J., Holmes,M.C., Zhang,H.S. and Gregory,P.D. (2010) Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.*, **11**, 636–646.
3. Carroll,D. (2011) Genome engineering with zinc-finger nucleases. *Genetics*, **188**, 773–782.
4. Ochiai,H., Sakamoto,N., Fujita,K., Nishikawa,M., Suzuki,K., Matsuura,S., Miyamoto,T., Sakuma,T., Shibata,T. and Yamamoto,T. (2012) Zinc-finger nuclease-mediated targeted insertion of reporter genes for quantitative imaging of gene expression in sea urchin embryos. *Proc. Natl. Acad. Sci. USA*, **109**, 10915–10920.
5. Watanabe,T., Ochiai,H., Sakuma,T., Horch,H.W., Hamaguchi,N., Nakamura,T., Bando,T., Ohuchi,H., Yamamoto,T., Noji,S. *et al.* (2012) Non-transgenic genome modifications in a hemimetabolous insect using zinc-finger and TAL effector nucleases. *Nat. commun.*, **3**, 1017.
6. Young,J.J., Cherone,J.M., Doyon,Y., Ankoudinova,I., Faraji,F.M., Lee,A.H., Ngo,C., Guschin,D.Y., Paschon,D.E., Miller,J.C. *et al.* (2011) Efficient targeted gene disruption in the soma and germ line of the frog Xenopus tropicalis using engineered zinc-finger nucleases. *Proc. Natl. Acad. Sci. USA*, **108**, 7052–7057.
7. Merlin,C., Beaver,L.E., Taylor,O.R., Wolfe,S.A. and Reppert,S.M. (2012) Efficient targeted mutagenesis in the monarch butterfly using zinc finger nucleases. *Genome research*, September 25 (doi:10.1101/gr.145599.112; epub ahead of print).
8. Straimer,J., Lee,M.C., Lee,A.H., Zeitler,B., Williams,A.E., Pearl,J.R., Zhang,L., Rebar,E.J., Gregory,P.D., Llinas,M. *et al.* (2012) Site-specific genome editing in Plasmodium falciparum using engineered zinc-finger nucleases. *Nat. Methods*, **9**, 993–998.

Downloaded from http://nar.oxfordjournals.org/ at University of Massachusetts Medical School on February 14, 2013

9. Perez,E.E., Wang,J., Miller,J.C., Jouvenot,Y., Kim,K.A., Liu,O., Wang,N., Lee,G., Bartsevich,V.V., Lee,Y.L. *et al.* (2008) Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.*, **26**, 808–816.
10. Yusa,K., Rashid,S.T., Strick-Marchand,H., Varela,I., Liu,P.Q., Paschon,D.E., Miranda,E., Ordonez,A., Hannan,N.R., Rouhani,F.J. *et al.* (2011) Targeted gene correction of alpha1-antitrypsin deficiency in induced pluripotent stem cells. *Nature*, **478**, 391–394.
11. Sebastiano,V., Maeder,M.L., Angstman,J.F., Haddad,B., Khayter,C., Yeo,D.T., Goodwin,M.J., Hawkins,J.S., Ramirez,C.L., Batista,L.F. *et al.* (2011) In situ genetic correction of the sickle cell anemia mutation in human induced pluripotent stem cells using engineered zinc finger nucleases. *Stem Cells*, **29**, 1717–1726.
12. Li,H., Haurigot,V., Doyon,Y., Li,T., Wong,S.Y., Bhagwat,A.S., Malani,N., Anguela,X.M., Sharma,R., Ivanciu,L. *et al.* (2011) In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature*, **475**, 217–221.
13. Urnov,F.D., Miller,J.C., Lee,Y.L., Beausejour,C.M., Rock,J.M., Augustus,S., Jamieson,A.C., Porteus,M.H., Gregory,P.D. and Holmes,M.C. (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*, **435**, 646–651.
14. Handel,E.M., Gellhaus,K., Khan,K., Bednarski,C., Cornu,T.I., Muller–Lerch,F., Kotin,R.M., Heilbronn,R. and Cathomen,T. (2012) Versatile and efficient genome editing in human cells by combining zinc-finger nucleases with adeno-associated viral vectors. *Hum. Gene Ther.*, **23**, 321–329.
15. Connelly,J.P., Barker,J.C., Pruett-Miller,S. and Porteus,M.H. (2010) Gene correction by homologous recombination with zinc finger nucleases in primary cells from a mouse model of a generic recessive genetic disease. *Mol. Ther.*, **18**, 1103–1110.
16. Cannon,P. and June,C. (2011) Chemokine receptor 5 knockout strategies. *Current Opin. HIV AIDS*, **6**, 74–79.
17. Greisman,H.A. and Pabo,C.O. (1997) A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites. *Science*, **275**, 657–661.
18. Rebar,E.J. and Pabo,C.O. (1994) Zinc finger phage: affinity selection of fingers with new DNA-binding specificities. *Science*, **263**, 671–673.
19. Isalan,M., Klug,A. and Choo,Y. (2001) A rapid, generally applicable method to engineer zinc fingers illustrated by targeting the HIV-1 promoter. *Nature Biotechnol.*, **19**, 656–660.
20. Segal,D.J., Dreier,B., Beerli,R.R. and Barbas,C.F. (1999) 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.
21. Sepp,A. and Choo,Y. (2005) Cell-free Selection of Zinc Finger DNA-binding proteins using in vitro compartmentalization. *J. Mol. Biol.*, **354**, 212–219.
22. Herrmann,F., Garriga-Canut,M., Baumstark,R., Fajardo-Sanchez,E., Cotterell,J., Minoche,A., Himmelbauer,H. and Isalan,M. (2011) p53 Gene repair with zinc finger nucleases optimised by yeast 1-hybrid and validated by Solexa sequencing. *PloS One*, **6**, e20913.
23. Jamieson,A.C., Kim,S.H. and Wells,J.A. (1994) In vitro selection of zinc fingers with altered DNA-binding specificity. *Biochemistry*, **33**, 5689–5695.
24. Meng,X.D., Noyes,M.B., Zhu,L.H.J., Lawson,N.D. and Wolfe,S.A. (2008) Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nat. Biotechnol.*, **26**, 695–701.
25. Maeder,M.L., Thibodeau-Beganny,S., Osiak,A., Wright,D.A., Anthony,R.M., Eichtinger,M., Jiang,T., Foley,J.E., Winfrey,R.J., Townsend,J.A. *et al.* (2008) Rapid Open-source engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol. Cell*, **31**, 294–301.
26. Durai,S., Bosley,A., Abulencia,A.B., Chandrasegaran,S. and Ostermeier,M. (2006) A bacterial one-hybrid selection system for interrogating zinc finger-DNA interactions. *Comb. Chem. High Throughput Screen.*, **9**, 301–311.
27. Pavletich,N.P. and Pabo,C.O. (1991) Zinc Finger DNA Recognition—Crystal-Structure of a Zif268-DNA Complex at 2.1-A. *Science*, **252**, 809–817.
28. Desjarlais,J.R. and Berg,J.M. (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.
29. Choo,Y. and Klug,A. (1997) Physical basis of a protein-DNA recognition code. *Curr. Opin. Struct. Biol.*, **7**, 117–125.
30. Elrod-Erickson,M., Rould,M.A., Nekludova,L. and Pabo,C.O. (1996) Zif268 protein–DNA complex refined at 1.6å: a model system for understanding zinc finger–DNA interactions. *Structure*, **4**, 1171–1180.
31. Segal,D.J., Beerli,R.R., Blancafort,P., Dreier,B., Effertz,K., Huber,A., Koksch,B., Lund,C.V., Magnenat,L., Valente,D. *et al.* (2003) Evaluation of a modular strategy for the construction of novel polydactyl zinc finger DNA-binding proteins. *Biochemistry*, **42**, 2137–2148.
32. Carroll,D., Morton,J.J., Beumer,K.J. and Segal,D.J. (2006) Design, construction and in vitro testing of zinc finger nucleases. *Nat. Protocols*, **1**, 1329–1341.
33. Liu,Q., Xia,Z. and Case,C.C. (2002) Validated zinc finger protein designs for all 16 GNN DNA triplet targets. *J. Biol. Chem.*, **277**, 3850–3856.
34. Bae,K.H., Kwon,Y.D., Shin,H.C., Hwang,M.S., Ryu,E.H., Park,K.S., Yang,H.Y., Lee,D.K., Lee,Y., Park,J. *et al.* (2003) Human zinc fingers as building blocks in the construction of artificial transcription factors. *Nat. Biotechnol.*, **21**, 275–280.
35. Blancafort,P., Segal,D.J. and Barbas,C.F. (2004) Designing transcription factor architectures for drug discovery. *Mol. Pharmacol.*, **66**, 1361–1371.
36. Bibikova,M., Beumer,K., Trautman,J.K. and Carroll,D. (2003) enhancing gene targeting with designed zinc finger nucleases. *Science*, **300**, 764.
37. Bibikova,M., Carroll,D., Segal,D.J., Trautman,J.K., Smith,J., Kim,Y.-G. and Chandrasegaran,S. (2001) Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol. Cell. Biol.*, **21**, 289–297.
38. Porteus,M.H. and Baltimore,D. (2003) Chimeric nucleases stimulate gene targeting in human cells. *Science*, **300**, 763.
39. Kim,H.J., Lee,H.J., Kim,H., Cho,S.W. and Kim,J.S. (2009) Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Res.*, **19**, 1279–1288.
40. Ramirez,C.L., Foley,J.E., Wright,D.A., Muller-Lerch,F., Rahman,S.H., Cornu,T.I., Winfrey,R.J., Sander,J.D., Fu,F., Townsend,J.A. *et al.* (2008) Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat. Methods*, **5**, 374–375.
41. Zhu,C., Smith,T., McNulty,J., Rayla,A.L., Lakshmanan,A., Siekmann,A.F., Buffardi,M., Meng,X., Shin,J., Padmanabhan,A. *et al.* (2011) Evaluation and application of modularly assembled zinc-finger nucleases in zebrafish. *Development*, **138**, 4555–4564.
42. Shimizu,Y., Sollu,C., Meckler,J.F., Adriaenssens,A., Zykovich,A., Cathomen,T. and Segal,D.J. (2011) Adding fingers to an engineered zinc finger nuclease can reduce activity. *Biochemistry*, **50**, 5033–5041.
43. Mandell,J.G. and Barbas,C.F. 3rd. (2006) Zinc finger tools: custom DNA-binding domains for transcription factors and nucleases. *Nucleic Acids Res.*, **34**, W516–W523.
44. Sander,J.D., Zaback,P., Joung,J.K., Voytas,D.F. and Dobbs,D. (2009) An affinity-based scoring scheme for predicting DNA-binding activities of modularly assembled zinc-finger proteins. *Nucleic Acids Res.*, **37**, 506–515.
45. Wolfe,S.A., Greisman,H.A., Ramm,E.I. and Pabo,C.O. (1999) Analysis of zinc fingers optimized via phage display: evaluating the utility of a recognition code. *J. Mol. Biol.*, **285**, 1917–1934.
46. Isalan,M., Choo,Y. and Klug,A. (1997) Synergy between adjacent zinc fingers in sequence-specific DNA recognition. *Proc. Natl. Acad. Sci. USA*, **94**, 5617–5621.
47. Dreier,B., Beerli,R.R., Segal,D.J., Flippin,J.D. and Barbas,C.F. 3rd. (2001) 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.
48. Wolfe,S.A., Grant,R.A., Elrod-Erickson,M. and Pabo,C.O. (2001) Beyond the "recognition code": structures of two Cys2His2 zinc finger/TATA box complexes. *Structure*, **9**, 717–723.

Downloaded from http://nar.oxfordjournals.org/ at University of Massachusetts Medical School on February 14, 2013

49. Doyon,Y., McCammon,J.M., Miller,J.C., Faraji,F., Ngo,C., Katibah,G.E., Amora,R., Hocking,T.D., Zhang,L., Rebar,E.J. *et al.* (2008) Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nat. Biotechnol.*, **26**, 702–708.
50. Gupta,A., Christensen,R.G., Rayla,A.L., Lakshmanan,A., Stormo,G.D. and Wolfe,S.A. (2012) An optimized two-finger archive for ZFN-mediated gene targeting. *Nat. Methods*, **9**, 588–590.
51. Kim,S., Lee,M.J., Kim,H., Kang,M. and Kim,J.S. (2011) Preassembled zinc-finger arrays for rapid construction of ZFNs. *Nat. Methods*, **8**, 7.
52. Sander,J.D., Dahlborg,E.J., Goodwin,M.J., Cade,L., Zhang,F., Cifuentes,D., Curtin,S.J., Blackburn,J.S., Thibodeau-Beganny,S., Qi,Y. *et al.* (2011) Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nat. Methods*, **8**, 67–69.
53. Cheng,C. and Young,E.T. (1995) A single amino acid substitution in zinc finger 2 of Adr1p changes its binding specificity at two positions in UAS1. *J. Mol. Biol.*, **251**, 1–8.
54. Miller,J.C. and Pabo,C.O. (2001) Rearrangement of side-chains in a zif268 mutant highlights the complexities of zinc finger-DNA recognition. *J. Mol. Biol.*, **313**, 309–315.
55. Westerfield,M. (1993) *The Zebrafish Book*. University of Oregon Press, Eugene, OR.
56. Christensen,R.G., Gupta,A., Zuo,Z., Schriefer,L.A., Wolfe,S.A. and Stormo,G.D. (2011) A modified bacterial one-hybrid system yields improved quantitative models of transcription factor specificity. *Nucleic Acids Res.*, **39**, e83.
57. Noyes,M.B., Meng,X., Wakabayashi,A., Sinha,S., Brodsky,M.H. and Wolfe,S.A. (2008) A systematic characterization of factors that regulate Drosophila segmentation via a bacterial one-hybrid system. *Nucleic Acids Res.*, **36**, 2547–2560.
58. Gupta,A., Meng,X., Zhu,L.J., Lawson,N.D. and Wolfe,S.A. (2011) Zinc finger protein-dependent and -independent contributions to the in vivo off-target activity of zinc finger nucleases. *Nucleic Acids Res.*, **39**, 381–392.
59. Bailey,T.L., Boden,M., Buske,F.A., Frith,M., Grant,C.E., Clementi,L., Ren,J., Li,W.W. and Noble,W.S. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.*, **37**, W202–W208.
60. Bailey,T.L., Williams,N., Misleh,C. and Li,W.W. (2006) MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.*, **34**, W369–W373.
61. Miller,J.C. (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. *Nature Biotechnol.*, **25**, 778–785.
62. Szczepek,M., Brondani,V., Buchel,J., Serrano,L., Segal,D.J. and Cathomen,T. (2007) Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nature Biotechnol.*, **25**, 786–793.
63. Reyon,D., Tsai,S.Q., Khayter,C., Foden,J.A., Sander,J.D. and Joung,J.K. (2012) FLASH assembly of TALENs for high-throughput genome editing. *Nat. Biotechnol.*, **30**, 460–465.
64. Guschin,D.Y., Waite,A.J., Katibah,G.E., Miller,J.C., Holmes,M.C. and Rebar,E.J. (2010) *Engineered Zinc Finger Proteins*, **649**, pp. 247–256.
65. Soldner,F., Laganiere,J., Cheng,A.W., Hockemeyer,D., Gao,Q., Alagappan,R., Khurana,V., Golbe,L.I., Myers,R.H., Lindquist,S. *et al.* (2011) Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell*, **146**, 318–331.
66. Christian,M., Cermak,T., Doyle,E.L., Schmidt,C., Zhang,F., Hummel,A., Bogdanove,A.J. and Voytas,D.F. (2010) Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*, **186**, 757–761.
67. Miller,J.C., Tan,S., Qiao,G., Barlow,K.A., Wang,J., Xia,D.F., Meng,X., Paschon,D.E., Leung,E., Hinkley,S.J. *et al.* (2011) A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.*, **29**, 143–148.
68. Bedell,V.M., Wang,Y., Campbell,J.M., Poshusta,T.L., Starker,C.G., Krug Ii,R.G., Tan,W., Penheiter,S.G., Ma,A.C. *et al.* (2012) In vivo genome editing using a high-efficiency TALEN system. *Nature.*, **491**, 114–118.
69. Mussolino,C., Morbitzer,R., Lutge,F., Dannemann,N., Lahaye,T. and Cathomen,T. (2011) A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.*, **39**, 9283–9293.
70. Boch,J., Scholze,H., Schornack,S., Landgraf,A., Hahn,S., Kay,S., Lahaye,T., Nickstadt,A. and Bonas,U. (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science*, **326**, 1509–1512.
71. Moscou,M.J. and Bogdanove,A.J. (2009) A simple cipher governs DNA recognition by TAL effectors. *Science*, **326**, 1501.
72. Mak,A.N., Bradley,P., Cernadas,R.A., Bogdanove,A.J. and Stoddard,B.L. (2012) The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science*, **335**, 716–719.
73. Deng,D., Yan,C., Pan,X., Mahfouz,M., Wang,J., Zhu,J.K., Shi,Y. and Yan,N. (2012) Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science*, **335**, 720–723.
74. Atkinson,H. and Chalmers,R. (2010) Delivering the goods: viral and non-viral gene therapy systems and the inherent limits on cargo DNA and internal sequences. *Genetica*, **138**, 485–498.

Downloaded from http://nar.oxfordjournals.org/ at University of Massachusetts Medical School on February 14, 2013