Directed Evolution of an Enhanced and Highly Efficient FokI Cleavage Domain for Zinc Finger Nucleases

Jing Guo, Thomas Gaj and Carlos F. Barbas III *

The Skaggs Institute for Chemical Biology and the Departments of Molecular Biology and Chemistry, The Scripps Research Institute, La Jolla, CA 92037, USA

Zinc finger nucleases (ZFNs) are powerful tools for gene therapy and genetic engineering. The high specificity and affinity of these chimeric enzymes are based on custom-designed zinc finger proteins (ZFPs). To improve the performance of existing ZFN technology, we developed an *in vivo* evolution-based approach to improve the efficacy of the FokI cleavage domain (FCD). After multiple rounds of cycling mutagenesis and DNA shuffling, a more efficient nuclease variant (*Sharkey*) was generated. *In vivo* analyses indicated that *Sharkey* is >15-fold more active than wild-type FCD on a diverse panel of cleavage sites. Further, a mammalian cell-based assay showed a three to sixfold improvement in targeted mutagenesis for ZFNs containing derivatives of the *Sharkey* cleavage domain. We also identified mutations that impart sequence specificity to the FCD that might be utilized in future studies to further refine ZFNs through cooperative specificity. In addition, *Sharkey* was observed to enhance the cleavage profiles of previously published and newly selected heterodimer ZFN architectures. This enhanced and highly efficient cleavage domain will aid in a variety of ZFN applications in medicine and biology.

Edited by J. Karn

Introduction

Zinc finger nucleases are artificial restriction enzymes generated by fusing the nonspecific DNA cleavage domain of the endonuclease FokI with site-specific DNA binding zinc finger proteins (ZFPs).[1] These chimeric enzymes have emerged as powerful tools for genome editing. Because of the flexible nature of ZFPs, ZFNs can be assembled that induce double-strand breaks (DSBs) site-specifically into genomic DNA. Targeted genes can be disrupted via mutagenic nonhomologous end joining (NHEJ)[2–5] or modified via homologous recombination (HR) if a closely related DNA template is supplied.[6–9] This method has been applied in many organisms, including plants,[7] *Drosophila*,[2,8] *Caenorhabditis elegans*,[9] zebrafish,[3,4] rats,[10] and various mammalian cells.[5,6,11] ZFNs can also be used in basic molecular research, providing diverse options for molecular cloning.[12]

The modular structure of C_2H_2 ZF motifs and their independent recognition capabilities make them an ideal framework for developing custom ZFPs with novel sequence specificities.[13] Each motif recognizes 3 or 4 bp via its α-helix, and engineering multiple ZFs in tandem enables the recognition of extended DNA sequences with unparalleled specificity.[13–15] Multiple methods have been developed toward the goal of generating highly selective ZFPs. In the modular-assembly approach, a strategy that has proven to be very effective for the generation of ZF transcription factors, individual ZF modules with preselected specificities are assembled using standard recombinant DNA technology.[15,16] Modular assembly is rapid, but constrained by sequence limitations, because there are not ZF motifs that recognize each of the 64 DNA triplets.[16–24] Alternatively, cell-based selection methods such as oligomerized pool engineering (OPEN) have been shown to be effective but require the construction and interrogation of large libraries and have similar sequence constraints.[25,26] When ZFPs are coupled with the nonspecific FokI cleavage domain (FCD),

*Corresponding author. E-mail address:
carlos@scripps.edu.

their affinity and specificity are major determinants of the activity and toxicity of the resulting ZFN.[6,27]

Unlike ZF transcription factors, which are usually composed of high-affinity six-finger ZFPs and can be designed to recognize a single site within the human genome,[14,15] ZFNs are typically composed of lower-affinity three- or four-finger ZFPs. To a certain extent, this is attributable to the nature of ZFN target sites. ZFN target sites are composed of two ZFP binding sites positioned in a tail-to-tail orientation and separated by a 5- to 7-bp spacer sequence,[28] and while theoretically every sequence can be targeted by custom ZFPs, in practice, not all can be targeted efficiently. Many factors affect the efficiency and precision of a chimeric nuclease, including the specificity and affinity of ZFPs,[6,27] the length and composition of the interdomain linker,[29,30] the length of spacer sequence DNA in between ZFP binding sites,[29,30] and the interaction between two FCDs.[31,32] Toward the goal of improving the performance ZFNs in mammalian cells, we sought to enhance the catalytic capabilities of the FCD through laboratory-directed protein evolution. We developed an *in vivo* cell-survival and evolution strategy to identify an FCD variant called *Sharkey*, which is >15-fold more active than wild-type FCD

on an array of diverse cleavage sites. When coupled with ZFPs, *Sharkey* derivatives stimulated a 3- to 6-fold increase in mutagenesis in mammalian cells compared to ZFNs constructed with wild-type FCD. In addition, we demonstrate that the *Sharkey* framework increases the rate of stimulated mutagenesis for both existing and newly selected heterodimeric ZFN architectures. We believe this novel FCD variant will serve as a useful reagent in a wide variety of ZFN applications.

Results

Selection strategy

To selectively enrich for catalytically improved FCD variants, we have based our design on a two-plasmid system developed by Chen and Zhao[33] for directing activity of the homing endonuclease I–SceI to a novel sequence. This system links DNA cleavage events with cell survival. The reporter plasmid contains the toxic gene *ccdB*[34] under tight control of the ara*BAD* promoter. Downstream of the *ccdB* gene, we engineered one copy of the desired ZFN cleavage site (Fig. 1a). A low-copy ZFN expression plasmid
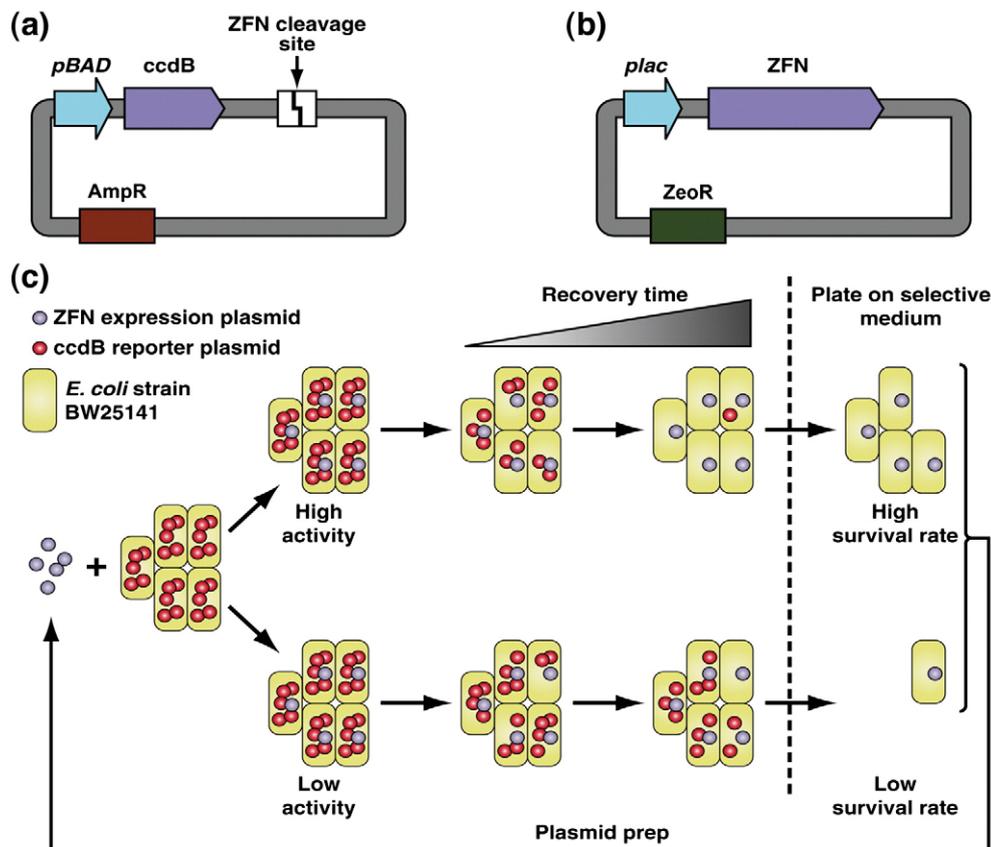

Fig. 1. Schematic representation of the selection strategy used for isolating novel FCD variants. (a and b) A two-plasmid approach using a reporter consisting of a single ZFN cleavage site downstream of *ccdB* and a ZFN expression plasmid under tight control of a modified *lac* promoter can be used to selectively enrich for catalytically improved FCDs. (c) A library of FCD variants can be transformed into the *ccdB* harboring BW25141 selection strain and enriched following ZFN-mediated reporter plasmid cleavage/degradation. Decreasing the recovery time following transformation facilitates the isolation of ZFN variants with enhanced catalytic properties.

encoding the ZFN gene under the control of a modified *lac* promoter, which possesses an additional *lac* operator sequence (Fig. 1b) for tighter control of ZFN expression, was constructed. Cleavage of reporter plasmid by a ZFN mutant linearizes the plasmid and results in degradation by RecBCD nuclease.[35] The process includes three steps: (1) expression of endonuclease mutants; (2) cleavage of reporter plasmids by these mutants; and (3) degradation of linearized plasmids. Time needed for step 3 is constant, but that for step 2 is variable. The DNA cleavage rate of a mutant will influence the rate of step 2 and, therefore, the time of linearization of the toxic gene. As a consequence, the survival rate (SR) of a mutant, calculated as the ratio of the number of colonies on an arabinose selection plate to that on a nonselective plate, is positively correlated with its catalytic activity. Mutants with higher catalytic activity linearize all reporter plasmids in an *Escherichia coli* cell within a shorter time window and will be enriched during evolution (Fig. 1c).

Directed evolution of an enhanced FCD

To establish proper selective pressure in our system, we first measured the SR of wild-type ZFN. An expression plasmid harboring wild-type FCD fused to ZF domain P3 (P3.wt) (Supplementary Fig. 1) was transformed into the p11LacY-sP3/P3 harboring strain BW25141.[33] The sequence of the ZFN site in the reporter plasmid p11LacY-sP3/P3 is shown in Table 1. The spacer sequence of the ZFN cleavage site is the restriction site MluI. Aliquots of transformants were withdrawn at regular time intervals after transformation and grown on solid media with and without arabinose to calculate SR. Without IPTG induction, the resulting SR curve was linear and ~10% survival was observed after 1 h of recovery (data not shown). With IPTG induction, however, the reaction was finished within 1 h and survival was 80–100%. We therefore used IPTG induction only for the first round of selection.

Error-prone PCR was used to introduce diversity into the catalytic domain of FokI prior to fusion to the P3 ZF domain. The mutation rate was determined to be ~4 amino acids per catalytic domain. In addition to cycling mutagenesis, DNA shuffling[36] was applied to the full ZFN gene to combine beneficial mutations every three rounds of selection. The ZFN library was cloned into pPDAZ and electroporated into *E. coli*. ZFN libraries were routinely composed of 10^7–10^8 transformants. Following transformation, ZFN containing plasmids were harvested from overnight culture and electroporated into the selection strain containing the reporter plasmid p11LacY-sP3/P3. As the evolution progressed, enrichment was observed (Fig. 2a). We measured SR at rounds 3, 6, and 9 (Fig. 2b). Relative to wild-type FokI, rounds 6 and 9 showed much higher SR at all measured time intervals. A more direct evaluation of ZFN activity was performed *in vitro*. Supercoiled plasmid pSub-P3, which contained a single copy of ZFN cleavage site, was incubated with cell extracts from rounds 3, 6, and 9 at room temperature. Round 9 cell extracts displayed the highest activity, linearizing all substrate within 10 min. In comparison, wild-type FokI linearized <10% of its substrate (Fig 2c).

Because round 9 displayed an SR of >50% at 1 h, higher selection stringency was desired for further evolution. One way of increasing stringency is to decrease the level of protein expression. The initiation codon of a transcript has a direct impact on its translation efficiency. Changing the start codon from the most frequently used ATG to GTG can reduce protein translation levels by fivefold.[37] We tested this strategy on clone FCD_{R9-3} and found that a change in the start codon reduced the SR of this variant from ~80% to ~8% (data not shown). A secondary library with a GTG initiation codon was then constructed from the 9th round via error prone PCR. The transformed library was composed of ~10^8 transformants. IPTG was added for the 10th round of selection. Recovery time was initially set at 2.5 h and shortened by 0.5 h after every three rounds of selection. Nine more rounds of evolution were executed in the absence of IPTG induction. Following every three rounds of selection, beneficial mutations were propagated by DNA shuffling. SR was measured following each round of selection (Fig. 2d). The selection was concluded when SR was measured to be >30%.

Table 1. ZFP binding characteristics and ZFN target sites

ZFP	K_d (nM)	ZFP binding site
P3	—	GCA GTG GCG
E3	35	GGG GCC GGA
E4	10.6±3.3	GGG GCC GGA GCC
E5	4.2±1.9	GGG GCC GGA GCC GCA
E6 (E2C)	0.85±0.2	GGG GCC GGA GCC GCA GTG

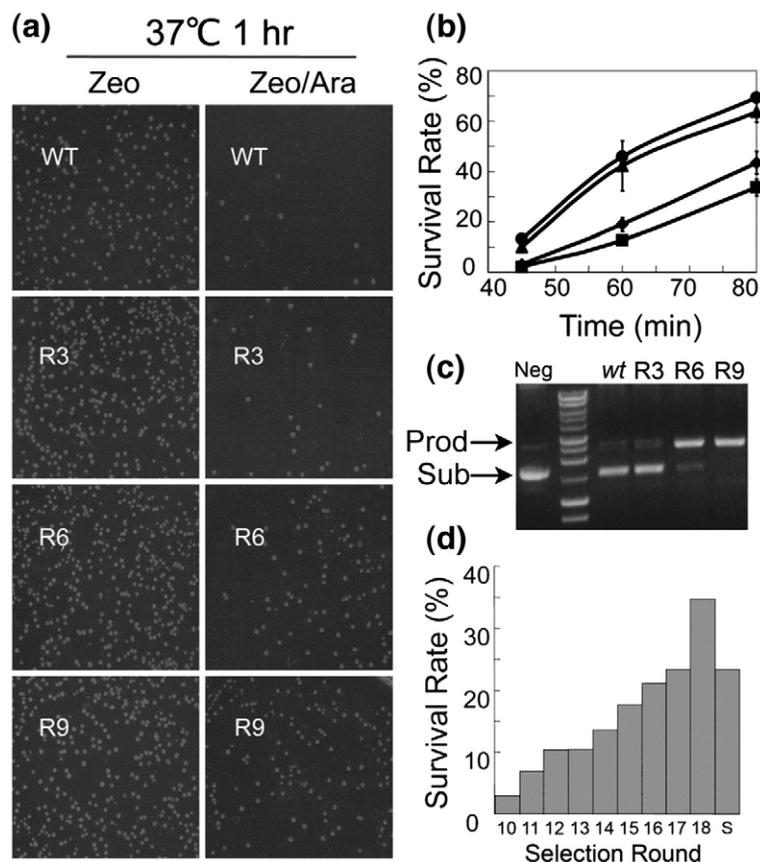| Cleavage site	DNA sequence (6-bp spacer)
E3/E3	TCC GGC CCC (ACGCGT) GGG GCC GGA
P3/P3	CGC CAC TGC (ACGCGT) GCA GTG GCG
P3/E6	CGC CAC TGC (ACGCGT) GGG GCC GGA GCC GCA GTG
E4/E4	GGC TCC GGC CCC (ACGCGT) GGG GCC GGA GCC
E5/E5	TGC GGC TCC GGC CCC (ACGCGT) GGG GCC GGA GCC GCA
E6/E6	CAC TGC GGC TCC GGC CCC (ACGCGT) GGG GCC GGA GCC GCA GTG

Fig. 2. Enhancing the FCD by directed evolution. A library of ZFNs was transformed into selection strain BW25141 and subjected to multiple rounds of evolution. (a) Survival rate (SR), which correlates directly with catalytic activity, was measured at 1 h for rounds 3, 6, and 9. (b) SR was observed to increase with recovery time. SR curves were measured for wt (♦), R3 (■), R6 (▲) and R9 (●). (c) The extent of substrate linearization from rounds 3, 6, and 9 was measured from cellular extracts prepared from overnight cultures. "Sub" indicates supercoiled substrate plasmid pSub-P3. "Prod" indicates linearized substrate plasmid pSub-P3. (d) SR for selection rounds 11–18 measured at 1 h. Bar 10 indicates FCD_{R18-28}.

Analysis of selected FCD variants

During sequence analysis of the final round, we noticed that a large number of clones contained mutations in the ZF domain and/or the interdomain linker. Those mutations were primarily in two areas: in the second linker of the ZF motif and in the four-amino acid interdomain linker. These two areas are removed from the catalytic center and mutations in these regions are likely to have altered the affinity and specificity of the ZFN rather than the catalytic domain. For a more direct comparison of FCD mutants, we reamplified the FCDs from the 18th round of selection, cloned them in the original framework, and screened 50 FCD variants for optimal performance *in vivo*. One of the most active catalytic domains, FCD_{R18-28} (S418P, F432L, K441E, Q481H, H523Y, N527D, and K559Q) was selected for further characterization. With a GTG start codon, this mutant had a 25% SR at 1 h, slightly lower than the overall 33% SR of the 18th round, but significantly higher than the 8% SR of the best clone from the 9th round. In comparison, wild-type FokI had an SR of <1% under the same conditions (data not shown).

To determine whether the presence of the FCD_{R18-28} domain enhanced the rate of DNA cleavage, we purified P3 nucleases with either the wild-type FCD or FCD_{R18-28} and performed an *in vitro* DNA cleavage analysis. Rates of DNA cleavage were determined using a constant concentration of ZFN

(12 nM) and increasing substrate concentrations (4, 6, 12, 24, or 36 nM). Linearized plasmid pSub-P3 DNA was used as a substrate. Cleavage of this substrate generates two product DNA molecules of the same size, simplifying the analysis. The progress of each reaction was monitored over time by measuring the initial velocity (Fig. 3a). The rate of cleavage for FCD_{R18-28} was four- to fivefold higher than that of wild-type FokI, demonstrating that FCD_{R18-28} has enhanced catalytic activity relative to wild-type ZFN. FCD_{R18-28} was also observed to have a faster turnover rate than wild-type FCD (Fig. 3b).

Having established that FCD_{R18-28} exhibits enhanced DNA cleavage and turnover rates *in vitro*, we next sought to identify the requisite amino acid mutations required for increasing the performance of FCD. We sequenced individual FCD variants from the 18th round of selection and identified four mutations (S418P, K441E, Q481H, and N527D) present in ≥70% of active mutants, including FCD_{R18-28} (Supplementary Fig. 2). Site-directed mutagenesis was performed to generate FCD variants containing these point mutations and SR was measured against ZFN target sites containing MluI (ACGCGT) and Nx6 (N = A, T, C or G) spacer sequences (Fig. 3c). Relative to wild-type FokI, FCD variants containing the single point mutations S418P and K441E exhibited increased rates of cleavage for each spacer sequence. Interestingly, an FCD variant containing the single point mutation Q481H exhibited modest activity on the MluI core cleavage site
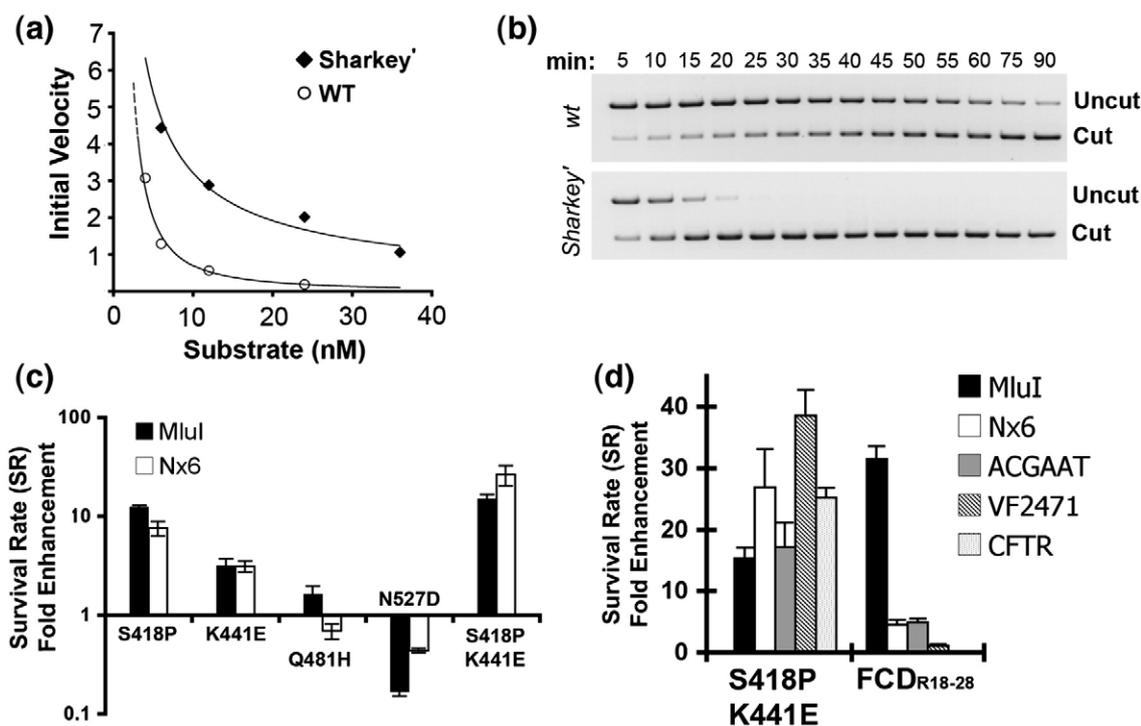
Fig. 3. Selected ZFN variants have enhanced catalytic profiles as demonstrated by *in vitro* DNA cleavage assays and *in vivo* activity assays. (a) *In vitro* cleavage of target DNA by P3.nuclease with either wild-type FCD (white) or FCD$_{R18-28}$ (black). Cleavage rates were determined by measuring the initial velocity of DNA cleavage. ZFN (12 nM) was added to prewarmed ZFN reaction buffer containing 4, 6, 12, 24, or 36 nM pSub-P3 substrate DNA. Reaction conditions were standardized. (b) *In vitro* cleavage of target DNA by P3 nuclease with either FCD$_{R18-28}$ or wild-type FCD. "Uncut" indicates linearized substrate plasmid pSub-P3. "Cut" indicates cleavage products. Cleavage was monitored incrementally over 90 min. (c) Activity analysis of FCD variants containing the selected mutations S418P, K441E, Q481H, N527D, and S418P::K441E with the P3 ZF domain. ZFN activity was measured against MluI (ACGGCT) and Nx6 (N = A, T, C, or G) spacer sequences and normalized to wild-type FCD. Error bars indicate standard deviation of three replicates. (d) Activity analysis of FCD variants S418P::K441E (*Sharkey*) and FCD$_{R18-28}$ (*Sharkey'*) with the P3 ZF domain. ZFN activity was measured against MluI, Nx6, ACGAAT, VF2471 (GAGAGT), and CFTR (TGGTGA) spacer sequences and normalized to wild-type FCD. Error bars indicate standard deviation of three replicates.

and decreased activity on the Nx6 spacer sequence. N527D FCD exhibited considerably reduced levels of activity on both MluI and Nx6 core sequences. An FCD variant composed of the beneficial mutations S418P and K441E (FCD$_{418P.441E}$) cleaved MluI and Nx6 containing ZFN target sites >15-fold more efficiently than did wild-type FCD (Fig. 3c). To probe the generalized cleavage capabilities of FCD$_{R18-28}$ relative to those of FCD$_{418P.441E}$ and wild-type FokI, we measured SR against a diverse array of spacer sequences. In addition to MluI and Nx6, SR was measured against ZFN target sites containing the endogenous *Drosophila* spacer sequence ACGAAT, the human *VEGF-A* spacer sequence GCGAGT (VF2471), and the human *CFTR* spacer sequence TGGTGA (CFTR).[25] Unexpectedly, FCD$_{R18-28}$ exhibited substantially increased levels of cleavage exclusively on MluI, displaying only a modest gain in activity against the remaining panel of spacer sequences (Fig. 3d). However, relative to wild-type FCD, FCD$_{418P.441E}$ exhibited >15-fold improvement in activity against each endogenous cleavage site (Fig. 3d). We anticipate the enhanced cleavage capabilities of FCD$_{418P.441E}$ (hereinafter referred to

as *Sharkey*) will have great utility in a wide variety of ZFN-based applications.

The complete nucleic acid and amino acid sequence of *Sharkey* is provided in Supplementary Information (Supplementary Fig. 3).

Sharkey enhances the efficiency of mutagenesis in mammalian cells

Because one major application of ZFNs is in gene therapy, we sought to evaluate the catalytic activity of *Sharkey* in a mammalian model system. In mammalian cells, the majority of DSBs are repaired by NHEJ, a somewhat error-prone process, resulting in small deletions and insertions at the site of DSBs.[38] It is expected that an increased frequency of DSBs at a given site will lead to an increased rate of mutagenesis. We constructed an enhanced green fluorescent protein (EGFP)-based reporter system to rapidly gauge the potential for ZFNs to create site-specific DSBs. The recognition site for E2C nuclease (Table 1) was inserted between amino acids 157 and 158 of the gene encoding EGFP and subsequently disabled with a frameshift. We chose to insert the

ZFN recognition site between EGFP residues 157 and 158 because it is a position that has been previously shown to accommodate 20-amino-acid peptide insertions.[39] The resulting nonfunctional transgene was stably integrated at a single location in the genome of human embryonic kidney (HEK) 293 cells using the Flp-In system. Certain deletions (e.g., 2, 5, or 8 bp) or insertions (e.g., 1, 4, or 7 bp) caused by NHEJ-mediated mutagenesis will restore the frame and consequently EGFP function (Fig. 4a). While this assay only reflects a small portion of the total mutation events, it has the advantage of being robust and high throughput with little background. Moreover, the rate of mutagenesis can also be measured by MluI cleavage. Because the 6-bp spacer sequence between ZFP binding sites is an MluI restriction site, any mutations within this spacer will abolish cleavage by MluI and can be easily evaluated by limited-cycle PCR/restriction digest analysis (Fig. 4d).

We constructed a series of ZFPs containing different numbers of ZF motifs. ZFP E2C (E6) is a six-finger protein and recognizes an 18-bp sequence, whereas ZFPs E5, E4, and E3 were made by deleting one, two, or three fingers from the N-termini of the protein, respectively. The affinity of each of the ZFPs for its target was determined by electrophoretic mobility shift assays (Table 1). ZFNs composed of these ZFPs target the same location in the mammalian genome, simplifying the analysis and reducing the potential background interference. We first fused these ZFPs to wild-type FCD, under the control of a cytomegalovirus (CMV) promoter, and compared their abilities to stimulate mutagenesis in mammalian cells by transient expression experiments. On day 3 after transfection, we analyzed these samples by flow cytometry and found that ZFNs with four or five fingers promoted mutagenesis with the highest efficiency ($6.87 \pm 0.8\%$ and $7.03 \pm 0.96\%$ EGFP positive, respectively). In contrast, the activity of E3.wt was barely above background ($0.55 \pm 0.10\%$) and E6.wt ($3.96 \pm 0.63\%$) was about half as active as E4.wt or E5.wt (Fig. 4b and c). The low level of activity of E3.wt was expected because the affinity of E3 is rather modest (35 nM). E6.wt also showed reduced activity even though E6 exhibits appreciable affinity for its target site (0.85 nM). It is possible that too high an affinity may obstruct downstream processes. MluI assays were in agreement with these results (Fig. 4d).

We then substituted the wild-type cleavage domain in these ZFNs with *Sharkey'* (previously FCD$_{R18-28}$), a derivative of the hyperactive *Sharkey* cleavage domain, and performed an identical mutagenesis assay. Following transfection, *Sharkey'*-based ZFNs increased EGFP expression three to sixfold relative to that of wild-type ZFN (Fig. 4b and c). A two to threefold increase was observed by MluI digestion resulting in up to ~64% targeted mutagenesis (Fig. 4d). Additionally, *Sharkey'* enhanced the performance of E6 nucleases. This may be due to the higher turnover rate of *Sharkey'*. To ensure that the improved activity was not achieved

at the cost of increased off-target cleavage, we used a well-established assay to measure genome-wide DNA cleavage levels. Phosphorylated histone H2AX (γH2AX) appears rapidly after DNA damage and can be used as a DSB indicator.[40] Using fluorescein-isothiocyanate-labeled antibody against γH2AX, we quantified the percentage of fixed antibody-stained cells by flow cytometry. We observed no appreciable difference in the levels of γH2AX staining resulting from *Sharkey'* or wild-type ZFN expression (Supplementary Fig. 4).

Sharkey is compatible with heterodimeric FCD architectures

The use of ZFNs can often be associated with considerable cytotoxicity, an undesirable result presumably linked to off-target cleavage events.[11] In recent years, structure-based design strategies have successfully been applied toward the generation of improved ZFN architectures with diminished levels of cleavage-competent homodimers.[31,32] Through iterative rounds of structure-guided mutagenesis and screening, Miller *et al.* were able to identify compatible FCDs with two substitutions in each cleavage domain, E490K:I538K and Q486E: I499L.[31] Relying primarily on *in silico* approaches, Szczepek *et al.* generated compatible FCD variants by remodeling the dimerization interface of the native enzyme, simply exchanging salt-bridge-participating Arg and Asp residues.[32] However, while these ZFN architectures are efficient for endogenous gene-targeting applications, oftentimes fewer mutations are introduced,[5] a result perhaps of reduced binding affinity between specialized cleavage domains.[31] We sought to address this issue and engineer enhanced heterodimeric ZFNs by (1) integrating the *Sharkey* framework within the existing ZFN architectures developed by Miller *et al.* and Szczepek *et al.* and (2) developing novel *Sharkey*-based architectures via directed evolution.

We began the process of evolving a *Sharkey*-based heterodimeric ZFN framework by introducing a ZFN expression cassette containing an E6.*Sharkey'* D483R mutant into the reporter plasmid p11-LacY-sE6/P3. The homodimeric P3/P3 cleavage site within this reporter plasmid was replaced with the heterodimeric E6/P3 cleavage site (Table 1). The *Sharkey* derivative FCD$_{9-3}$ (S418P, K448E, H523Y, N527D, and R570Q) was used as a starting point for our selections. To limit the library size, we focused on introducing diversity exclusively into the FCD$_{9-3}$ dimerization interface, holding E6.*Sharkey'* constant over the course of our evolutions. Randomizing FCD$_{9-3}$ positions 483–487 with all 20 amino acids (theoretical library size, 3.4×10^7 variants), we cloned the P3.FCD$_{9-3}$ library into the ZFN expression plasmid pPDAZ, generating ~10^8 transformants. Active ZFN variants were enriched as described earlier. Following three rounds of selection, we observed Asp as the consensus residue at position 483. We subsequently generated additional diversity within FCD$_{9-3}$ at positions 484–487, this
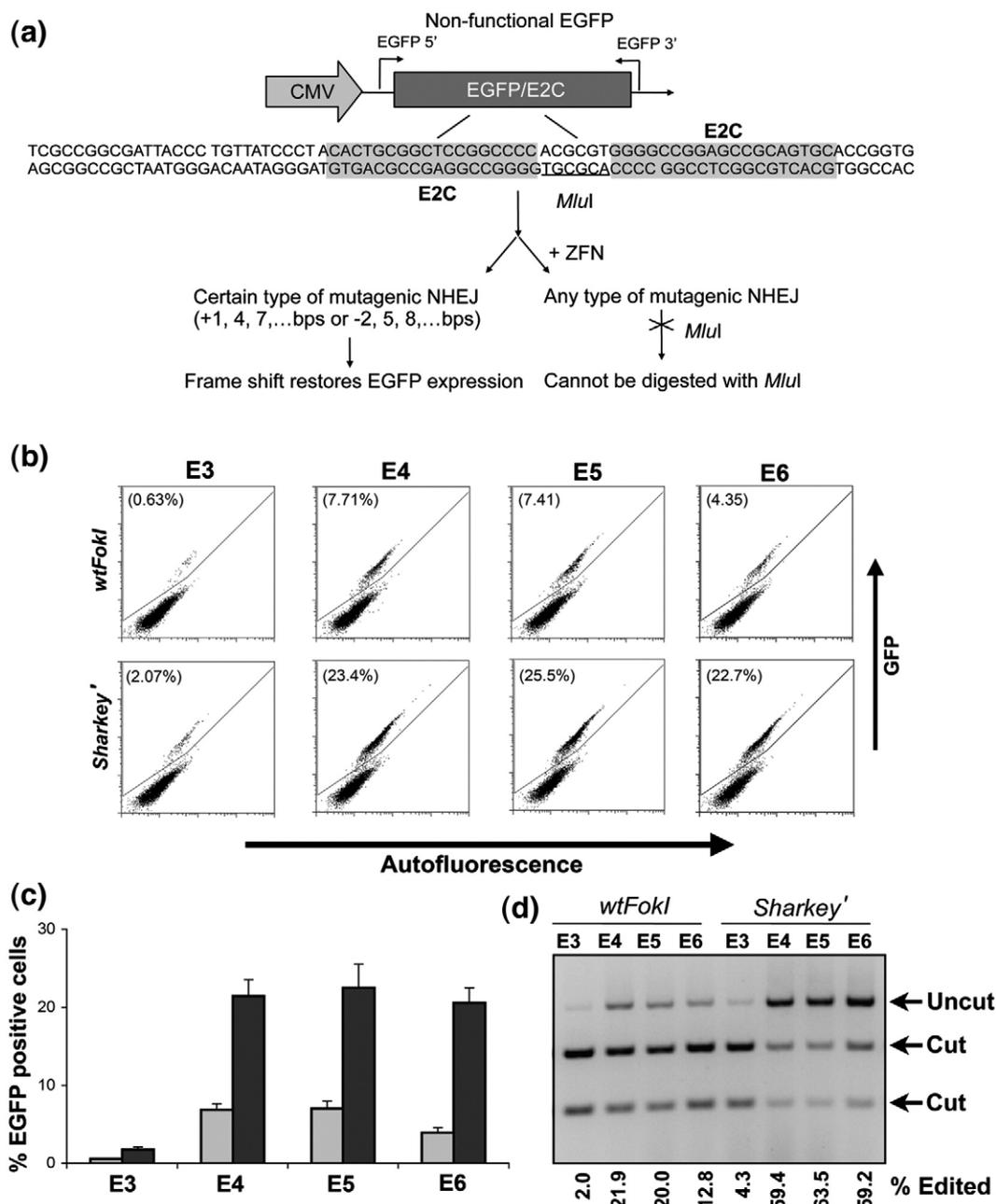
Fig. 4. *Sharkey'* increases the rate of mutagenesis in a mammalian model system. (a) Schematic overview of the reporter system used to evaluate the efficiency of mutagenesis in mammalian cells. The model system consists of a HEK 293 cell line containing a modified and disabled EGFP transgene stably integrated in a single locus. An MluI restriction site flanked by E2C ZF recognition sites was inserted between EGFP residues 157 and 158. Select deletions (e.g., 2, 5, or 8 bp) or insertions (e.g., 1, 4, or 7 bp) result in frame restoration and EGFP expression. (b) Representative flow cytometry data for reporter cells transfected with CMV-controlled wild-type FokI and *Sharkey'* cleavage domains with 3, 4, 5, and 6-finger ZF DNA binding domains. Mutagenesis is measured by counting the percentage of EGFP-positive cells. (c) Quantification of EGFP-positive reporter cells following transfection with ZFN. Error bars denote standard deviation of three replicates (d) MluI restriction digest assay of HEK 293 reporter cells transfected with ZFN. "Cut" indicates the presence of unmodified reporter gene. "Uncut" indicates the presence of ZFN-modified reporter gene. The percentage of modified reporter cells is indicated.

time holding D483 constant. All 20 amino acids were represented at each position (theoretical library size, 1×10^6 variants). Following five rounds of selection, we observed the SR to be ~27% (Supplementary Fig. 5). Sequence analysis of selected mutants led to

the identification of the consensus motif DAMQS (hereinafter referred to as DS) (Supplementary Fig. 1f).

We next sought to integrate *Sharkey* with the existing heterodimeric ZFN architectures. We hypothesized that introducing *Sharkey* mutations into

the FCD variants designed by Miller *et al*. and Szczepek *et al*. would lead to increased rates of stimulated mutagenesis in mammalian cells. Toward the goal of validating our hypothesis, we constructed an EGFP reporter cell line containing the heterodimeric P3/E4 ZFN cleavage site (Table 1). This cell line is identical to the previous reporter line described, with the exception of the P3/E4 cleavage site. *Sharkey*′ mutations were positioned alongside the heterodimeric KK::EL and RR::DD interfaces and fused to either P3 or E4 ZFPs. In addition to these existing architectures, we analyzed the mutagenesis-stimulating characteristics of our selected DS domain with wild-type FCD and *Sharkey*′ residues. Each ZFN was under control of a CMV promoter, and mutagenesis was measured by flow cytometry following transient ZFN expression. The pairwise analysis of these three architectures shows that ZFN heterodimers containing *Sharkey*′ mutations are greater than sixfold more efficient at stimulating mutagenesis than wild-type FokI-based KK::EL and RR:DD ZFNs (Fig. 5a and b). In addition, *Sharkey*′-based ZFN heterodimers composed of our selected DS interface are greater than twofold more efficient at stimulating mutagenesis than the RR::DD heterodimers described by Szczepek *et al.* (Fig 5a). None of the characterized variants were observed to be cleavage-competent homodimers. These results demonstrate that *Sharkey* muta-

tions are compatible with existing heterodimer ZFN architectures and may further increase the rate of stimulated mutagenesis.

Discussion

We have improved the efficiency of ZFNs through the use of an evolutionary strategy geared toward optimizing the catalytic activity of the FCD. Numerous factors affect the proficiency of a chimeric nuclease, including the specificity and affinity of ZFPs,[6,27] the length and composition of the interdomain linker,[29,30] the length of spacer sequence DNA in between ZFP binding sites,[29,30] and the interaction between two FCDs.[31,32] We hypothesized that by increasing the catalytic activity of FCD through directed protein evolution, we could improve the performance of ZFNs even if they were constructed using lower-affinity ZFPs. Using a sensitive *in vivo* selection methodology in parallel with cycling mutagenesis, *in vitro* DNA shuffling, and site-directed mutagenesis, we were able to identify an FCD mutant with enhanced catalytic activity relative to that of the wild-type domain, as demonstrated by bacterial genetic assays, *in vitro* DNA cleavage assays, and targeted mammalian genome mutagenesis assays (Figs. 2–4).

Sharkey, the most efficient FCD variant we interrogated, consists of two amino acid mutations, S418P and K441E, which were shown to independently enhance the cleavage capabilities of wild-type FCD. Situated within 8 Å of the catalytic center Lys 469 and within 4 Å of the phosphate backbone of bound substrate DNA (Fig. 6), S418P was observed to increase activity >10-fold relative to that of wild-type FCD. In comparison, K441E increased catalytic performance >5-fold relative to that of wild-type FCD (Fig. 3c). The S418P mutation appears at a turn region in the protein and may contribute to the observed increase in catalytic efficiency by fine-tuning the structure of the enzyme. The introduction of Pro at the onset of helix α3 may influence its structural rigidity and enable rapid substrate turnover and recognition. In addition to the advantageous S418P and K441E mutations, both Q481H and N527D were observed in ≥70% of sequenced FCD variants. However, while Q481H lies within 7 Å of the catalytic center (Fig. 6), introduction of either itself or N527D to FokI diminishes activity between 2- and 8-fold, relative to that of the wild-type cleavage domain. Indeed, it appears that the accumulation of either of these point mutations over the course of our *Sharkey*′ evolutions resulted in preferential activity, specifically toward the MluI core sequence. The ability to refine the substrate specificity of the FCD by identification of mutations that contribute to substrate discrimination may have great utility for various endogenous gene-targeting applications. Saturation mutagenesis of Gln481 and Asn527, two amino acids that may play a role DNA recognition, as well as neighboring amino acids may yield the identification of selective FCD variants
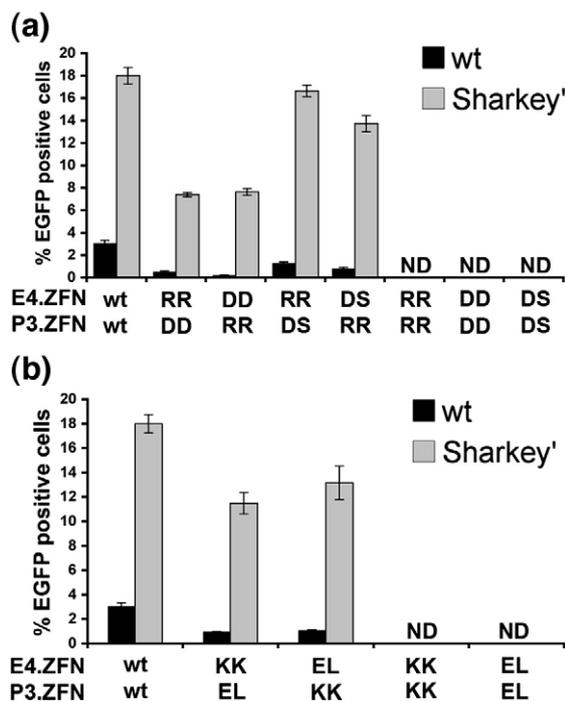

Fig. 5. *Sharkey* is compatible with alternative ZFN architectures. Mutagenesis efficiencies for asymmetric ZFN scaffolds comprising wild-type FCD (black) or *Sharkey* mutations (white). (a) RR::DD, RR::DS and (b) KK::EL FCD architectures were fused to E4 or P3.ZFPs. Stimulated mutagenesis was measured by counting the percentage of EGFP-positive cells. Error bars indicate standard deviation of three replicates.
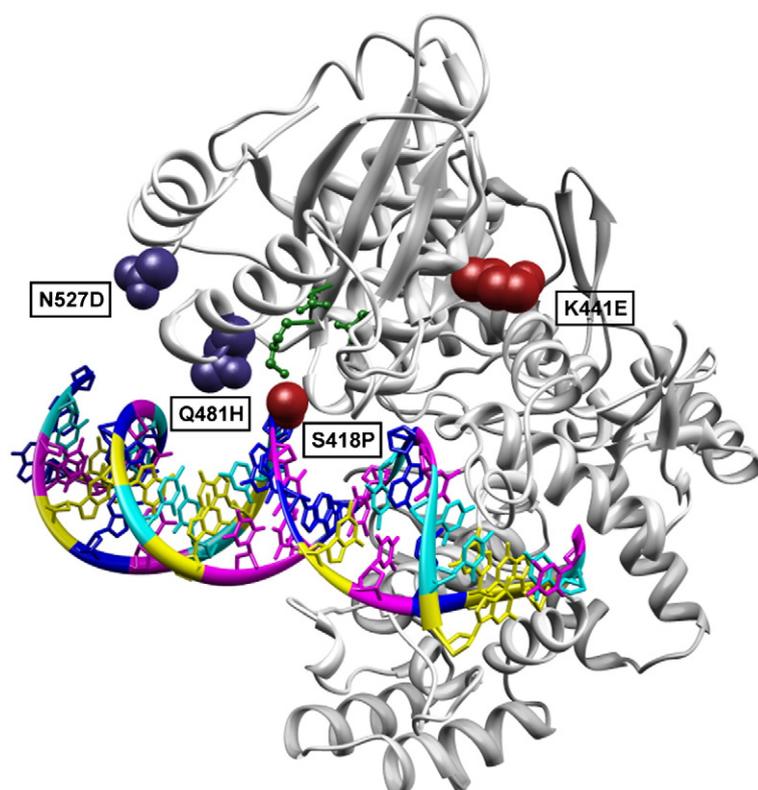
Fig. 6. Selected mutations mapped on the crystal structure of the restriction endonuclease FokI. Mutations identified in ≥70% sequenced variants are highlighted. Activating mutations S418P and K441E are depicted as red spheres on full-length FokI bound to substrate DNA (PDB ID 1FOK).[41] The selected mutations Q481H and N527D are shown as blue spheres. The catalytic center amino acids Asp 450, Asp 467 and Lys 469 are depicted as green balls and sticks.

capable of discriminating between highly homologous endogenous target sites, thus refining overall specificity in conjunction with the ZF domains through cooperative specificity.[42] The development of ZFNs with preferential FCD activities may further reduce off-target cleavage events and, consequently, toxicity. Additionally, over the course of these evolutions, we were able to identify mutations outside the cleavage domain, which may contribute to the activity of ZFNs by affecting the affinity and specificity of custom ZFPs. Our results suggest that our adapted selection system may be used for optimizing the ZF domains of the ZFN as well.

In addition to demonstrating the enhanced cleavage capabilities of *Sharkey* in a variety of contexts, including a mammalian mutagenesis assay, we showed that *Sharkey* mutations are compatible with the ZFN architectures developed by Miller *et al.* and Szczepek *et al.*[31,32] Moreover, toward the goal of improving the efficiency of existing asymmetric ZFN scaffolds, we were successfully able to amend our selection system for the directed evolution of novel heterodimeric architectures. Using saturation mutagenesis to target two critical hydrogen bonds within the dimer interface of FokI, we were able to identify a novel FokI interface, DAMQS, that in conjunction with the D483R ZFN architecture was able to stimulate mutagenesis more efficiently than the RR::DD ZFN scaffolds developed by Szczepek *et al.* We suspect the original interface contained suboptimal hydrogen-bonding networks as a result of the *in silico*-based approach used to generate the asymmetric interface. We believe targeting neighboring amino acids by saturation mutagenesis enabled us to experimentally survey favorable regions of sequence space inaccessible through computational approaches and identify increasingly stable ZFN heterodimers with optimal side-chain configurations.

In summary, with increased catalytic capabilities and compatibility with the promising and potentially powerful heterodimeric ZFN scaffolds, we anticipate the *Sharkey* domain will prove indispensable for a growing list of ZF nuclease related applications throughout biology and medicine.

Methods

Directed evolution of ZFN variants

The expression plasmid pPDAZ was constructed by first removing nucleotides 5–33 from pPROLar.A322 (Clontech Laboratories, Inc) to generate pPROLar.del.ara. This modification abolished arabinose control over protein expression. The zeocin resistance gene was next PCR-amplified from pcDNA3.2/Zeo(−) and cloned into pPROLar.del.ara with AatII and SacI to form pPDAZ. ZFNs were cloned into pPDAZ with the restriction sites KpnI and XbaI. The P3.FN recognition sequence was PCR-amplified using the overlapping primers P3.FN forward (Fwd) and P3.FN reverse (Rev) and cloned into the reporter plasmid p11-LacY (kindly provided by Dr. H. Zhao) to generate p11-LacY-sP3/P3. Additional DNA beyond the ZFN recognition site was included to avoid formation of a target-site hairpin during PCR. Primer sequences are reported in Supplementary Fig. 6. The reporter plasmid p11-LacY-sP3/P3 was transformed into

E. coli BW25141 (kindly provided by Dr. H. Zhao). Electrocompetent cells containing the reporter plasmid p11-LacY-sP3/P3 were prepared using standard protocols.

The FCD was amplified by PCR from the template pcDNA.FCD with the primers FokI Fwd and FokI Rev. The P3 ZFP was amplified by PCR from the template pMal-PBS2.HS1 with the primers ZIF-KpnI Fwd and 3ZF-FokI Rev. Overlap PCR with ZIF-KpnI Fwd and FokI Rev was used to fuse FCD to the P3 ZFP. P3.wt ZFN was cloned into pPDAZ with KpnI and XbaI to generate pPDAZ-P3.FCD. Libraries of ZFN mutants were generated by error-prone PCR. Amplification of FCD was performed over 20 cycles with the template pPDAZ-P3. FCD and the primers ZFN Lib Fwd and Lib Restore Rev in the presence of 12.5 μM dPTP and 12.5 μM 8-oxo-dGTP. An error-free copy of the P3 ZFP was PCR-amplified with ZIF-KpnI Fwd and ZFN Lib Rev. Subsequent overlap PCR with ZIF-KpnI Fwd and Lib Restore Rev was used to fuse cleavage domain (with an average of four amino acid mutations) to the error-free copy of P3 ZFP. The resulting ZFN library was cloned into pPDAZ with KpnI and XbaI and electroporated into *E. coli*. ZFN libraries were routinely composed of 10^7–10^8 members. Following transformation, ZFN containing plasmids were isolated from overnight culture and electroporated into the selection strain BW25141. Transformed cells were recovered in SOC medium at 37 °C for 1 h before plating on solid selection medium containing zeocin (25 ng/mL) and 10 mM arabinose. Survival rate (SR) was calculated as the ratio of the number of colonies on an arabinose plate to that of an arabinose-free plate. Following nine rounds of selection, the start codon ATG was replaced with GTG, and the recovery time following electroporation was increased to 3 h. Subsequent rounds of selection saw an incremental decrease in recovery time.

ZFN purification and *in vitro* DNA cleavage assay

N-terminal His$_6$-tagged P3.wt and P3.*Sharkey*′ were cloned into the p11-LacY-wtx1 expression vector, replacing the Nhe1- and Xba1-flanked *ccdB* gene, and transformed into *E. coli* TOP10F′. Single colonies were picked and grown overnight at 37 °C in SB medium with carbenicillin (100 μg/mL), an ampicillin analogue, and 1% glucose. The overnight culture was then diluted 1:100 in SB medium containing 90 μM ZnCl$_2$ and carbenicillin (100 μg/mL) at 37 °C with shaking until an OD$_{600}$ of 0.4, at which point each culture was incubated at room temperature. At an OD$_{600}$ of 0.6, protein expression was induced with 10 mM arabinose. After 5 h, cells were harvested via centrifugation and sonicated in lysis buffer (50 mM NaH$_2$PO$_4$, 500 mM NaCl, 0.1 mM ZnCl$_2$, 1 mM PMSF, 5 mM DTT, 5 mM imidazole, and 10% glycerol, pH 7.9). Proteins were purified using Ni–NTA agarose resin (Qiagen) and eluted in 50 mM NaH$_2$PO$_4$, 500 mM NaCl, 0.1 mM ZnCl$_2$, 400 mM imidazole, and 10% glycerol (pH 7.9) and were further purified by ion-exchange chromatography with SP Sepharose Fast Flow resin (Amersham Pharmacia Biotech AB). Protein was concentrated and stored at −80 °C in 20 mM NaH$_2$PO$_4$, 50 mM NaCl$_2$, 100 μM ZnCl$_2$, 5 mM DTT, and 50% glycerol (pH 7.4) until use.

The *in vitro* cleavage assay was performed against linearized substrate plasmid pSub-P3. ZFN (12 nM) was added into prewarmed ZFN reaction buffer (20 mM Tris–acetate, 10 mM magnesium acetate, 50 mM potassium acetate, 90 μM ZnCl$_2$, and 5 mM DTT, pH 7.9) containing 4, 6, 12, 24, or 36 nM substrate DNA. Samples were incubated at 37 °C. Aliquots (5–10 μL) were withdrawn at regular time intervals and quenched with 2 μL of stop solution [10 mM EDTA (ethylenediaminetetraacetic acid), 39% glycerol, 0.5% SDS, 0.025% bromophenol blue, and 0.025% xylene cyanol]. ZFN-mediated DNA cleavage was monitored by gel electrophoresis and analyzed by ImageJ.

In vivo ZFN activity assay

P3.FCD variants containing a GTG initiation codon and the mutations S418P, Q481H, N527D, and K441E were generated by Quickchange site-directed mutagenesis PCR. Site-directed mutagenesis was confirmed by DNA sequencing. MluI (ACGGCT), Nx6 (N = A, T, C, or G), ACGAAT, VF2471 (GAGAGT), and CFTR (TGGTGA) containing target sites were amplified by the overlapping synthetic oligonucleotides sPBSm-SceI-X Fwd and sPBSm-SceI-X Rev, where X is the ZFN cleavage site core sequence. ZFN target sites were subsequently cloned into the reporter plasmid p11-LacY-X and transformed into *E. coli* BW25141. Individual colonies containing the reporter plasmid p11-LacY-X were cultured overnight and electrocompetent cells for each ZFN cleavage site were prepared. ZFN expression plasmid was transformed into the appropriate BW25141 selection strain. ZFN activity was measured by calculating the SR at the time point at which wild-type ZFN SR was observed to be 0.5–2%. ZFN variant SR was normalized to wild-type ZFN.

Directed evolution of *Sharkey*′-based ZFN heterodimers

To construct the heterodimer reporter plasmid p11-LacY-sE6/P3, we PCR-amplified an additional expression cassette from pROLar.A322 and cloned it into the NsiI site of p11-LacY-sE6/P3-ΔXbaI. The mutation D483R was introduced into E6.*Sharkey*′ by Quickchange site-directed mutagenesis PCR. The resulting fixed ZFN variant was cloned into the reporter plasmid p11-LacY-sE6/P3-ΔXbaI and transformed into the selection strain BW25141. The FCD variant FCD$_{9-3}$ was used as a starting point for diversity introduction, as *Sharkey* had yet to be identified. Diversity was introduced into FCD$_{9-3}$ at amino acids 483–487 with the oligonucleotides Rd DEMQR Fwd and Rd DEMQR Rev. The Rev primer contained an NNK codon at each of the corresponding targeted amino acids (where N = A, T, C, or G and K = G or T). The NNK codon yields 32 possible codons, which encode 20 amino acids. The ZFN library was subsequently cloned into pPDAZ with KpnI and XbaI and electroporated into *E. coli*. The ZFN library was composed of >10^8 transformants. ZFN containing plasmids were isolated from overnight culture and electroporated into the selection strain BW25141, which contained the E6.FCD$_{9-3}$ D483R expressing reporter plasmid p11-LacY-sE6/P3-ΔXbaI. Active ZFN variants were enriched as previously described. The secondary FCD$_{9-3}$ library was generated with the oligonucleotides Rd DEMQR Fwd and Rd EMQR Rev.

Construction of mammalian cell lines and measurements of mutagenesis

To generate an EGFP reporter gene containing either E6/E6 or E6/P3 target sites, a single SgrAI site was inserted between EGFP residues 157 and 158 using the overlapping primers midEGFP SgrAI Fwd and midEGFP SgrAI Rev. The overlapping oligonucleotides E2C.I-SceI Fwd and E2C.I-SceI Rev were subsequently used to clone into the engineered SgrAI site to generate the E6/E6 EGFP

reporter gene. The overlapping oligonucleotides PBS2/E2C.I-SceI Fwd and E2C.I-SceI Rev were used to clone into SgrAI to generate the E6/P3 EGFP reporter gene. Reporter cell lines containing a single CMV promoter and a modified EGFP reporter gene were generated in Flp-In-293 cells using the Flp-In system (Invitrogen).

Wild-type FCD and *Sharkey*′ were subcloned into the ZFN expression plasmid pVAX1 (Invitrogen) using the restriction enzymes PstI and XbaI. E3, E4, E5, and E6 ZFPs were constructed as previously described[21] and subcloned into ZFN expression plasmid using the restriction enzymes XmaI and AgeI. To measure the rates of stimulated mutagenesis, we seeded reporter cells onto polylysine-coated 24-well plates at a density of 1.5×10^5 per well. After 24-h incubation, reporter cells were cotransfected with 100 ng of ZFN expression plasmid and 500 ng of pcDNA3.1/Zeo(−) using Lipofectamine 2000 (Invitrogen) under conditions specified by the manufacturer. Similarly, for heterodimer ZFN assays, cells were cotransfected with 100 ng of ZFN expression plasmid and 400 ng of pcDNA3.1/Zeo(−) carrier DNA. Transfection efficiencies were measured to be between 70% and 80%. Three days after transfection, 30,000 cells were analyzed by flow cytometry (FACScan Dual Laser Flow Cytometer, BD Biosciences) to measure the percentage of EGFP-positive cells. Additionally, the rate of mutagenesis was measured by MluI cleavage. Briefly, 3 days after transfection; genomic DNA was harvested and purified with a QIAmp DNA Mini Kit (Qiagen). Modified EGFP gene was amplified over 30 cycles by PCR (Expand High Fidelity, Roche) with 1 μg of template DNA and 10% (v/v) dimethyl sulfoxide with an annealing temperature of 72 °C. The PCR products were digested overnight with MluI and visualized by gel electrophoresis.

γ-H2AX-based cytotoxicity assay

γ-H2AX cytotoxicity assay is described in Supplementary Information.

Accession numbers

The nucleic acid sequence of the *Sharkey* cleavage domain has been deposited in GenBank with accession number HM130522.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2010.04.060

References

1. Kim, Y. G., Cha, J. & Chandrasegaran, S. (1996). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl Acad. Sci. USA*, **93**, 1156–1160.
2. Bibikova, M., Golic, M., Golic, K. G. & Carroll, D. (2002). Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. *Genetics*, **161**, 1169–1175.
3. Doyon, Y., McCammon, J. M., Miller, J. C., Faraji, F., Ngo, C., Katibah, G. E. *et al.* (2008). Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nat. Biotechnol.* **26**, 702–708.
4. Meng, X., Noyes, M. B., Zhu, L. J., Lawson, N. D. & Wolfe, S. A. (2008). Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nat. Biotechnol.* **26**, 695–701.
5. Perez, E. E., Wang, J., Miller, J. C., Jouvenot, Y., Kim, K. A., Liu, O. *et al.* (2008). Establishment of HIV-1 resistance in CD4$^+$ T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.* **26**, 808–816.
6. Urnov, F. D., Miller, J. C., Lee, Y. L., Beausejour, C. M., Rock, J. M., Augustus, S. *et al.* (2005). Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*, **435**, 646–651.
7. Wright, D. A., Townsend, J. A., Winfrey, R. J., Jr, Irwin, P. A., Rajagopal, J., Lonosky, P. M. *et al.* (2005). High-frequency homologous recombination in plants mediated by zinc-finger nucleases. *Plant J.* **44**, 693–705.
8. Beumer, K., Bhattacharyya, G., Bibikova, M., Trautman, J. K. & Carroll, D. (2006). Efficient gene targeting in *Drosophila* with zinc-finger nucleases. *Genetics*, **172**, 2391–2403.
9. Carroll, D., Beumer, K. J., Morton, J. J., Bozas, A. & Trautman, J. K. (2008). Gene targeting in *Drosophila* and *Caenorhabditis elegans* with zinc-finger nucleases. *Methods Mol. Biol.* **435**, 63–77.
10. Geurts, A. M., Cost, G. J., Freyvert, Y., Zeitler, B., Miller, J. C., Choi, V. M. *et al.* (2009). Knockout rats via embryo microinjection of zinc-finger nucleases. *Science*, **325**, 433.
11. Porteus, M. H. & Baltimore, D. (2003). Chimeric nucleases stimulate gene targeting in human cells. *Science*, **300**, 763.
12. Zeevi, V., Tovkach, A. & Tzfira, T. (2008). Increasing cloning possibilities using artificial zinc finger nucleases. *Proc. Natl Acad. Sci. USA*, **105**, 12785–12790.
13. Beerli, R. R. & Barbas, C. F., III (2002). Engineering polydactyl zinc-finger transcription factors. *Nat. Biotechnol.* **20**, 135–141.
14. Beerli, R. R., Segal, D. J., Dreier, B. & Barbas, C. F., III (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.
15. Liu, Q., Segal, D. J., Ghiara, J. B. & Barbas, C. F., III (1997). Design of polydactyl zinc-finger proteins for unique addressing within complex genomes. *Proc. Natl Acad. Sci. USA*, **94**, 5525–5530.
16. Gonzalez, B., Schwimmer, L. J., Fuller, R. P., Ye, Y., Asawapornmongok, L. & Barbas, C. F. (2010). Modular system for the construction of zinc-finger libraries and proteins. *Nat. Protoc.* **5**, 791–810.
17. Segal, D. J., Dreier, B., Beerli, R. R. & Barbas, C. F., III (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.

18. Dreier, B., Segal, D. J. & Barbas, C. F., III (2000). Insights into the molecular recognition of the 5′-GNN-3′ family of DNA sequences by zinc finger domains. *J. Mol. Biol.* **303**, 489–502.

19. Dreier, B., Beerli, R. R., Segal, D. J., Flippin, J. D. & Barbas, C. F., III (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.

20. Dreier, B., Fuller, R. P., Segal, D. J., Lund, C. V., Blancafort, P., Huber, A. *et al.* (2005). Development of zinc finger domains for recognition of the 5′-CNN-3′ family DNA sequences and their use in the construction of artificial transcription factors. *J. Biol. Chem.* **280**, 35588–35597.

21. Mandell, J. G. & Barbas, C. F., III (2006). Zinc Finger Tools: custom DNA-binding domains for transcription factors and nucleases. *Nucleic Acids Res.* **34**, W516–W523.

22. Liu, Q., Xia, Z., Zhong, X. & Case, C. C. (2002). Validated zinc finger protein designs for all 16 GNN DNA triplet targets. *J. Biol. Chem.* **277**, 3850–3856.

23. Bae, K. H., Kwon, Y. D., Shin, H. C., Hwang, M. S., Ryu, E. H., Park, K. S. *et al.* (2003). Human zinc fingers as building blocks in the construction of artificial transcription factors. *Nat. Biotechnol.* **21**, 275–280.

24. Sander, J. D., Zaback, P., Joung, J. K., Voytas, D. F. & 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.

25. Maeder, M. L., Thibodeau-Beganny, S., Osiak, A., Wright, D. A., Anthony, R. M., Eichtinger, M. *et al.* (2008). Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol. Cell*, **31**, 294–301.

26. Foley, J. E., Yeh, J. R., Maeder, M. L., Reyon, D., Sander, J. D., Peterson, R. T. & Joung, J. K. (2009). Rapid mutation of endogenous zebrafish genes using zinc finger nucleases made by Oligomerized Pool ENgineering (OPEN). *PLoS ONE*, **4**, e4348.

27. Cornu, T. I., Thibodeau-Beganny, S., Guhl, E., Alwin, S., Eichtinger, M., Joung, J. K. & Cathomen, T. (2008). DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. *Mol. Ther.* **16**, 352–358.

28. Smith, J., Bibikova, M., Whitby, F. G., Reddy, A. R., Chandrasegaran, S. & Carroll, D. (2000). Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic Acids Res.* **28**, 3361–3369.

29. Bibikova, M., Carroll, D., Segal, D. J., Trautman, J. K., Smith, J., Kim, Y. G. & Chandrasegaran, S. (2001). Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol. Cell. Biol.* **21**, 289–297.

30. Handel, E. M., Alwin, S. & Cathomen, T. (2009). Expanding or restricting the target site repertoire of zinc-finger nucleases: the inter-domain linker as a major determinant of target site selectivity. *Mol. Ther.* **17**, 104–111.

31. Miller, J. C., Holmes, M. C., Wang, J., Guschin, D. Y., Lee, Y. L. & Rupniewski, I. (2007). An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat. Biotechnol.* **25**, 778–785.

32. Szczepek, M., Brondani, V., Büchel, J., Serrano, L., Segal, D. J. & Cathomen, T. (2007). Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat. Biotechnol.* **25**, 786–793.

33. Chen, Z. & Zhao, H. (2005). A highly sensitive selection method for directed evolution of homing endonucleases. *Nucleic Acids Res.* **33**, e154.

34. Bahassi, E. M., Salmon, M. A., Van Melderen, L., Bernard, P. & Couturier, M. (1995). F plasmid CcdB killer protein: *ccdB* gene mutants coding for non-cytotoxic proteins which retain their regulatory functions. *Mol. Microbiol.* **15**, 1031–1037.

35. Kowalczykowski, S. C., Dixon, D. A., Eggleston, A. K., Lauder, S. D. & Rehrauer, W. M. (1994). Biochemistry of homologous recombination in *Escherichia coli*. *Microbiol. Rev.* **58**, 401–465.

36. Stemmer, W. P. (1994). Rapid evolution of a protein in vitro by DNA shuffling. *Nature*, **370**, 389–391.

37. O'Donnell, S. M. & Janssen, G. R. (2001). The initiation codon affects ribosome binding and translational efficiency in *Escherichia coli* of *cI* mRNA with or without the 5′ untranslated leader. *J. Bacteriol.* **183**, 1277–1283.

38. Valerie, K. & Povirk, L. F. (2003). Regulation and mechanisms of mammalian double-strand break repair. *Oncogene*, **22**, 5792–5812.

39. Abedi, M., Caponigro, G. & Kamb, A. (1998). Green fluorescent protein as a scaffold for intracellular presentation of peptides. *Nucleic Acids Res.* **26**, 623–630.

40. Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S. & Bonner, W. M. (1998). DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J. Biol. Chem.* **273**, 5858–5868.

41. Wah, D. A., Hirsch, J. A., Dorner, L. F., Schildkraut, I. & Aggarwai, A. K. (1997). Structure of the multi-modular endonuclease *Fok*I bound to DNA. *Nature*, **388**, 97–100.

42. Gordley, R. M, Gersbach, C. A. & Barbas, C. F., III (2009). Synthesis of programmable integrases. *Proc. Natl Acad. Sci. USA*, **106**, 5053–5058.